THE UNIVERSITY OF SHEFFIELD

DOCTORAL THESIS

Investigations of the dynamics and mechanism of β-phosphoglucomutase

Author:
Angus ROBERTSON

Supervisor:
Prof. Jon WALTHO

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Biomolecular NMR group
Department Molecular Biology and Biotechnology

December 11, 2018

Declaration of Authorship

I, Angus ROBERTSON, declare that this thesis titled, "Investigations of the dynamics and mechanism of β-phosphoglucomutase" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Acknowledgements

To all that enjoy the active first-person writing voice ... this is all that there is in this whole book, so make the most of it.

While I was supposed to learn how to be a scientist over the course of my PhD, I certainly feel that I have learned so much more than that [cue self aggrandizing soliloquy about what it is to be a scientist]. But seriously, I would like to thank my supervisor **Jon W**, not just for the opportunity to do this PhD, but for engaging on such a broad range of topics that now I start to see just how much there is to know. And also for the near hundred vivas over the last four years (conveniently disguised as project meetings) while we were trying to figure out how this damned enzyme really works. Thank you for giving me the opportunity to pursue things that I found interesting and trusting me to take responsibility for my work.

On the technical side, thank you to **Andrea H** and **Matt C** for always fitting me in for yet another "final experiment", even when it was the fifth or sixth iteration. Thank you to both for general advice about science, lots of laughter, and looking the other way when my RTSG was beyond empty [possibly ... it depends who's reading]. Thank you to Matt for all of the computer geekiness and the occasionally intelligible README files. On the topic of computers, **Alex W**, thank you for running the endless calculations and bearing with us biologists while we repeatedly and repeatedly reframed the question that we were asking, and to **Andy A** for engaging with yet another βPGM simulation. Thank you to **Mike W** for answering NMR related questions when I had them, for excellent introductions, for fielding questions no matter how simple, and for always giving your insight into the research going on in the lab.

To all of the D10 lab members, it has been a pleasure, particularly being able to see such a wide range of science going on under one roof. It was always nice to chat with **Rosie S**, **Pete D**, or **Alex T** about science, maths, and scientific politics, I always left having learned something about how it all "really" works no matter what it really was. Needless to say thank you all for the laughter too ... I think that this might be a recurring theme .... To **Nicky B** and **Clare T**, you two have been everything that a PhD student needed, kind, supportive, patient, and incredibly smart! Thank you for the Felix macros and the time explaining them all at the beginning (and for keeping your magic lists updated!), but more than that, for instilling the patience to be both precise and accurate and for reading drafts of papers and a draft of this thesis (**Clare T**). Talking of patience, thank you to **Luke J** for showing me the ropes in the lab (and **Abi W**), especially when my wet-lab technique was rather rusty at the beginning, at least I didn't set fire to the lab! Speaking of ... **Henry W**, it has been a genuine pleasure passing the βPGM tips and tricks on to you, the project is certainly in safe hands ... well, as long as they're wearing a lab coat.

To the crystallographers. Thank you for taking me under your wing and using my crystals to "test the beam line", the marriage of X-ray structure and solution NMR data made the questions that we could answer so much more interesting (at least to those who are interested in βPGM). Thank you to **Dave R** and **Pat B** for some great crystallography discussions and to **Jon R** and **Jeremy C** for pointing me in the right direction on several occasions. To **Sam D** , **Alicia CA** and **Adli A**, I could apologize for the distraction caused over the last couple of years, but I'm not going to. Helpful discussions and technical assistance doesn't quite cover it, you lot have been ace! Speaking of ace, **Claudine B**, thank you for the time and patience

showing me the ropes of crystallography, the immaculate looping, the belaying, and all of the laughter. I'm nearly a crystallographer now, nearly...

To long suffering **Liam A**, I'm sorry for all of the word play. If you give it a second, I think that you'll reckon, and be sad that I'm going away. That and puns ... wayyy too many puns. But seriously, **Liam A**, **Helen R** and **Henry W**, it was a great house to share and I'm really gonna miss you guys, well ... maybe not Henry. To **Daria S**, **Claudine B**, **Liam A** and **Cath H**, thank you for making the snowboarding breaks happen, and then so much fun! Still no broken bones, although some bruised knees, (grapefruit) and pride. Thanks to such an awesome MBB PhD soc. committee too, although it seems like so long ago! **James T** I miss the hugs, and the dream duo of **Dani M** and **Cath H** always managed to make organizing fun! As for fun, thanks to **Nate A** (fun personified) for all of the chats about science and otherwise, and especially for encouraging me to engage with out-reach, and for helping me to better explain what I do.

Så här går det ... Tack till **Mikael A** och **Göran C** för att ni lärde mig en hel del svår NMR. Tack för programmen, avhandlingarna och diskussionerna, jag lärde mig så mycket medan jag var där, jag känner mig nästan som en riktig NMR-spektroskopist. Till **Johan W**, tack för cykeln och för 'the inside scoop' på distans, och **Uli W**, tack för ditt tålamod med βPGM. **Olof O**, min kära Olof, vad skulle jag ha gjort utan dig? **Olof O**, **Sven W**, **Filip P** och **Sam B** tack för skrattet, luncherna, ölerna, vetenskapen, men mestadels för att få att ni fick mig att känner mig så välkommen. Ni kommer alltid ha en skål vid mitt bord! Till alla de andra underbara människorna på CMPS, **Mattias W**, **Baxter B**, **CJ H**, **Mattis T**, **Viktoria B**, **Veronica R**, **Rebecca F** och alla ansikten utan namn (på mitt huvud), tack för det roliga och må all er vetenskap publiceras! Slutligen, tack Olivia för att kontrollera denna översättning, tack för uppmuntran och många kloka ord!

And my family, thanks for giving birth to me [ref. birth certificate], and raising me [ref. grocery bills], and sending me to school [ref. school fee debt]. It definitely helped...

THE UNIVERSITY OF SHEFFIELD

Abstract

Faculty of Science

Department Molecular Biology and Biotechnology

Doctor of Philosophy

Investigations of the dynamics and mechanism of β-phosphoglucomutase

by Angus ROBERTSON

This thesis uses a multidisciplinary approach of nuclear magnetic resonance (NMR) spectroscopy, X-ray crystallography, and enzyme kinetics to further investigate how β-Phosphoglucomutase (βPGM; EC 5.4.2.6) , an archetypal phosphoryl transfer enzyme from the HAD superfamily, catalyses the inter-conversion of β-glucose 1-phosphate (βG1P) with glucose 6-phosphate (G6P) via a β-glucose 1,6-phosphate (βG16P) intermediate. The use of metal fluorides to mimic positions along the reaction coordinate of phosphoryl transfer enzymes has been well established and allows for a critical investigation of the the role of enzyme dynamics, electrostatics, conformation, and intrinsic organization of the enzyme in catalysis.

In a series of papers, this thesis demonstrates several elements of how βPGM has evolved to perform its function. Firstly, mutation of the enzymatic general acid-base (GAB) allowed the investigation of native substrate in the active site of the enzyme. This ground state model was closed around the substrate, with transferring phosphate and nucleophile in van der Waals contact, but without overall transition state architecture. Furthermore, a weakened magnesium affinity in this ground state suggests a mechanism for dissociation of such a high affinity ligand, essential for efficient catalysis. Secondly, using the same GAB mutation, the role of proton transfer in phosphoryl transfer reactions is investigated in pre- and post- proton transfer models. Using a combination of NMR, X-ray crystallography and DFT calculation, it is determined that the proton transfer event is not synchronous with phosphorous transfer, and several key themes are elucidated; before, during, and after the chemical transfer. Each of which contribute to the capacity of βPGM to break and form phosphate monoester bonds on a viable timescale. Thirdly, a mechanism is presented to explain a previously modelled enzymatic lag phase prior to steady state catalysis. Mutation of a key arginine residue is sufficient to alleviate this lag phase and does not perturb the chemical step of the reaction which indicates that such perturbations are not transmitted through substrate to the catalytic center. Finally, it is observed that the phospho-enzyme state of βPGM, when compared to a specific phosphatase (phosphoserine phosphatase (PSP)), displays several features in order to stabilize the phospho-enzyme state that are not present in PSP. Together these features further describe how βPGM has evolved both specificity and to achieve high levels of catalytic rate enhancement.

Contents

List of Figures

List of Tables

List of Abbreviations

NMR	Nuclear Magnetic Resonance
DFT	Density Functional Theory
QM	Quantum Mechanics
EVB	Empirical Valence Bond
βPGM	β-Phosphoglucomutase
PSP	Phosphoserine Phosphatase
DHFR	Dihydrofolate Reductase
G6PDH	Glucose 6-phosphate dehydrogenase
HADSF	Haloacid Dehydrogenase Superfamily
WT	Wild Type
PDB	Protein Data Bank (usually in reference to accession code)
TS	Transition State
TSA	Transition State Analog
GS	Ground State
GSA	Ground State Analog
NAC	Near Attack Complex
GAB	General Acid-Base
βG1P	β-Glucose 1-Phosphate
G6P	Glucose 6-Phosphate
βG16BP	β-Glucose 1,6-bisphosphate.
AcP	Acetylphosphate
αG16BP	α-Glucose 1-Phosphate
NAD	Nicotinamide adenine dinucleotide (oxidized form)
NADP	Nicotinamide adenine dinucleotide phosphate (oxidized form)
NADH	Nicotinamide adenine dinucleotide (reduced form)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
Pi	Inorganic phosphate
NOE	Nuclear Overhauser Effect
FID	Free Induction Decay
CSA	Chemical Shift Anisotropy
TROSY	Transverse Relaxation Optimized Spectroscopy
HSQC	Heteronuclear Single Quantum Coherence
MF	Model Free
CPMG	Carr Purcell Meiboom Gill
PEG	polyethylene glycol
TRIS	tris(hydroxymethyl)aminomethane
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
SDS	Sodium Dodecyl Sulphate

PAGE	Polyacrylamide Gel Electrophoresis
DEAE	Diethylaminoethyl
LB	Luri Bertani
SDM	Site Directed Muatgenesis
PCR	Polymerase Chain Reaction
IPTG	Isopropyl β-D-1-thiogalactopyranoside
EDTA	Ethylenediaminetetraacetic acid
MWCO	Molecular Weight Cut Off
TEMED	Tetramethylethyldiamine
APS	Ammonium Persulphate
DTT	Dithiothreitol
TSP	Trimethylsilylpropanoic acid

List of Symbols

k_{cat}	The catalytic rate constant for a reaction	(s^{-1})
k_{obs}	The observed rate constant for a reaction	(s^{-1})
κ	The transmission coefficient for a reaction.	
K_m	The Michaelis constant for a reaction	
k_{ex}	The exchange rate for a chemical exchange process	Hz
G	Gibbs free energy	XXX
ΔG	change in Gibbs free energy	
ΔH	change in enthalpy	
ΔS	change in entropy	
T	Temperature	K
R	Gas constant	8.314 ($J\,K^{-1}\,mol^{-1}$)
h	Planck's constant	6.626 x 10^{-34} ($m^2\,kg\,s^{-1}$)
$\hbar$	h $/2\pi$	1.055 x 10^{-34} ($m^2\,kg\,s^{-1}$)
k_B	Boltzmann's constant	1.381 x 10^{-23} ($m^2\,kg\,s^{-2}\,K^{-1}$)
$\mathbf{I}$	nuclear spin angular momentum	
I	nuclear spin angular momentum quantum number	
m	magnetic quantum number	
μ	nuclear magnetic moment	
$\mathbf{B}$	Magnetic field vector	
γ	the gyromagnetic ratio of a nucleus	
B_0	Strength of static magnetic field	T
ω_0	Larmor frequency of a nucleus	
σ	isotropic shielding	
η	asymmetry of isotropic shielding	
δ	chemical shift	
p_A	population of species A	
p_B	population of species B	
$j(\omega)$	the spectral density function a frequency ω	
$\sigma_{\|\|}$	the shielding tensor parallel to the principal axis	
$\sigma_{\perp}$	the shielding tensor perpendicular to the principal axis	
μ_0	the magnetic permeability of free space	
τ_c	global correlation time of a molecule	(rad s^{-1})
τ_e	local correlation time of a molecule	(rad s^{-1})
D^R	the overall tumbling coefficient	
$G(\tau)$	the correlation function of a molecule	
$C(\tau)$	the reduced correlation function of a molecule	
R_1	Longitudinal relaxation rate	

$R_{1\rho}$	Rotating frame relaxation rate	
R_2	Transverse relaxation rate	
T_1	Longitudinal relaxation time	
$T_{1\rho}$	Rotating frame relaxation time	
T_2	Transverse relaxation time	
S^2	Modelfree order parameter	
ϵ	the extinction coefficient for βPGM at 280 nm	19940 $\mathrm{L\,mol^{-1}\,cm^{-1}}$

Publications and Manuscripts

- Paper I: Van der Waals contact between nucleophile and transferring phosphorus is insufficient to achieve enzyme transition state architecture (Johnson et al., 2018)
- Paper II: X-ray, NMR and QM approaches reveal the relationship between protein conformational change, proton transfer, and phosphoryl transfer in an archetypal enzyme (manuscript)
- Paper III: Arg - phosphate interaction in β-phosphoglucomutase improves substrate affinity, but introduces inhibition (manuscript)
- Paper IV: Mechanisms of phosphatase activity in good and bad phosphatases of the HAD superfamily (manuscript)

Contribution to publications

- Paper I: I expressed and purified protein, I performed most of the crystallography and some of the NMR, I analysed and interpreted the data and designed further experiments, I wrote processing scripts for analysis of kinetic data, I took part in writing of the manuscript alongside NJB, CRT, and JPW.
- Paper II: I expressed and purified protein with different isotope enrichment schemes and enzymatically synthesized the U[^{13}C]C1-G6P for the NMR experiments. I performed the crystallography, NMR spectroscopy (including backbone assignment) and backbone model free analysis. I analysed and interpreted the data, I wrote programs to analyse and interpret chemical shift perturbations and display the output. I wrote the manuscript together with AW and JPW.
- Paper III: I performed much of the crystallography, I designed the experiments and analysed and interpreted the NMR data, I wrote the manuscript with contributions from HPW and JPW.
- Paper IV: I expressed and purified proteins with different isotope enrichment schemes. I performed the crystallography on βPGM with some technical assistance from CB. I performed the NMR experiments on βPGM and I analysed and interpreted the acquired data. I wrote the manuscript with early contributions from NJB and JPW.

Chapter 1

Introduction

1.1 Foreword - Why bother?

On reading the research directions for most bioscience funding bodies in the UK, a key theme recurs with subtle variations, and with different voicings. The theme is to try and keep human beings alive for as long as possible, with the highest quality of life, for the minimal amount of effort. Necessarily these mandates will attract academic researchers to work on the human condition, nutrition, and preventing infectious diseases, often (although to a lessening extent) in a highly reductionistic manner. This pursuit lends itself well to applied science, to screening, to brute force investigations, often (because of the nature of scientific research) in a way that promotes speed over quality. Where in all of this, is the time for basic research? In fully understanding how an enzyme functions and drawing conclusions about how enzymes have evolved to function?

Our research on β-phosphoglucomutase (βPGM) from *L. lactis* has essentially a twofold implication. The first prong points towards developing a fundamental understanding of enzyme function in an enzyme that performs a very specific, but complicated task, balancing enormous catalytic rate enhancements (*ca.* 10^{21}) of at least 3 different substrates while preventing inhibition by a high affinity intermediate. The hope is that this understanding will feed into biotechnological applications and enzyme design. The second prong of the investigation pertains to the further development of metal fluoride based transition state analogues to investigate phosphoryl transfer enzymes. Fluorine is a highly sensitive spin half nucleus with a huge chemical shift range and a high sensitivity relative to proton. This sensitivity serves as a great training dataset for computational chemistry with subsequent quantum mechanical (QM) and density functional theory (DFT) calculations permitting the investigation of key interactions at the atomic level. This makes ^{19}F NMR an incredibly strong candidate for screening drug molecules against phosphoryl transfer / kinase targets, with a highly sensitive and informative readout. It would be very exciting to see this approach taken more widely, and at the outset, it appears to be an extremely powerful tool.

1.2 Introduction

This text will not attempt to re-define enzymology, catalysis, and protein biochemistry, as there are several excellent texts available to that end (Jencks, 1969; Lodish et al., 2007; Voet and Voet, 2010; Williamson, 2012; Berg, Tymoczko, and Stryer, 2012). This text will however highlight some key themes necessary for engagement with the results that follow.

1.2.1 Free Energy and Catalysis

Chemical reactions principally involve the movement of atoms with concomitant redistribution of electron density and reconfiguration of electronic bonding orbitals. Product states are generally more energetically favourable than reactant states, while both are referred to as ground states (GSs), compared to the transition state (TS) of a reaction which is defined as least energetically favourable state. The intermediate states (where applicable) provide an alternate GS which lies between the product/reactant and transition state in terms of free energy (Fig. 1.1 A,B). These states are typically associated with a free energy and population at equilibrium (Eq. 1.1, 1.2, respectively). In these classic equations, R is the universal gas constant, ΔG is the free energy change, ΔH is the enthalpy change, T is the temperature (in K), and ΔS is the entropy change, and K_{eq} is the equilibrium constant of a system.

$$\Delta G = \Delta H - T\Delta S \tag{1.1}$$

$$\Delta G = -RTln(K_{eq}) \tag{1.2}$$

Enzymes act to lower the activation energy barrier by stabilizing the transition state of the reaction, without perturbing the equilibrium population of reactants and products (Pauling, 1948; Jencks, 1969). Monoclonal antibodies raised against transition state analogues (TSAs) of chemical reactions do catalyze the reaction (Tramontano, Janda, and Lerner, 1986), however, the rate enhancement is often much less than the corresponding enzyme (Hilvert, 2000). Bi-bi ping-pong reaction schemes (such as the reaction catalyzed by βPGM) display several substrate/intermediate bound states across the reaction trajectory, usually with two active states of the enzyme, each with different substrate affinities (Fig. 1.1).

1.2.2 Transition State Theory

The time for the chemical reaction step in an enzyme is of the order of a bond vibration (ca. 10^{-13} s) thus, it is difficult to either observe or structurally characterize transition states. The transition state of a reaction can be described using the general form of the Eyring Polyani equation (Eq. 1.3) where the the transmission coefficient κ plays a role in limiting the overall rate of transfer.

$$k = \kappa \frac{k_B T}{h} e^{-\frac{\Delta G^{\ddagger}}{RT}} \tag{1.3}$$

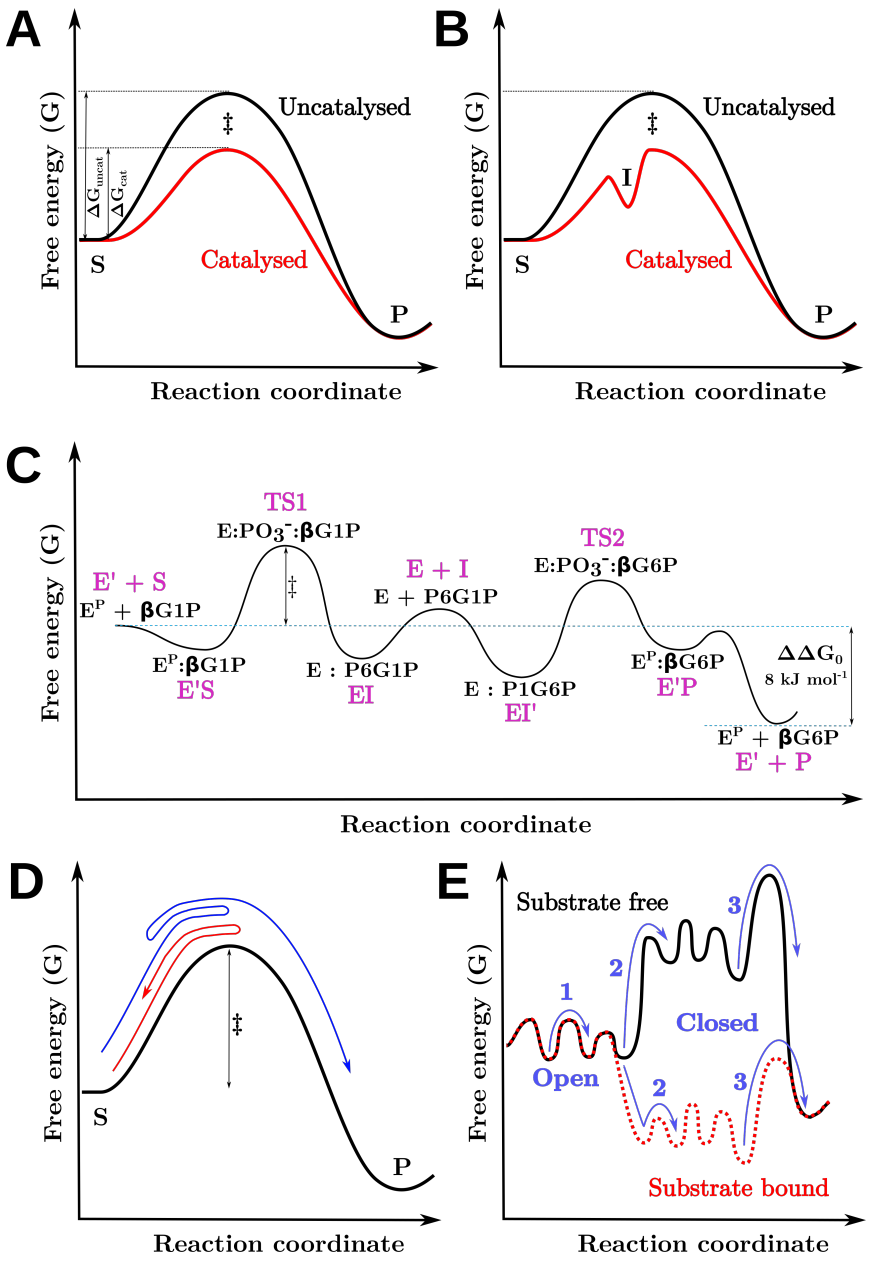


FIGURE 1.1: Gibbs free energy diagrams illustrating the reaction of substrate (S) reacting with enzyme (E) product (P). **A)** Represents the free energy of an uncatalyzed reaction (ΔG_{uncat}) vs. the free energy of a catalyzed reaction (ΔG_{cat}), here the effect of introducing a catalyst is to lower the activation barrier, rather than change the energy of the two resting states. **B)** Indicates the presence of an intermediate state **I** which is lower in free energy than the transition state, but higher than the two resting states. One important role of catalysts is to ensure that these intermediate states are not so stable that they hinder efficient catalysis. **C)** Illustrates the free energy profile of the reaction catalyzed by βPGM with relative energetic values given in (Jin et al., 2014). In this case E' denotes the phosphorylated enzyme E'S denotes the Michaelis complex for the step 1 reaction, and TS1 denotes the transition state for the step 1 reaction. EI is the ground state complex with the βG16BP intermediate, E + I denotes the state where the intermediate has dissociated from the enzyme. EI' indicates the Michaelis complex where the intermediate has reoriented to start the step 2 reaction, TS2 is the TS for the step 2 reaction, and E'P denotes the product ground state where phospho-enzyme has been regenerated. E'+P indicates the dissociated product state with the enzyme re-primed for catalysis with and overall free energy change of the reaction at *ca.* 8 kJ mol^{-1}. **D)** Illustrates the transmission coefficient κ from Eq. 1.3, where not every molecule with sufficient activation energy crosses the energy barrier, instead only a certain proportion of molecules do. **E)** Illustrates the protein conformational selection model using a free energy landscape. Transition 1 has the same small free energy barrier in both substrate free (black) and bound (red) states, so it will likely happen quickly. However, transition 2 is much more energetically favourable if the ligand is present, than if it is absent, thus the transition from an "open" to a "closed" conformation has a large activation energy which is significantly reduced if ligand is present in the active site. If transition 3 was to represent the formation of product, while this step is possible in the "open" state, the activation energy of such a transition likely restricts the chemical step to the "closed" conformer.

Here k is the rate constant, κ is the transmission coefficient, k_B is Boltzmann's constant, T is the temperature (in K), h is Planck's constant, R is the universal gas constant, and $\Delta G^{\ddagger}$ is the free energy change of the transition state. This can be conceptualised using Fig. 1.1 where a single molecule does not always proceed to the product stage of the reaction despite having sufficient energy to do so. However, the observed rate of catalysis (k_{obs}) for most enzyme catalysed reactions is typically of the order 1-1000 s^{-1}, which is often much smaller than the 10^{13} s^{-1} bond vibrations associated with chemical group transfer. An often cited reason for the disparity in rates is the need for the protein to undergo a conformational rearrangement in order to release product/intermediate from the active site. These conformational changes typically occur on this ms-s timescale (Bae and Phillips, 2006) but this raises the important question of what processes occur on what timescales in catalysis (Fig. 1.2).

1.3 Enzyme catalysis

While Fischer's classic "lock and key" model (Fischer, 1909) was a useful initial way to visualize the 3D complementarity between enzyme and substrate in the active site, there were several limitations to this model. Firstly, water exclusion from the active site upon ligand binding (and the entropic contribution therein) was not accounted for, nor the capacity for "induced fit" or "conformational selection" models of ligand binding where lock and key would not perfectly fit together, but could be distorted in order to accommodate one another. These models initially did not account for the dynamic nature of enzymes, from picosecond to >second, which was worsened by a the classic one-fold - one-structure dogma of early structural biology. While it is still the case that on the timescale of a bond vibration necessary for chemical transfer - the protein is static, resulting in electrostatics dominating the chemical step of catalysis. It is not necessarily the case that there is only one pathway to the transition state, rather, there are likely multiple parallel pathways stochastically explored by substrate across a free energy landscape.

Modern structural biology approaches the protein (and ligand) as a dynamic entity across a range of timescales (Fig. 1.2), with the "structure" best represented as a conformational ensemble with variable occupancies and associated degree of "order". Depending on the timescale of the process concerned, from ligand binding, to domain swap dimerization, the population of protein conformations is going to be driven by different intrinsic processes (Fig. 1.2). Fortunately, many of these timescales are accessible to NMR spectroscopy. Three specific examples are relevant to this text, namely the characterization of dihydrofolate reductase (DHFR), Galectin-3 (GAL3), and Abl kinase, as well as triose phosphate isomerase (TIM), glycerol 3-phosphate dehydrogenase (GPDH), and orotidine 5'-monophosphate decarboxylase (OMPDC) from the Richardson lab. Each highlights key themes that will appear in this work, for DHFR it is the dynamic characterization of the enzyme ligand complex and its relationship to observed catalytic rates. For GAL3, the interest lies in the entropy-enthalpy trade-off in ligand binding, while for DHFR, it is the protein energy landscape characterization that is particularly relevant. Work from the Richardson lab is particularly relevant when considering the binding of phosphate to the guanidinium group of arginine sidechains, as this has been investigated in several systems.

DHFR demonstrates a range of conformational dynamics in both open and substrate bound states (Schnell, Dyson, and Wright, 2004). The authors report that slow exchange between

two species in the substrate-free state occludes the binding pocket and presents with different ligand affinities. Upon substrate binding, a loop dynamic occurred at approximately the observed rate of catalysis. When mutations were made to this loop, the rate of catalysis was significantly perturbed. This loop when characterized using the Lipari-Szabo (LS) modelfree (MF) formalism indicates that the loop acts as a flexible gate, helping to stabilize the substrate in a geometry conducive to chemical transfer. Furthermore, the authors report explicit dynamic fluctuations for residues in the active site that occur at the rate of the chemical step in catalysis. Taken together, this extensive characterization of multiple steps in the reaction cycle with dynamic information corroborated using multiple techniques highlights some of the key themes that proteins may use to optimize the energy landscape for catalysis.

The characterization of ligand binding to GAL3 highlights an often overlooked concept in ligand design, namely, the role of protein conformational entropy (Diehl et al., 2010). The authors use a combination of NMR spectroscopy, isothermal titration calorimetry, and X-ray crystallography, to investigate the binding of 3 different ligands to GAL3. The authors conclude that "The estimated change in conformational entropy is comparable in magnitude to the binding enthalpy, demonstrating that it contributes favourably and significantly to ligand binding." Furthermore, the authors demonstrate that in their case, ligand binding is enthalpically favoured but entropically disfavoured. This highlights a key theme in ligand binding which is particularly relevant to enzymes with high affinity intermediate states - the enzyme must encode some mechanism of ligand dissociation. The alternative, a ligand where both enthalpic and entropic contributions are favoured, will likely result in a high affinity complex that does not dissociate on a meaningful timescale for efficient catalysis.

However, if the purpose is to design such high affinity inhibitors, then a full free energy landscape characterization of the protein and ligand is necessary. An example of this is the case of Gleevec selectively binding to Abl kinase but not Src kinase, both targets for the inhibition of chronic myelogenous leukemia (Wilson et al., 2015). In this case a dynamic characterization of the enzyme indicated that the effectiveness of the drug was not due to the binding of the open state of the Abl and Src enzymes, rather, the drug bound and through an "induced-fit" mechanism, stabilized a minor state of the Abl (not Src) with a low k_{off} rate, resulting in effective inhibition. Furthermore, a previously asserted gatekeeper mutation that drastically reduced the effectiveness of gleevec binding to Abl kinase was characterised not to change the open - closed ratio and rate, rather, it drastically affected the rate of induced fit conformational change.

Sidechain guanidinium - phosphate interactions (between arg and substrate) have been reported to provide substantial binding energies in the range 11-13 kcal/mol for GPDH (Tsang, Amyes, and Richard, 2008), for TIM (Amyes, OĎonoghue, and Richard, 2001), and for OMPDC (Amyes, Richard, and Tait, 2005). For the enzymes GPDH (Tsang, Amyes, and Richard, 2008; Go, Amyes, and Richard, 2010; Reyes, Amyes, and Richard, 2016) and TIM (Go, Amyes, and Richard, 2010; Zhai, Amyes, and Richard, 2014), it was demonstrated that substrate and active site could be assembled from constituent pieces and displayed highly similar transition states to the native reaction (reviewed (Amyes and Richard, 2013)). Furthermore, investigtation of Guanidinium - phosphate interactions provided sufficient energy to induce conformational change in OMPDC (Desai et al., 2012; Reyes, Amyes, and Richard, 2016) and GPDH (Reyes et al., 2015). This is particularly relevant as our enzyme (βPGM) uses this sidechain guanidinium - phosphate interaction to bind both βG1P and G6P in the

active site. This work demonstrates that it is a well conserved binding mechanism, with significant energetic implications.

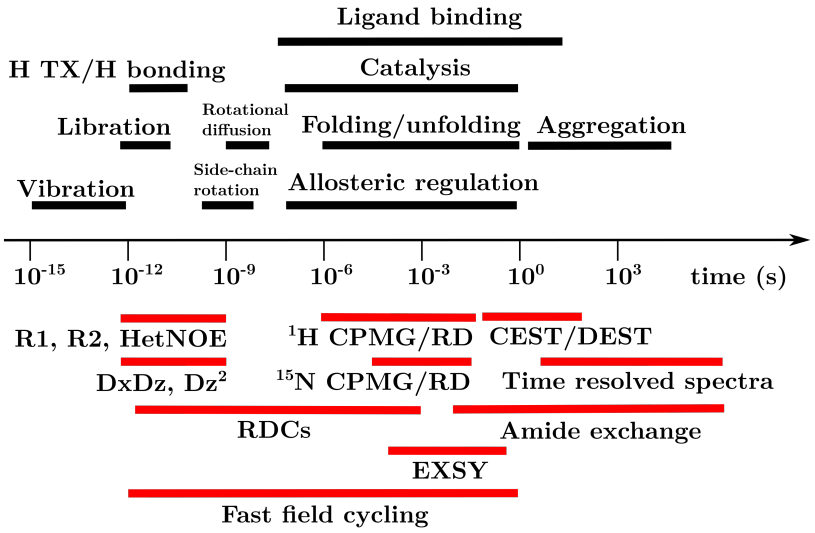


FIGURE 1.2: Timescales for a selection of biological processes (top) and the NMR methods suited to their investigation (bottom).

1.3.1 Free Energy and Conformational Rearrangement

While the reaction coordinate diagram illustrated in Fig. 1.1 gives an elegant overview to the mechanism by which enzymes achieve catalysis, this is by no means the whole picture. This reaction coordinate could refer to a number of things, from bond orientations and electronic distributions, to localisation of substrates and their relative orientations. Consequently, multi dimensional free energy landscapes have developed in attempt to better model the reaction pathway. Like in a reaction coordinate diagram (Fig. 1.1 A), a free energy diagram has a barrier to surmount (the TS) and the relative populations on either side of the barrier (Fig. 1.1 E). As protein conformational changes are often slower than the chemical step of catalysis, it is often the height of the TS in this free energy landscape that is rate limiting.

This free energy landscape model has also been used in attempt to understand enzyme sampling of conformational space in protein folding pathways (Okazaki and Takada, 2008). This has led to the concept of a protein folding funnel, whereby, a protein randomly samples conformational space yet tends towards the lowest energetic conformation (Dill and Chan, 1997). This has recently been corroborated by the direct observation of parallel folding pathways in ubiquitin (Charlier et al., 2018). While the energetic minimum is the most populated at equilibrium, there are often several minima, and it is the exchange between these states that we observe when we observe conformational dynamics in proteins.

Proteins are dynamic across a range of timescales, from electronic fluctuations in chemical bonds on the ns–ps timescale, to domain reorientations and proline isomerisms on the ms–μs and s timescales respectively (Benkovic and Hammes-Schiffer, 2003). Proteins maintain their

dynamism even when substrates are bound. It is frequently a challenge to determine if dynamics occurring on the timescale of the rate of catalysis are implicated in the catalytic cycle of the enzyme as in DHFR case (Schnell, Dyson, and Wright, 2004), or rather, if the dynamics are just an artefact of the enzyme and unrelated to catalysis. Frequently large-scale dynamic processes such as domain-reorientation are involved in product/intermediate release and are the rate limiting step in catalysis (Bae and Phillips, 2006). While dynamics on faster timescales may serve to rigidify the protein for catalytic specificity (Pabis, Duarte, and Kamerlin, 2016), or to encode an efficient mechanism of release of high affinity intermediates. It is argued that dynamics play a minimal role in the chemical step of catalysis, rather, the electrostatic environment plays a much more important role in the chemical step (Pisliakov et al., 2009). This was hotly contested in DHFR, where a 'promoting motion' dynamic was asserted to play a role in the chemical step itself (Hay and Scrutton, 2012).

1.3.2 Phosphoryl transfer enzymes

Under biological conditions, phosphate mono- and di-ester bonds have half-lives of millions of years (Lad, Williams, and Wolfenden, 2003). This property makes the phosphate ester bond essential for many core biological processes such as both long and short term information storage (DNA and cell signalling respectively), and the storage of chemical potential energy (metabolism). Phosphorylation is a potent mechanism of post translational modification of cellular proteins, with a predicted 100,000 phosphorylation sites in the human proteome (Zhang et al., 2002). Furthermore, phosphorylation of intrinsically disordered proteins has recently been observed to impart secondary and tertiary structure essential to function (Bah et al., 2014), but unchecked phosphorylation can lead to pathology, for example the hyperphosphorylation of tau which has been implicated in neurofibrillary degeneration in Alzheimer's disease (Gong et al., 2006).

In order to overcome the immense stability of the phosphate ester bond, the cell requires highly proficient and often specific enzymes to break and reform these stable bonds on appropriate timescales (Todd, 1959; Manning et al., 2002; Graauw, Hensbergen, and Water, 2006). A table outlining some of the key roles of phosphate in biology is presented in Table 1.1. Phosphoryl transfer enzymes have evolved with some of the largest rate accelerations known to biology, with typical catalytic rate enhancements (k_{cat} / k_{uncat}) approaching 10^{21} (Lad, Williams, and Wolfenden, 2003) which makes them prime targets for the investigation of how to catalyze such unfavourable reactions. The phosphoryl transfer field spans nearly 70 years, with excellent reviews of the development of the field published recently (Lassila, Zalatan, and Herschlag, 2011; Kamerlin et al., 2013) in addition to the seminal review by Westheimer (Westheimer, 1987) entitled "Why nature chose phosphates". Several archetypal phosphoryl transfer enzymes have emerged to investigate kinase, phosphatase, and mutase activity over the last six decades, with each contributing to the general understanding and paving way to further discovery (Kamerlin et al., 2013).

1.4 The β-phosphoglucomutase enzyme

Phosphoglucomutase enzymes are one class of phosphoryl transfer enzymes, found in both prokaryotes and eukaryotes, that generate important precursors for glycolysis and anabolism

Phosphate containing component	Biological role
DNA/RNA	Genetic material - information storage
ADP/ATP	Intracellular energy storage and transfer
Pyridoxal phosphate	Coenzyme
Nicotine adenine dinucleotide phosphate	Ca^{2+} signalling
Glucose-6-phosphate	Metabolism
Dihydroxyacetone phosphate	Calvin cycle
Inositol phosphates	Cellular signalling

TABLE 1.1: Biological uses of phosphate

in cells. β-Phosphoglucomutase (βPGM) from *Lactococcus lactis* is a magnesium-dependent phosphoryl transfer enzyme (βPGM, EC. 5.4.2.6) which has been well-characterized physiologically (Qian et al., 1994; Qian et al., 1997; Levander, Andersson, and Rådström, 2001), kinetically (Lahiri et al., 2004; Zhang et al., 2005; Dai et al., 2006; Golicnik et al., 2009) and mechanistically (Lahiri et al., 2004; Dai et al., 2006; Baxter et al., 2006; Dai et al., 2009; Baxter et al., 2010; Griffin et al., 2012; Jin et al., 2014; Johnson et al., 2018). βPGM catalyzes the reversible isomerization of β-glucose 1-phosphate (βG1P) to glucose 6-phosphate (G6P) via a β-glucose 1,6-bisphosphate (βG16BP) intermediate using a ping-pong bi-bi reaction mechanism (Fig. 1.3).

The active site of βPGM is located at the interface between the helical cap domain (T16-V87) and the α/β core domain (M1-D15, S88-K216). Cap opening and closing relative to the core domain occurs during the catalytic cycle, which exposes the active site to solvent and facilitates release of the substrates and the βG16BP intermediate. βPGM transfers a phosphate group from the phospho-enzyme (βPGM^P, phosphorylated at residue D8) to the physiological substrate, βG1P, forming an enzyme-bound βG16BP intermediate. In this complex, the 6-phosphate group remains in the *proximal* catalytic site, while the 1-phosphate group occupies a *distal* phosphate binding site (Fig. 1.4). Release of βG16BP from the active site allows rebinding in the alternate orientation (with the 1-phosphate in the *proximal* site and the 6-phosphate in the *distal* site), leading to phosphoryl transfer from βG16BP to βPGM, generating G6P and βPGM^P as products.

A catalytic Mg^{2+} ion (Mg_{cat}) is located adjacent to residue D8 in the *proximal* site and is coordinated by the sidechain carboxylate groups of D8 and D170, the backbone carbonyl group of D10 and typically two water molecules. The *distal* phosphate binding site is removed from the catalytic center and has a role in anchoring ligands in the active site via interactions with several conserved residues (*e.g.* R49 and K117). The interplay between the two phosphate binding sites allows βPGM^P to bind either βG1P or G6P as substrates, and βPGM to bind the βG16BP intermediate in either orientation, thus facilitating mutase activity.

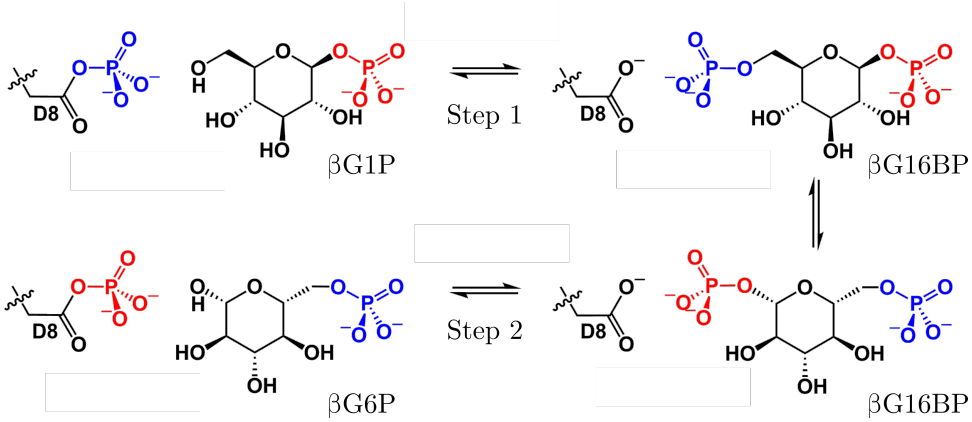


FIGURE 1.3: Reaction scheme for the enzymatic conversion of βG1P to G6P via a βG16BP intermediate adapted from (Johnson et al., 2018). The phosphoryl transfer reaction between the phospho-enzyme (βPGM^P, phosphorylated at residue D8) and βG1P is termed Step 1 and is illustrated with the transferring 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 termed Step 2 and is shown with the transferring phosphate (red) in the *proximal* site and the 6-phosphate (blue) of G6P in the *distal* site. The two intermediate complexes are labeled βPGM:P6G1P and βPGM:P1G6P to explicitly denote the orientation of βG16BP bound in the active site.

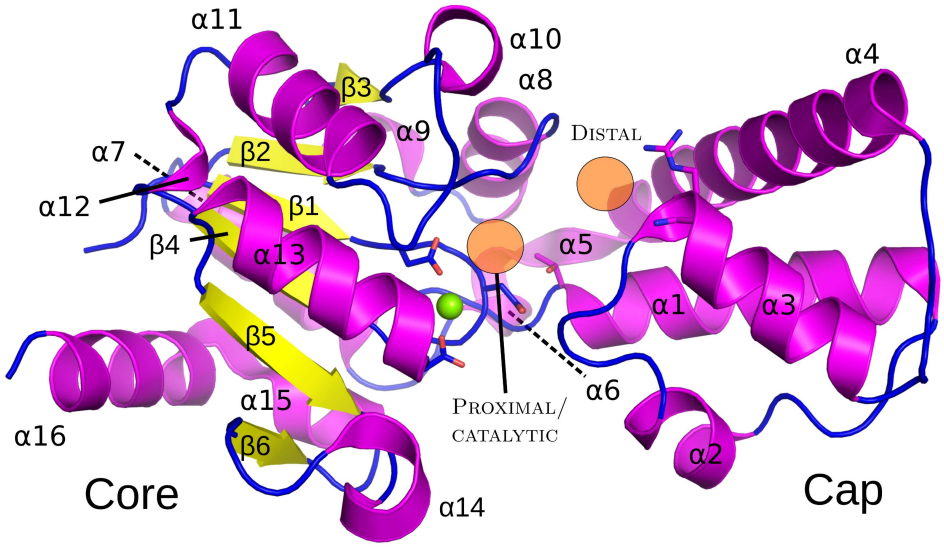


FIGURE 1.4: A ribbon representation of βPGM (PDB: 2WHE) showing the core (left) and cap (right) domains with α-helices coloured purple, β-sheets yellow, with standard CPK colors for atoms. The proximal and distal phosphate binding sites are indicated by orange circles.

1.5 Chronology of the studies on βPGM

Here a brief overview of the βPGM publication history is presented with authors, dates and manuscript titles for reference (Fig. 1.5). Some quotations are used to report directly what was said when key phenomena were presented for the first time, and additionally, to try and avoid misrepresentation on potentially contentious issues. Readers are strongly encouraged to look in the respective manuscripts for further context. A description of the key findings from each paper is presented, along with key tables and kinetic data. In the next section a retrospective analysis is made which discusses many of the findings in the context of work from other groups involved in the βPGM narrative.

1.5.1 "Purification and characterization of two phosphoglucomutases from *Lactococcus lactis* subsp. lactis and their regulation in maltose- and glucose-utilizing cells" (Qian et al., 1994)

In this paper the authors report:

- The first identification of α and β- PGM enzymes in *L. lactis* .
- The first use of KCl gradient to elute βPGM off anion exchange column.
- A parametrization of sample longevity at room temperature and thermostability of the open-βPGM form.
- The phenomenon that growth of *L. lactis* on glucose induces lactate production and acidifies the growth medium. Growth of the cells on maltose, results in production of less lactate, and consequently less acidification of the media.
- That switching growth media of bacteria from maltose to glucose represses the specific activity of both α- and β- PGM.
- The βPGM enzyme appears to have the same function in both *Lactococcus* and *Euglena* cells.

1.5.2 "Product formation and phosphoglucomutase activities in *Lactococcus lactis*: cloning and characterization of a novel phosphoglucomutase gene" (Qian et al., 1997)

In this paper the authors report:

- The characterisation of the PgmB gene and gives a detailed account of its cloning resulting. The authors calculated a molecular mass of 24210 Da which was reportedly in close agreement with the molecular mass of the purified βPGM (25 kDa).
- Exploration of growth on different media building from previous paper. The authors observed that maximum specific activity of βPGM was increased by a factor of 4.6 (and the specific growth rate by a factor of 7), when the maltose concentration was raised from 0.8 to 11.0 g l^{-1}. They concluded that growth on maltose upregulated the activity of βPGM.

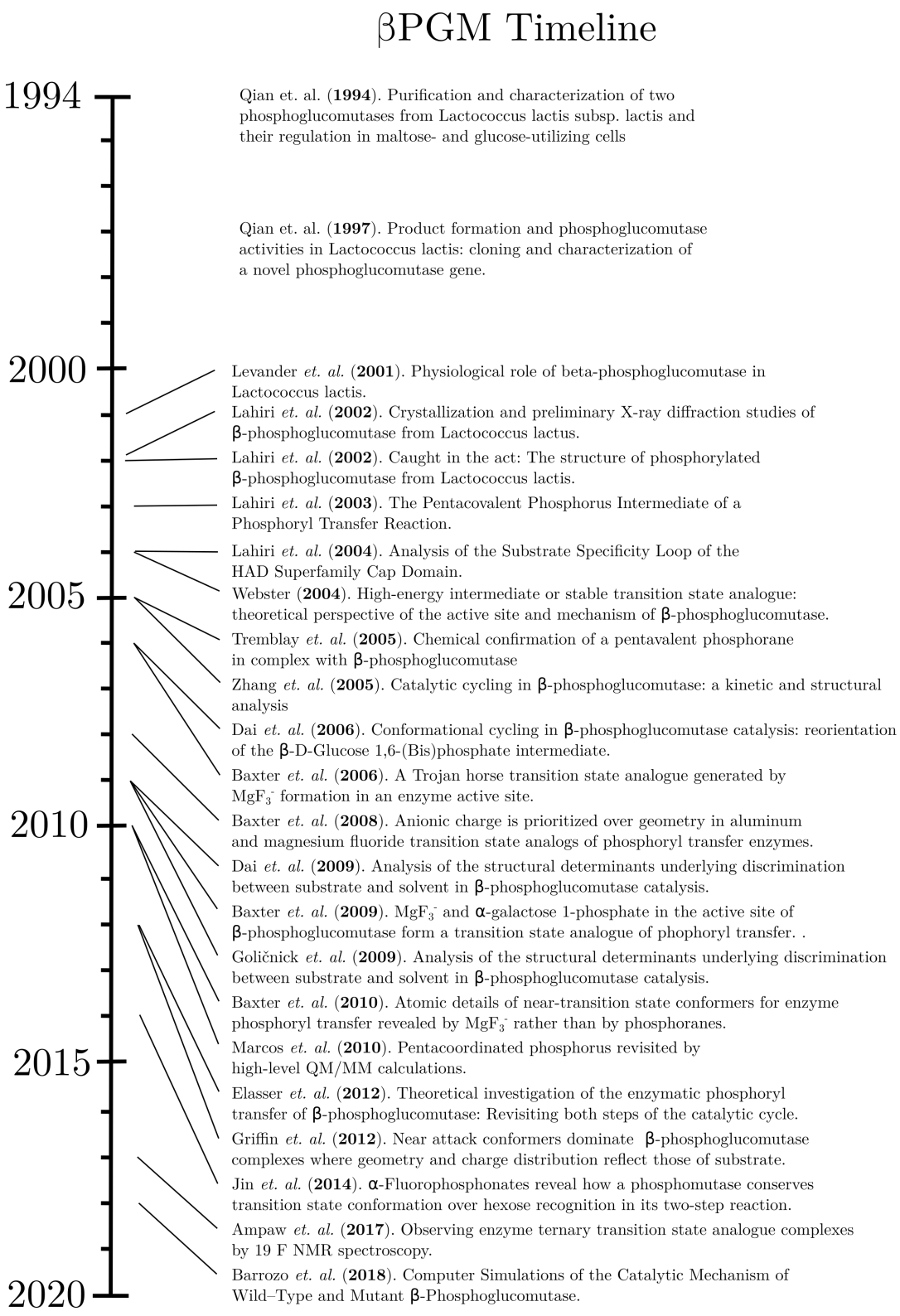


FIGURE 1.5: βPGM publication timeline.

- The first identification of the importance of a divalent metal cation in βPGM where the authors state "A short sequence at the N-terminus was found to be similar to known metal-binding domains..."

1.5.3 "Physiological role of beta-phosphoglucomutase in *Lactococcus lactis*." (Levander, Andersson, and Rådström, 2001)

In this paper the authors report:

- The necessity of βPGM for growth on trehalose and a distinct growth rate improvement on maltose.
- While maltose catabolism in *B. subtilis* could involve βPGM , this enzyme does not have an apparent role in the metabolism of *E. coli*. In the present study, they found evidence that βPGM is a central enzyme in the maltose and trehalose catabolic pathways of *L. lactis* and also that trehalose is assimilated by a novel pathway in this bacterium.

1.5.4 "Crystallization and preliminary X-ray diffraction studies of β-phosphoglucomutase from *Lactococcus lactus*" (Lahiri et al., 2002b)

In this paper the authors report:

- The first crystallization of βPGM from *L. lactis*.
- The use of a SeMet labelling approach, with structure solution attempted using multiple wavelength anomalous dispersion (MAD) phasing method, using three wavelengths.

Crystals were obtained in two different conditions:

- 0.2M ammonium acetate, 0.1M trisodium citrate dihydrate pH 5.6, 30%(w/v) PEG 4000.
- 0.2M ammonium fluoride, 20% (w/v) PEG 3350 (unbuffered; measured pH 6.5).
- Further optimization of these conditions to 0.15M ammonium acetate, 0.1M trisodium citrate dihydrate pH 4.5, 25% (w/v) PEG 4000 for the former and 0.1M ammonium fluoride, 16% (w/v) PEG 3350 for the latter resulted in good-quality plate-like crystals suitable for diffraction.

1.5.5 "Caught in the Act: The Structure of Phosphorylated β-Phosphoglucomutase from *Lactococcus lactis*," (Lahiri et al., 2002a)

Associated PDB files: 1LVH

In this paper the authors report:

- The structure solution from the MAD phasing reported previously. The authors build Cα chain through density then add sidechains. All but terminal 4 residues were modelled.
- The electron density in the active site and model both Mg^{2+} and phosphate in the active site. They assert that they have crystallized phosphorylated enzyme that co-purified in the phosphorylated form.
- The authors note that no phosphate, sulfate, or similar ions were included in the crystallization solution, which prompted the assertion that "...the assignment of the electron density to a phosphoryl group covalently linked to Asp8 is unequivocal."
- The first residue-specific mechanistic prediction "...the β-PGM active site might be compatible with a two- nucleophile mechanism in which Asp8 and Asp10 function as the mediators of phosphoryl transfer."
- A description of how the βPGM fold relates to the rest of the HADSF fold and elaborate on how several phosphotransferases may have specialized to perform function.
- Due to the 'out' orientation of D10, it cannot bind and activate a water molecule for in-line attack of the aspartyl phosphate group, which may indicate why the phosphoaspartate ($D8^P$) was so stable.
- A half life of AcP was previously determined to be 21h under pseudo-physiological conditions (Di Sabato and Jencks, 1961) which the authors rationalize to be the reason that they could observe a phosphorylated aspartate residue crystallographically.

1.5.6 "The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction" (Lahiri et al., 2003)

Associated PDB files: 1O03,1O08.

In this paper the authors report:

- The crystal structure of βPGM in a closed conformer with a trapped pentavalent phosphorane intermediate.
- The phosphorane intermediate is trapped when βPGM is co-crystallized with either βG1P or G6P ligands.
- "Inversion matrix calculations of the [electron density] data show that the accuracy of the bond lengths is ± 0.11 Å and that of the bond angles is ± 3°. (Sheldrick and Schneider, 1997)
- The "...in-line (174°± 3°angle) arrangement of the Asp8 OD1 nucleophile and C(1)O leaving group, conforms to the trajectory expected for the reaction coordinate of an associative nucleophilic substitution reaction at phosphorus".

1.5.7 "Comment on "The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction"" (Blackburn et al., 2003)

In this technical comment the authors discuss:

- The apical bond lengths for the trigonal bipyramidal intermediate (2.0 and 2.1 Å) are longer than bond lengths in small molecule phosphoranes, such as pentacyclohexyloxyphosphorane where bond lengths are 1.65 Å and 1.67 Å. Given that the observed species resembles a transition state, these distances may not be unreasonable.
- At 93 K, the half life for decomposition of the activated complex is $\leq 10^{-11}s$, so it is unlikely to be a trapped intermediate state.
- The reduction in temperature from 291 K to 93 K would have led to a decrease in $\Delta G‡$ of *ca.* 12 kcal mol^{-1} for the reaction within the ES complex.
- Previously, it was determined that MgF_3^- assembled in the active site of a small GTPase (Graham et al., 2002), and given the 10 mM $MgCl_2$ and 100 mM NH_4F used in the crystallization conditions, the authors suggest that the observed species may in fact be a MgF_3^- TSA.
- While a change in interpretation was essential, the model can still be interpreted in terms of dissecting the nature of the TS (concerted/dissociative/intermediate, etc...)

1.5.8 "Response to Comment on "The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction"" (Allen and Dunaway-Mariano, 2003)

In this technical comment the authors discuss five key points to defend their initial interpretation:

- Bradford protein and Malachite Green phosphate assays on washed crystals define an enzyme:phosphate stoichiometry of 1:2 with 10% error.
- Mg–F bond distances are in the range 1.9-2.0 Å, whereas the equatorial P–O bond distances reported in (Lahiri et al., 2003) were 1.7 ± 0.1
- Anomalous-difference electron density shows electron density of equal magnitude at both assigned phosphorus positions, which is not consistent with a MgF_3 TSA as the scattering from the Mg ion is less than half that of P (at the wavelength of data collection).
- "...crystals of the β-PGM complex are formed in crystallization solutions containing as little as 1 mM ammonium fluoride, yet ammonium fluoride at three times this concentration does not inhibit β-phosphoglucomutase catalysis."
- "... the phosphorane intermediate observed in the β-PGM-complex structure has precedent in chemical models (7,8). The magnesium trifluoride species cited by Blackburn *et al.* (1), by contrast, has no proven chemical model."
- "In summary, the study in (9) [(Graham et al., 2002)] offers no proof of the existence of magnesium trifluoride in solution or bound to the G protein, nor any explanation of why Mg(II) would form magnesium trifluoride in neutral solution or in the active site of an enzyme."

1.5.9 "Analysis of the Substrate Specificity Loop of the HAD Superfamily Cap Domain" (Lahiri et al., 2004)

The authors kinetically characterize several variants of βPGM (Table 1.2). The authors report:

- A compare and contrast between βPGM and phosphonatase.
- A kinetic characterisation of several variants of βPGM using the initial rate of turnover of a spectrophotometric coupled assay: WT, K45A, K45R, G46P, G46V, G46A, R49K, R49A, S52A (Table 1.2).
- Identification of the salt bridge between K45 and D170.
- That R49 plays an important role due to the mutations effect on k_{cat} / Km and the authors suggest that the interaction between R49 and the phosphate in the non-catalytic site " ... is important to substrate binding and domain-domain closure." and conclude the paragraph stating that "Arg49 is therefore essential to efficient β-PGM catalysis."

enzyme	k_{cat} (s^{-1})	Km (μM)	k_{cat} /Km (s^{-1} M^{-1})
wild type	17.6 ± 0.6	4.6 ± 0.5	3.7×10^{6}
K45A	0.037 ± 0.002	240 ± 40	1.5×10^{2}
K45R	0.61 ± 0.01	8.5 ± 0.7	7.2×10^{4}
G46P	0.126 ± 0.003	2200 ± 100	5.7×10^{1}
G46V	0.0046 ± 0.0001	18.8 ± 0.8	2.6×10^{2}
G46A	0.018 ± 0.002	2000 ± 100	9.0×10^{0}
R49K	0.0580 ± 0.0009	180 ± 10	3.2×10^{2}
R49A	0.078 ± 0.002	13300 ± 900	5.9×10^{0}
S52A	9.8 ± 0.6	4.04 ± 0.04	2.5×10^{6}

TABLE 1.2: Kinetic characterisation of several variants of βPGM presented in (Lahiri et al., 2004) using a spectrophotometric coupled assay

1.5.10 "High-Energy Intermediate or Stable Transition State Analogue: Theoretical Perspective of the Active Site and Mechanism of β–Phosphoglucomutase" (Webster, 2004)

The author reports:

- The use of ONIOM(B3LYP:PM3MM) calculations to characterise the nature of the previously reported phosphorane intermediate using the crystal structure as a starting model.
- The calculated TS for the MgF_3 analog "...geometry matches that of the reported crystal quite well". Optimization for a five coordinate phosphorus "...does not converge to a stable minimum but to a transition state (TS) for phosphoryl transfer." Distances are given in table 1.3
- TS energy barrier predicted to be +14.0 kcal mol^{-1}, 147*i* cm^{-1}.

Bond	Reactant (Å)	Transition state (Å)	Bis-phospho intermediate (Å)
D8 – P	1.78	2.24	2.94
1-OH – P	3.07	1.99	1.70

TABLE 1.3: Calculated bond lengths for a transferring phosphate group in step 2 of the βPGM reaction scheme (assembled from numbers in text), presented in (Webster, 2004)

- The author summarizes the calculation work with "In any case, the concerted transfer of the phosphoryl group is clear; there is no five-coordinate phosphorane or three-coordinate metaphosphate intermediate"

1.5.11 "Catalytic cycling in β-phosphoglucomutase: A kinetic and structural analysis" (Zhang et al., 2005)

Associated PDB files: 1ZOL

The authors present kinetically determined parameters (Table 1.4) and the authors report:

- The defence of the assertion in PDB: 1LVH (Lahiri et al., 2002a) that the enzyme is phosphorylated, they suggest that phosphate in solution may have spontaneously phosphorylated the enzyme.
- The essential role of D8 was identified when catalytic inactivity was observed in the D8N and D8E protein variants.
- A Km for Mg^{2+} activation of 270 $\pm$ 20 μM.
- Observed activity in the E169A/D170A double loop mutant, as well as the D170A mutant.
- The authors use varying concentrations of sodium tungstate (0, 125, 250, and 500 μM) or molybdate (0, 160, and 320 μM) in 50 mM K^+ HEPES (pH 7.0) to inhibit βPGM catalysis.
- The re-crystallization of the open enzyme (PDB: 1ZOL).
- The prediction that "Arg49 with the 'nontransferring' phosphoryl group of the substrate ligand might stabilize the cap-closed conformation, as required for active site desolvation and alignment of Asp10 for acid-base catalysis."
- A k_{cat} as a function of pH profile was calculated for βPGM from pH 4.5 – 9.5. [The buffer was seemingly HEPES (buffering range of *ca.* 6.8–8.3.) although it is unclear.]
- It was identified that phosphoryl transfer occurs from C(6) of α-G16BP, rather than C(1), when catalysis is primed with the bisphospho-substrate.
- The authors identify a phospho-β-PGM (βPGM^P) hydrolysis rate of 3 min^{-1}.

1.5.12 "Chemical Confirmation of a Pentavalent Phosphorane in Complex with β-Phosphoglucomutase" (Tremblay et al., 2005)

Associated PDB files: 1Z4N, 1Z4O.

enzyme	Km – Mg^{2+} (μM)	Km – βG1P (μM)	k_{cat} (s^{-1})
wild type	270 ± 20	14.6 ± 0.5	17.1 ± 0.6
D8A	inactive		($< 10^{-5}$ s^{-1})
D8E	inactive		($< 10^{-5}$ s^{-1})
D170A	ND	7.8 ± 0.2	(3.84 ± 0.03) $\times 10^{-3}$
E169A/D170A	ND	390 ± 20	(1.20 ± 0.02) $\times 10^{-3}$

TABLE 1.4: Kinetic Constants for Wild-Type and Mutant β-PGMs (in the presence of 50 μM α-G16P as the activator) presented in (Zhang et al., 2005) using a spectrophotometric coupled assay.

The authors report:

- The crystallization of α-galactose 1-phosphate in the active site of βPGM which does not appear to form an MgF_3^- TSA despite the assertion that the same concentrations of ammonium fluoride and $MgCl_2$ were used as in the phosphorane intermediate structure (Lahiri et al., 2003).
- Malachite green phosphate binding assay data are also presented in attempt to validate the reported phosphorane complex. (PDB: 1O08; (Lahiri et al., 2003))

1.5.13 "Conformational cycling in β-phosphoglucomutase catalysis: Reorientation of the β-D-glucose 1,6-(bis)phosphate intermediate" (Dai et al., 2006)

The authors report:

- β-Phosphoglucomutase catalyzes the conversion of βG1P to G6P (in the presence of 2 mM Mg^{2+} (Km = 270 ± 20 μM)) with a k_{cat} = 177 ± 9s^{-1} and Km = 49 ± 4 μM at pH 7.0 and 25 °C.
- A steady-state k_{cat} = 177 ± 9s^{-1} for multiple turnover of βG1P when βPGM is activated by the βG16BP intermediate.
- The use of a radiolabelled substrate approach to investigate whether the βG16BP intermediate dissociates from the active site at the mid point in the reaction, or maintains contact with the enzyme and is "flipped" in the active site. The authors conclude that the ligand is released into solution to freely dissociate.
- Kinetics of the autophosphorylation rate of βPGM by βG1P with k_{cat} = 0.83 ± 0.01 s^{-1} and Km = 400 ± 40 μM.
- Single turnover reactions of [U-^{14}C]βG16BP with excess βPGM demonstrated that phosphoryl transfer (rather than ligand binding) is rate-limiting in catalysis, and corroborate the observation that the βG16BP intermediate binds to the active site both orientations with roughly the same efficiency.

1.5.14 "A Trojan horse transition state analogue generated by MgF_3^- formation in an enzyme active site" (Baxter et al., 2006)

The authors report:

- For open βPGM, the resting state was not phospho-enzyme (βPGM^P) using ^{31}P 1D NMR. They further demonstrated that there was no accumulation of enzyme bound phosphorane intermediate in solution either using the same approach.
- ^{31}P NMR was used to demonstrate that G6P and MgF_3^- formed an observable complex, this observation was then corroborated by the observation of MgF_3 peaks by ^{19}F NMR.
- ^{19}F–^1H NOE spectra were performed to assign the fluorides in the active site.
- Backbone NMR assignments of both the apo form and the βPGM:MgF_3:G6P complex were performed.
- Inhibition of catalysis by μM fluoride concentrations is demonstrated in contrast to previous reports (Allen and Dunaway-Mariano, 2003; Tremblay et al., 2005)

1.5.15 "Anionic charge is prioritized over geometry in aluminum and magnesium fluoride transition state analogs of phosphoryl transfer enzymes" (Baxter et al., 2008)

The authors report:

- The use of ^{19}F NMR to demonstrate that for βPGM, a pH-jump does not induce the fluoride coordination number of an AlF_4^- moiety to drop to AlF_3, with concomitant charge neutralization. Instead, AlF_4^- is progressively replaced by MgF_3^- as the pH increases. The authors conclude this point by indicating that the enzyme prioritized anionic charge over native trigonal geometry over a broad pH range.
- Backbone NMR assignments of the βPGM_{WT}:AlF_4:G6P complex, as well as the four fluoride peaks by ^{19}F:^1H–NOE.
- Through competition study, the authors demonstrate that any physiological effects of enzyme inhibition by metal fluorides are the result of aluminum fluoride species. This directly correlates with previous work where it was observed that high F concentrations leech Al out of glassware (which typically comprises >4% (Sternweis and Gilman, 1982; Wittinghofer, 1997)).
- The authors finish by noting "... that several, if not a majority, of the high-pH AlF_3^0 transition state analogue complexes reported for nucleotide kinases, should be treated with caution".

1.5.16 "Analysis of the Structural Determinants Underlying Discrimination between Substrate and Solvent in β-Phosphoglucomutase Catalysis" (Dai et al., 2009)

Associated PDB files: 3FM9.

The authors perform steady state kinetics on several enzyme variants of βPGM (Table 1.5). The authors present βG16BP hydrolysis rates by several enzyme variants (Table 1.6), and the authors also perform further single turnover reactions (Table 1.7).

The authors report:

- Replacement of D10 with A, S, C, N, or E resulted in no observable activity.
- The authors also present a T16P open-structure with D10 in a strained rotamer.
- The authors present a model whereby repositioning of D10 (from 'out' to 'in') is concerted with domain closure, and a transition from hydrogen bonding with T16-A17 in the open state to the H20-K76 pair in the closed state.
- The authors state that "The model predicts that Asp10 is required for general acid/base catalysis and for stabilization of the enzyme in the cap-closed conformation. It also predicts that hinge residue Thr16 plays a key role in productive domain-domain association, that hydrogen bond interaction with the Thr16 backbone amide NH group is required to prevent phospho-Asp8 hydrolysis in the cap-open conformation, and that the His20-Lys76 pair plays an important role in substrate-induced cap closure."
- In the T16P variant the authors observe a reduced rate of D8 phosphorylation by βG16BP, a reduced rate of equilibration βG1P with G6P, and an enhanced rate of phosphoryl transfer from phospho-Asp8 to water.
- The authors predict that "Taken together, the results support a substrate induced-fit model of catalysis in which βG1P binding to the core domain facilitates recruitment of the general acid/base Asp10 to the catalytic site and induces cap closure."

β-PGM	Km (μM)	$k_{cat}(s^{-1})$	$k_{cat}/Km\ (M^{-1}\ s^{-1})$
WT	31 ± 2	175 ± 5	6×10^6
D10N	ND	<0.001	ND
D10S	ND	<0.001	ND
D10C	ND	<0.001	ND
T16P	4.8 ± 0.3	0.026 ± 0.001	5×10^3
H20Q	45 ± 1	21.9 ± 0.1	5×10^5
H20N	170 ± 10	0.62 ± 0.02	4×10^3
H20A	41 ± 3	0.026 ± 0.004	6×10^2
K76A	66 ± 1	1.56 ± 0.01	2×10^4

TABLE 1.5: Kinetics taken from (Dai et al., 2009). ND = Not Determined, Steady-State Kinetic Constants for Wild-Type and Mutant β-PGM Measured Using Assay Solutions Containing Varying Concentrations of βG1P, 5 μM βG1,6bisP, 2 mM $MgCl_2$, 0.2 mM NADP, and 2.5 Units/mL Glucose 6-Phosphate Dehydrogenase in 50 mM K^+ HEPES (pH 7.0 and 25 °C)

β-PGM	Km (μM)	k_{cat} (s^{-1})
wild-type	0.63 ± 0.07	0.0298 ± 0.008
H20N	1.49 ± 0.09	0.0303 ± 0.0005
H20A	2.6 ± 0.1	0.0197 ± 0.003
T16P	10 ± 1	0.38 ± 0.01

TABLE 1.6: Kinetics taken from (Dai et al., 2009). Steady-State Kinetic Constants of β-PGM-Catalyzed βG1,6bisP Hydrolysis in 50 mM K^+ HEPES (pH 7.0, 25 °C) Containing 2 mM $MgCl_2$ and βG1,6bisP at Various Concentrations.

β-PGM	[βG1,6bisP] (μM)	kobs(βG1P) (s^{-1})	kobs(βG6P) (s^{-1})	G6P/βG1P
wild-type	5	12.2 ± 0.3	9.3 ± 0.4	10
T16P	50	0.014 ± 0.002	0.014 ± 0.002	0.6
H20Q	5	0.96 ± 0.04	0.78 ± 0.04	10
H20N	50	0.042 ± 0.002	0.018 ± 0.002	10
H20A	50	0.0055 ± 0.0004	0.0029 ± 0.003	1
K76A	5	0.164 ± 0.006	0.157 ± 0.004	10

TABLE 1.7: Kinetics taken from (Dai et al., 2009). Apparent Rate Constants of the Wild-Type and Mutant *L. lactis* β-PGM (40 or 20 μM)-Catalyzed Single-Turnover Reactions of [14C]βG1P (5 μM) in the Presence of βG1,6bisP (5 or 50 μM). The k_{obs} values for [^{14}C]G6P or [^{14}C] βG1,6bisP formation and [^{14}C]βG1P consumption were obtained by fitting the individual sets of time course data to first-order rate equations.

1.5.17 "Kinetic Analysis of β-Phosphoglucomutase and Its Inhibition by Magnesium Fluoride." (Golicnik et al., 2009)

The authors kinetically characterize and model the reaction scheme in βPGM (Table 1.8). The authors report:

- The first documentation of the lag phase prior to steady state kinetics in βPGM.
- The addition of βG16BP to the reaction eliminated the lag phase prior to steady state kinetics and modelled the lag phase as a competition reaction between the non-productive binding of βG1P to un-phosphorylated enzyme, and βG16BP binding to the enzyme active site and generating phospho-enzyme.
- Kinetic modelling of the βPGM reaction scheme with fitting the observed data to two reaction schemes (Table 1.8; Fig. 1.6, 1.7).
- When fluoride and magnesium ions are present, time-dependent inhibition of the βPGM is observed.
- A βPGM:MgF_3:βG1P inhibitory complex is formed when the equilibration of βG1P and G6P is performed in the presence of magnesium and fluoride with inhibitions occurring at slow (minutes) and fast (ms) timescales (illustrated in Fig. 1.8; Table 1.9). The overall stability constant for this complex is approximately 2×10^{-16} M^5. When a conservative estimate of the association constant of MgF_3 for the active site is made, a Kd of the MgF_3^- moiety for this transition-state analogue (TSA) of *ca.* 70 nM.

- The authors conclude that "The preference for TSA formation when fluoride is present, and the hydrolysis of substrates when it is not, rules out the formation of a stable pentavalent phosphorane intermediate in the active site of βPGM."

parameter	fit to Scheme 3	fit to Scheme 5	calculated
k_{cat}	$64.7 \pm 0.7\ s^{-1}$		$74 \pm 15\ s^{-1}$
Km (βG1P)	$14.7 \pm 0.5\ \mu M$		$15 \pm 4\ \mu M$
Km (βG16BP)	$0.72 \pm 0.04\ \mu M$		$0.8 \pm 0.2\ \mu M$
Ki (βG1P)	$122 \pm 15\ \mu M$	$122 \pm 8\ \mu M$	$122 \pm 8\ \mu M$
k1		$253 \pm 22\ s^{-1}$	
K1 (βG1P)		$51 \pm 5\ \mu M$	
k2		$105 \pm 3\ s^{-1}$	
K2 (βG16BP)		$1.1 \pm 0.1\ \mu M$	
k5		$2.9 \pm 0.3\ s^{-1}$	
K5 (αG16BP)		$91 \pm 6\ \mu M$	
Kiα (αG16BP)		$21 \pm 2\ \mu M$	
kH2O	$0.026 \pm 0.001\ s^{-1}$		

TABLE 1.8: Kinetics presented in (Golicnik et al., 2009). Kinetic Parameters Determined under Steady-State Conditions (from Initial Rates in the Presence of βG16BP) and Non-Steady-State (from Progress Curves in the Presence of αG16BP). k_{H_2O} determined from steady-state rates when only βG16BP was present in the reaction mixture and the rate-limiting step at steady state is the hydrolysis of phosphoenzyme. Conditions: 2 mM $MgCl_2$, 50 mM K^+ HEPES, pH 7.2.

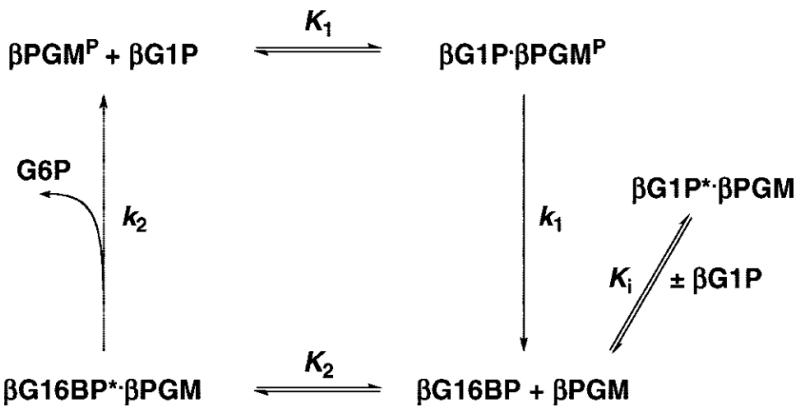


FIGURE 1.6: Scheme 3 adapted from (Golicnik et al., 2009).

	fluoride-dependent	magnesium-dependent
K_f	2.8 ± 0.1	
$k_{i1} M^{-1} s^{-1}$	166 ± 10	173 ± 5
$k_{i2} s^{-1}$	0.037 ± 0.007	0.035 ± 0.008
$k_{i3} s^{-1}$	0.010 ± 0.006	0.011 ± 0.004
$k_{i4} s^{-1}$	0.0056 ± 0.0016	0.0049 ± 0.0016

TABLE 1.9: Kinetic constants calculated determined from the time-dependent inhibition progress curves reported in (Golicnik et al., 2009).

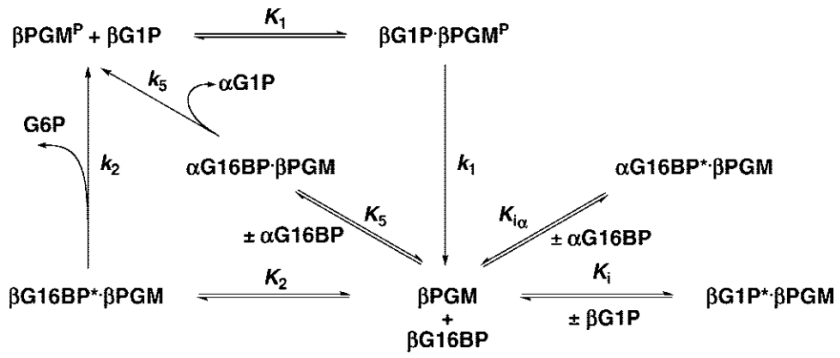


FIGURE 1.7: Scheme 5 adapted from (Golicnik et al., 2009).

1.5.18 "MgF_3^- and α-Galactose 1-Phosphate in the Active Site of β-Phosphoglucomutase Form a Transition State Analogue of Phosphoryl Transfer" (Baxter et al., 2009)

The authors report:

- ^{19}F 1D spectra demonstrating that MgF_3^- forms in the active site of βPGM in the presence of α-galactose 1-phosphate, contrasting to the evidence presented previously (Tremblay et al., 2005).
- ^{19}F–^{1}HNOE data are used to corroborate the fact that the MgF_3 species is in the active site of βPGM, along with J_{HF} couplings present in the observed backbone amide ^{1}H-^{15}N–2D spectra.
- This manuscript also notes that there is the loss of a hydrogen bond to Fc [now termed F1] in the MgF_3^- TSA moiety on binding α-galactose 1-phosphate rather than G6P (based on ^{19}F 1D NMR spectra). This is due to the loss of the 2-OH group of G6P which can stably coordinate the Fc fluorine position [now termed F1].

1.5.19 "Atomic details of near-transition state conformers for enzyme phosphoryl transfer revealed by MgF_3^- rather than by phosphoranes" (Baxter et al., 2010)

Associated PDB files: 2WF5, 2WF6, 2WHE.

The authors report:

- The ^{31}P NMR spectrum of βPGM expressed and purified according to established procedures by both groups (Allen/Dunaway-Mariano and Waltho/Blackburn/Hollfelder) and showed that freshly prepared protein has no phosphate moiety covalently bound to D8. Thus, neither the phosphorane species or the metastable phospho-enzyme species were directly observed at equilibrium.
- The authors addressed "...the remarkable hypothesis that the presence of Pi is sufficient to cause phosphorylation of D8 before or after crystallization (Zhang et al., 2005)". The authors reported no aspartyl-phosphate peak in the presence of high concentrations of

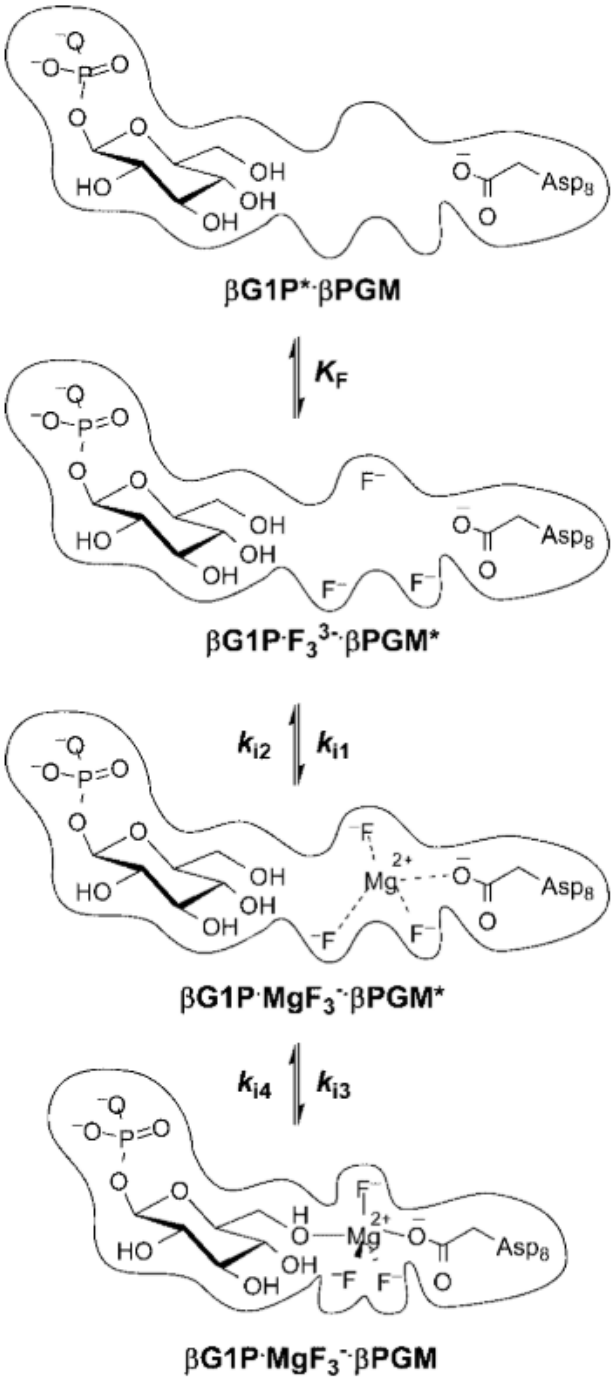


FIGURE 1.8: Scheme 6 adapted from (Golicnik et al., 2009).

phosphate. The authors predict that what was observed crystallographically may be an AlF_4^- moiety and present a 1D ^{19}F NMR of AlF_4^- binding to the open-enzyme.

- The βPGM:MgF_3:G6P complex under near identical conditions to the NMR solution and observed an MgF_3^- moiety in the active site. The authors also re-refine the initial "phosphorane" structure with MgF_3 and demonstrate the removal of difference map peaks.

- Several Kd values for TSA formation were determined for solutions containing magnesium and fluoride by ^{1}H NMR and ITC: G6P = 1 μM, 2-deoxyG6P = 80 μM, 6-deoxy-6-(phosphonomethyl)-D-glucopyranoside = 300 μM.
- Solvent isotope shifts for the ^{19}F resonances in the MgF_3 TSA complexes were plotted against both chemical shift and F-H distance in the crystal complex. The authors conclude the paper stating that "...the metal fluoride complexes offer opportunities to measure properties of near-TS complexes that are currently unmeasurable for phosphorus oxide species, in particular the independent measures of local electrostatic and hydrogen-bonding distributions using ^{19}F 1D NMR."

1.5.20 "Pentacoordinated phosphorus revisited by high-level QM/MM calculations" (Marcos, Field, and Crehuet, 2010)

The authors report:

- A determination of the reaction path of the phosphorylation step using high-level QM/MM calculations, and calculate the geometry of the transition state analogue complex, which is in good agreement with the βPGM:MgF_3:G6P TSA crystal structure.
- "Our results reveal that the rate limiting step for the production of G6P from the phosphorylated enzyme is the chemical process of phosphoryl transfer, with an activation energy that corresponds well to the experimental rate constant obtained by Waltho and co-workers".
- "We also show that the TSA is a good mimic of the true TS" and that the timing of proton transfer from the nucleophilic hydroxyl group of G6P (to D10) occurs after much of the P–O(C1)-G6P bond formation has occurred (ie. when considering the reverse of the step 2 process, going from phospho-D8 to form the βG16BP).

1.5.21 "Theoretical investigation of the enzymatic phosphoryl transfer of β-phosphoglucomutase: revisiting both steps of the catalytic cycle" (Elsässer, Dohmeier-Fischer, and Fels, 2012)

The authors report:

- The investigation of both steps (βG6P to βG16BPand βG16BPto βG1P) of the reaction using QM/MM theoretical method at the DFT/PBE0 level of theory, as well as NEB (nudged elastic band) and free energy calculations to identify transition states and free energies.
- "Ser114 and Lys145 also play important roles in stabilizing the large negative charge on the phosphate through strong coordination with the phosphate oxygens and guiding the phosphate group throughout the catalytic process."
- The calculated energy barrier for the βG1P to βG16BP step is only slightly higher than for the βG16BP to βG6P step (16.10 kcal mol^{-1} versus 15.10 kcal mol^{-1}) and is in excellent agreement with experimental findings (14.65 kcal mol^{-1}).

- The timing of proton transfer from the nucleophilic hydroxyl group of G6P (to D10) occurs after very little of the P–O(C1)-G6P bond formation has occurred (ie. when considering the reverse of the step 2 process, going from phospho-D8 to form the βG16BP).

1.5.22 "Near attack conformers dominate β-phosphoglucomutase complexes where geometry and charge distribution reflect those of substrate" (Griffin et al., 2012)

Associated PDB files: 2WF8, 2WF9, 2WFA.

The authors report:

- ^{19}F 1D NMR spectra that demonstrate that a beryllium trifluoride moiety (BeF_3^-) readily assembles in the active site of βPGM without requirement for chemical synthesis.
- Backbone NMR assignment of the βPGM:BeF_3^- complex, with indication that the solution data predict an open conformer like the open state. The authors also note that generation of a phospho-enzyme mimic brings residues out of intermediate exchange (ms timescale) in the active site as is observed in the open state of the enzyme.
- Backbone NMR assignment of the βPGM:BeF_3^-:G6P complex and predict that the solution state is *ca.* 50-60% closed (between open and MgF_3:G6P reference points).
- The crystal structures of: βPGM:$BeF_3^-$and βPGM:BeF_3:G6P in a hydrogen-bonded (NACI) and aligned (NACII) near attack complexes.
- Despite the addition of BeF_3, significant mutase activity persists in the enzyme, as there is a partial occupancy of βG1P in the active site.
- ^{19}F NMR is presented of both the βPGM:BeF_3^- complex (3 protein bound peaks) and the βPGM:BeF_3^-:G6P complex (2 observed protein bound peaks).
- "In the active site, the fluoride ions (average charge = –0.67) carry ∼70% of the negative charge of the corresponding oxygen atoms (average charge = –0.95), and the beryllium atom (charge = +1.2) carries ∼60% of the positive charge of the corresponding phosphorus atom (charge = +2.1). Hence, though the fluoroberyllate moiety carries approximately the same overall charge as the genuine reacting species, it is somewhat scaled down in terms of internal charge separation."

1.5.23 "α-Fluorophosphonates reveal how a phosphomutase conserves transition state conformation over hexose recognition in its two-step reaction" (Jin et al., 2014)

Associated PDB files: 2WF7, 4C4R, 4C4S, 4C4T.

The authors report:

- The design and synthesis of novel phosphonate analogues of βG1P to probe the first phosphoryl transfer step of βPGM. The authors conclude that the enzyme necessarily generates alignment of the scissile O–P bond with the nucleophile in all of the TSA complexes observed for both steps of the reaction in βPGM.

- That βPGM accomplishes step 1 and step 2 of its reaction within a near-identical, closed protein conformation. The primary differences between the TSA complexes for steps 1 and 2 lie in how the enzyme accommodates substrate. The authors apportion the binding event into three zones; the inert (non-transferring phosphate), the hexose ring, and the transferring phosphate.
- The inert phosphate is coordinated identically across all of the TSAs studied (PDB ID codes 4C4R, 4C4S, 4C4T, 2WF5, 2WF6, and 2WF7). This shows that phosphate is coordinated in the non-transferring site site in essentially the same way for both steps of the reaction.
- The hexose ring is accommodated differently in step 1 and step 2 of the reaction as a result of a $\sim$1.5Å translation towards the transferring phosphate site. A cavity emerges between the hexose ring and the protein on binding of βG6P (as opposed to βG1P) and the resulting space is occupied by two water molecules not present in step 1. Only one direct hydrogen bond to Gly46 maintained between the two steps.
- The analogue of the transferring phosphate group is accommodated nearly-identically across the two steps, with highly comparable O_{ax}–Mg^{2+}–O_{ax} bond lengths and angles (176° for both; 4.1 Å and 4.3 Å for step 1 and 2 respectively).

1.5.24 "Observing enzyme ternary transition state analogue complexes by 19F NMR spectroscopy" (Ampaw et al., 2017)

Associated PDB files: 5OLW, 5OLX, 5OLY.

The authors present inhibition data for fluoro-phosphonate analogs of βG1P (Table 1.10). The authors report:

- The use of a 5-Fluoro Tryptophan labelling strategy to oberve ternary complex formation with MgF_3^- and AlF_4^- TSAs, with *ca.* 85 % labelling efficiency.
- The crystal structure of the 5FWβPGM:MgF_3:G6P complex and demonstrate some non-essential ring flips of W216 which is distant from the active site.
- The conclusion that "This data unequivocally demonstrates that the concentration of the metal fluoride complexes is equivalent to the concentration of enzyme and ligand in the TSA complex in aqueous solution."
- The authors present ^{19}F 1D spectra of 5FWβPGM in both open forms and with non-hydrolysable fluoro-phosphonate analogues of βG1P and MgF_3.

1.5.25 "Computer simulations of the catalytic mechanism of wild-type and mutant β-phosphoglucomutase" (Barrozo et al., 2018)

The authors report:

- A substrate-assisted mechanism of phosphoryl transfer is viable for many phospho-transferases.

	Wild-type βPGM	5FWβPGM
Km	9.0 ± 0.7 μM	10.1 ± 2.1 μM
k_{cat}	7.7 ± 0.1 s^{-1}	3.8 ± 0.1 s^{-1}
IC_{50} (βG1CP)	18 ± 3 μM	13 ± 5 μM
IC_{50} (βG1CFsP)	15 ± 2 μM	11 ± 2 μM
$Ki_{(comp)}$ (βG1CP)		4.67 ± 0.04 μM
$Ki_{(comp)}$ (βG1CFsP)		4.03 ± 0.03 μM

TABLE 1.10: Kinetic parameters for wild-type and 5FWβPGM with native substrate and competitive inhibitors described in (Ampaw et al., 2017).

- Empirical valence bond (EVB) calculations of the catalysis of the wild type (WT), D10N, D10S, D10C, H20A, H20Q, T16P, K76A, D170A and E169A/D170A protein variants.
- " ... calculated activation free energies confirm that D10 is likely to serve as the general base/acid for the reaction catalyzed by the WT enzyme and all its variants, in which D10 is not chemically altered." Namely, that while a substrate assisted mechanism may be a viable choice in the WT enzyme, it is "substantially higher in free energy" than the D10-assisted mechanism .
- The prediction from calculation that D10 plays a key role in both structural organization and maintaining electrostatic balance in the active site.
- The prediction that the phosphorus transfer and the proton transfer in both steps of the reaction are synchronous. This contrasts to QM predictions (Marcos, Field, and Crehuet, 2010; Elsässer, Dohmeier-Fischer, and Fels, 2012).
- Metadynamics calculations which indicated that there is a free energy difference of 0.8 ± 0.9 kcal mol^{-1} between the two rotameric states of the D10 sidechain (in favour of the “cap-open” conformation - ie. "out"), with an activation barrier of 3.1 ± 0.6 kcal mol^{-1}.
- The prediction that the increased stability of the βPGM:βG16BP (Michaelis complex) in D10 variants to the loss of charge repulsion between the phosphate and the D10 sidechain carboxylate. The authors conclude later that this electrostatic repulsion may play a role in ground state destabilization, and promotion of βG16BP intermediate release.
- "...every amino acid that makes a significant electrostatic contribution (>1.0 kcal mol^{-1}) to the calculated activation free energies appears to interact differently with the transition states for each of the two different pathways".
- The observation of pre-organisation of the enzyme active site with multiple key residues maintaining hydrogen bonds in both open and closed states.

1.6 Discussion of PGM narrative

Given a description of what are the key findings were each of these papers and the conclusions that the authors drew from them, how do all of these pieces fit together? In discussing the narrative some quotations are used to report directly what was said when key phenomena were observed and discussed for the first time, and additionally, to try and avoid misrepresentation on potentially contentious issues. Readers are strongly encouraged to look in the respective manuscripts for further context.

1.6.1 Pentavalent phosphorane

In (Lahiri et al., 2003), the authors state that they observe for the first time a pentavalent phosphorane intermediate in the active site of βPGM the vast importance of which was commented upon by the late great enzymologist Jeremy Knowles (Knowles, 2003). A response letter from Blackburn and Williams (Blackburn et al., 2003), identified several issues with the interpretation of the data. Key objections were difference map peaks in the electron density, inappropriate bond lengths around the phosphorane group, and the presence of MgF_3^- components in the crystallization conditions. This information when taken with the precedent for MgF_3^- formation in the active site of a phosphoryl transfer enzyme (Graham et al., 2002), indicated an incorrect interpretation of the data.

A rebuttal letter by the original authors in response to the letter by Blackburn and Williams was presented in the same journal (Allen and Dunaway-Mariano, 2003) and outlined 5 reasons why their initial interpretation was correct. These reasons were: 1) Bradford protein and Malachite Green phosphate assays indicated two phosphates present per washed enzyme crystal. 2) The equatorial P–O bond lengths 1.7 $\pm$ 0.1 Å in both the 1.2 Å structure and an additional 0.9 Å structure. 3) A SAD dataset contoured at 3.5 σ indicated an identical number of electrons are present at peaks corresponding to the C(1)P and C(6)P positions. 4) Crystals of the βPGM complex are formed in crystallization solutions containing 1 mM ammonium fluoride, and the authors assert that three times this concentration does not inhibit catalysis by βPGM. 5) The authors cite two examples of pentavalent phosphoranes in the literature and criticise the evidence presented in the Blackburn comment.

The presence of a MgF_3^- in the active site of βPGM was supported by computational work a year later (Webster, 2004) as well as a rationale as to how the TSA could form in the active site of the enzyme.

In 2005 Allen and Dunaway-Mariano publish a paper titled "Chemical Confirmation of a Pentavalent Phosphorane in Complex with β-Phosphoglucomutase" (Tremblay et al., 2005) where βPGM crystal are grown in the presence of α-galactose 1-phosphate and 100 mM Ammonium fluoride and 10 mM $MgCl_2$ - the conditions that had led to the previously reported phosphorane structure. The authors observe no MgF_3^- TSA and combine that observation with a Malachite green assay to defend their initial interpretation in response to the scrutiny from Blackburn and Williams.

In 2006 the Waltho group attempt to elucidate the phosphorane controversy by presenting ^{19}F NMR under very similar buffer conditions to those used to grow the controversial crystals. The ^{19}F spectrum indicated three peaks that were consistent with the MgF_3 group being

present in the active site (Baxter et al., 2006). The backbone of the complex was assigned by 3D NMR, and NOE measurements demonstrated NOE transfer between the three ^{19}F peaks and the backbone amides of several residues (D10, A115) in the active site. The only way that this could occur is in the MgF_3^- group was in the active site, demonstrating that the species observed was not a phosphorane species, it was infact the MgF_3^- TSA predicted by Blackburn (Blackburn et al., 2003). Kinetic inhibition of βPGM by magnesium and fluoride was also presented here. A more detailed inhibition study of βPGM by fluoride was presented in (Golicnik et al., 2009) and in both cases, it was observed that in the presence of Mg^{2+}, βPGM was inhibited by fluoride. Furthermore, the claim that α-galactose 1-phosphate does not form a MgF_3 TSA on addition of fluoride was disproved by ^{19}F NMR in (Baxter et al., 2009). Here the authors presented ^{19}F NMR spectra of the TSA complex, as well as $^1J_{HF}$ couplings between the fluorides of the MgF_3^- TSA and the backbone amides of residue D10 and A115 which coordinate the TSA in the active site.

The final element to the narrative is the recrystallization of βPGM in the presence of magnesium, fluoride, and G6P (Baxter et al., 2010), the initial buffer conditions that led to the reported phosphorane complex in (Lahiri et al., 2003). In the 2010 study, the authors demonstrate that MgF_3^- is sufficient to explain the observed electron density in the transferring phosphate position. These observations, coupled to the solution NMR observations effectively put an end to the debate.

1.6.2 Phospho-enzyme

In (Lahiri et al., 2002b; Lahiri et al., 2002a) the authors present a structure of a reportedly phosphorylated βPGM enzyme (phospohrylated at residue D8). The authors note that "... no phosphate, sulfate, or similar ions were included in the crystallization solution, and thus the assignment of the electron density to a phosphoryl group covalently linked to Asp8 is unequivocal."

In (Zhang et al., 2005) the authors concede that the half life of the phospho-enzyme in βPGM is less than a minute, "Thus emerged the paradox of the aspartyl group in the structure ...". The authors note that phosphate had not been directly added to the crystallization solution, however the PEG 3350 stock used in the crystallization was contaminated with phosphate as determined by Malachite Green phosphate binding assay. The authors suggest that the resulting *ca.* 0.5 mM Pi in solution (in a 2:1 ratio with protein) and "Because the active site of the crystalline enzyme is open to solvent, phosphorylation could have occurred before or after crystallization. [due to the inorganic phosphate]"

In (Baxter et al., 2010) the authors partly addressed "the remarkable hypothesis that the presence of Pi is sufficient to cause phosphorylation of D8 before or after crystallization (Zhang et al., 2005) ... [which could be] discounted by ^{31}P NMR showing the absence of an aspartyl phosphate peak in the presence of a large excess of Pi ... ". The authors go on to suggest that "It is particularly likely that the observed density is the result of the formation of an aluminum fluoride adduct of β-PGM.". The rationale was that "Although the crystallization conditions reported contained no added aluminum, the levels of fluoride used (100 mM) are sufficient to leach aluminum from laboratory glassware, as shown previously (Sternweis and Gilman, 1982; Wittinghofer, 1997)."

While this prediction seemed more plausible than spontaneous phosphorylation of the apo βPGM enzyme, it ultimately proved to be incorrect. Re-refinement of the intial structure with AlF_4 present in the active site was not sufficient to remove difference map peaks at the phosphate site. [See paper 4 for details]

1.6.3 The role of the general acid-base (GAB)

Due to the energetic favourability of step 2 complexes over step 1 complexes in the overall reaction (Fig. 1.3), the role that D10 plays as a GAB was primarily investigated using step 2 complexes in either the ground state or the transition state.

In (Lahiri et al., 2002b) the authors comment on their reportedly phosphorylated enzyme (see previous section) that given the DXDXV/T motif present in the active site, and the proposed phosphorylation of the first aspartate (D8) in the sequence: "This result suggests a novel two-base mechanism for phosphoryl group transfer in a phosphorylated sugar." Namely, they identified that D10 may play a role in phosphoryl group transfer.

In (Lahiri et al., 2003) the authors in Fig 2B through illustration indicate that a protonated D10 sidechain carboxylate coordinates the bridging oxygen between the sugar and the proposed phosphorane-intermediate complex. Thus without explicitly stating it, they suggest that D10 plays a role as a GAB. In (Webster, 2004) the author commented on the "... concomitant proton transfer from the hydroxyl group of glucose to ASP10 ..." in his QM modelling. The author concluded by stating that "Site-directed mutagenesis studies could establish the vital role of the conserved ASP10 residue as a proton acceptor/donor for the OH group of glucose."

It was not until 2009 that these mutagenesis studies were performed (Dai et al., 2009) and the authors reported that "Replacement of Asp10 with Ala, Ser, Cys, Asn, or Glu resulted in no observable activity." In this paper, the authors concluded that "Taken together, the results support a substrate induced-fit model of catalysis in which βG1P binding to the core domain facilitates recruitment of the general acid/base Asp10 to the catalytic site and induces cap closure." This was the first suggestion that the transition from "out" rotamer occupied in the open state (with D10 oriented away from the active site), to "in" (with D10 engaged in the active site), was correlated with global domain motion.

In (Griffin et al., 2012), the authors crystallize a phosphate surrogate (BeF_3^-) covalently bound to D8 in ground state complexes of βPGM with and without βG1P/G6P ligand. Here the authors observe that D10 adopts the "out" rotamer when only the BeF_3^-(phospho-enzyme surrogate) is bound, and they suggest that "Such disruption of general base catalysis should reduce the rate of hydrolysis of the phospho-enzyme resulting from the inadvertent activation of water molecules." Furthermore, the authors observe two step 2 complexes with BeF_3^- and G6P bound crystallographically, one with alignment of the G6P 1-OH hydroxyl nucleophile, and one without. This indicated that the enzyme could close (although without adopting TS geometry) without full engagement of the GAB.

In 2018 Barozzo *et al.* calculated the free energy of D10 and substrate-assisted catalysis in βPGM (Barrozo et al., 2018). The EVB calculations of the authors predict that it is more energetically favourable for D10 to act as a GAB, than the transferring phosphate, in a substrate-as-base paradigm.

Previous DFT work on the phosphoryl transfer between the phospho-enzyme and the G6P substrate (ie. the reverse direction of the step 2 reaction) presented conflicting timings for the proton transfer to the GAB, with some indicating "early" (Elsässer, Dohmeier-Fischer, and Fels, 2012), some indicating "concerted" (Kamerlin et al., 2013) and some indicating "late" proton transfer events (Marcos, Field, and Crehuet, 2010; Webster, 2004). In the first paper of this thesis, the authors observe that the D10N variant of βPGM in-fact traps the βG16BP intermediate in the active site of the enzyme, with the transferring phosphorus and nucleophilic oxygen at Van der Waals contact distance. Trapped immediately prior to proton transfer to the βG16BP intermediate, this variant allows a further investigation into the role that D10 plays in catalysis. Furthermore, the authors demonstrate that it is highly likely that the D10N variant is active, with a *ca.* 360 fold reduced k_{cat} compared to wild type enzyme.

1.6.4 (Not so) standard operating procedures – Investigations of catalysis

A source of opacity in the βPGM narrative are the range of k_{cat} values (not k_{obs} values) reported for wild type βPGM. Initial reports placed k_{cat} at 18 s^{-1} in (Lahiri et al., 2002a) and then again in (Lahiri et al., 2004) and (Zhang et al., 2005). However, k_{cat} jumps to 180 s^{-1} in (Dai et al., 2006) and (Dai et al., 2009), while k_{cat} was fitted to be 65 s^{-1} in (Golicnik et al., 2009). Several years later k_{cat} was reported to be 8 s^{-1} (Ampaw et al., 2017), and 25 s^{-1} a year later (Johnson et al., 2018). Given that these k_{cat} values for wild type enzyme vary by two orders of magnitude, it is pressing to understand why.

Initial reports of k_{cat} values in (Lahiri et al., 2004) and then again in (Zhang et al., 2005) report taking steady state values of βG1P turnover and fitting them to Michaelis-Menten kinetics. In this case α-glucose 1,6-(bis)phosphate (αG16BP) was used as a priming agent to phosphorylate βPGM at residue D8. The increase in k_{cat} in (Dai et al., 2006) and (Dai et al., 2009) followed a change of priming agent from αG16BP (in (Lahiri et al., 2004)) to βG16BP and the authors again fit the kinetic data to a Michaelis-Menten equation for k_{cat} determination. In these subsequent cases, a coupled assay was used where the product of the reaction (G6P) was converted to glucose 6-phospholactone by glucose 6-phosphate dehydrogenase (G6PDH) which reduces NAD to NADH in the process. A spectrophotometer was used to follow the formation of NADH at 340 nm.

Furthermore, [U-^{14}C]β-glucose 1-phosphate and [U-^{14}C]β-D-glucose 1,6- bisphosphate were prepared in (Dai et al., 2006) and used to follow both βG1P turnover and the dissociation and re-orientation of βG16BP from the enzyme active site. These observations were performed over a rapid timescale using a stopped flow apparatus and the radioactivity was determined by liquid scintillation counting.

2009 was the first mention of the kinetic lag phase in βPGM and is in fact the first paper to present a graph showing the formation of substrate against time (Golicnik et al., 2009). The authors characterise the equilibration of βG1P and G6P substrates by βPGM using αG16BP as a priming agent, while the inclusion of βG16BP as priming agent eliminates this lag phase. The authors fit the reaction profile to a ping-pong bi-bi reaction scheme and determine a higher k_{cat} value than was previously reported using αG16BP as a priming agent. The authors determine that as well as the inhibition by fluoride, the initial substrate βG1P - serves to inhibit catalysis and presents as a lag phase. It is surprising that 2009 is the first report of this lag phase as papers in 2004 (Lahiri et al., 2004), 2005 (Zhang et al., 2005), 2006 (Dai et al.,

2006) and 2009 (Dai et al., 2009) all report k_{cat} values derived from βPGM equilibration of βG1P with G6P observed by spectrophotometric coupled assay. Furthermore, the third paper in this thesis demonstrates that mutations to residue R49 eliminate the βG1P dependent lag phase in βPGM. Residue R49 was predicted to play a key role in ligand recognition and the R49A and R49K mutations were previously investigated in 2004, however, no comment was made about any effects on the catalytic lag phase (Lahiri et al., 2004).

With AcP as a priming agent, linear regions of a range of βG1P turnover profiles were fitted in (Johnson et al., 2018) to give k_{cat} values (Paper I). In Zhang et al., 2005 the authors observe that increasing the concentration of phosphate donors, αG16BP ,α-fructose 1,6-(bis)phosphate, and AcP as a priming agents inhibits the initial velocity of the reaction, while ATP and Pi did not either serve to activate or inhibit catalysis. The inclusion of AcP in (Johnson et al., 2018) was to ensure that the lag phase was not dependent on the rate of phosphorylation by the priming agent, and to be directly translatable to the structural observations by NMR.

In conclusion, a range of techniques have been used to determine k_{cat}, from spectrophotometric coupled assays, to stopped flow radioactivity assays, to direct observation by NMR spectroscopy. While this has allowed the reaction scheme to be investigated from a multitude of angles, some time dependent, and priming agent dependent elements have become slightly obscure over time. Ultimately, the best approach it seems is to initially reduce the bi-bi ping-pong reaction mechanism to a standard Michaelis-Menten reaction to investigate k_{cat} for the reaction (this is achieved here through the use of βG16BP as a priming agent). Following this, a more robust fit of the reaction scheme can be performed with a phosphorylating agent of known efficiency, which will allow for the investigation of feed-forward and feed-back inhibitions and activations within the system. This is especially pressing as bi-bi ping-pong reaction mechanisms are particularly susceptible to substrate inhibition of unprimed enzymes (Cornish-Bowden, 2012).

1.7 Brief review of metal fluoride TSAs in the context of protein crystallography and NMR

Excellent reviews of the use of metal fluorides to investigate phosphoryl transfer have been published very recently (Jin et al., 2017b; Jin, Molt, and Blackburn, 2017a) and the reader is directed to those texts for a comprehensive overview of the field. A short summary is be presented here that relates to the current work. Many BeF_3^-, MgF_3^-, and AlF_4^- TSAs with substrate spontaneously assemble in the active site of the enzyme, which permits crystallographic investigation. Furthermore, due both to the spin half property of the ^{19}F nucleus, and the high gyromagnetic ratio (*ca.* 94% of the gyromagnetic ratio of proton), these fluorine nuclei provide highly sensitive reporters of active site electrostatics in enzymes (Jin et al., 2016; Jin et al., 2017a). Solvent induced isotope shift (SIIS) values can be determined for fluorine resonances by replacing H_2O in the buffer with D_2O and observing the change in chemical shift of the fluorine resonance between H_2O and D_2O samples. SIIS values report on the tightness of hydrogen bonding to a particular fluoride, and can be useful in both resonance assignment and investigation of active site structure.

1.7.1 BeF_3^- complexes

Beryllium forms stable complexes with fluoride ions in solution, forming several species, including: $BeF_2 \cdot 2H_2O$, $BeF_3^- \cdot H_2O$, BeF_4^{2-}. Due to the obligate tetrahedral geometry of beryllium fluorides in solution, and the monoanionic nature of the BeF_3^- moiety, BeF_3^- has been used to probe the ground state (GS) of phosphoryl transfer processes. To date >120 BeF_3^- containing structures have been deposited in the PDB, with *ca.* 70 coordinated to an aspartyl group or carboxylate (Fig. 1.9), and *ca.* 50 coordinated to the terminal phosphate group of a nucleotide (Fig. 1.10)(Jin et al., 2017b). In both cases, water from the $BeF_3^- \cdot H_2O$ species in solution is displaced by either a carboxylate oxygen, or a phosphate oxygen and the BeF_3^- moiety serves as an additional phosphate. The stable nature of these phospho-enzyme/phospho-substrate mimics permitted the investigation of near attack complexes (NACs) in βPGM, where it was argued that two separate NACs were partially-closed energetic minima on-pathway to catalysis (See (Griffin et al., 2012) and references within). In these cases the nucleophilic hydroxyl group was either aligned to hydrogen bond to one of the fluorides (non-productive NAC1), or aligned to attack the beryllium atom (productive NAC2).

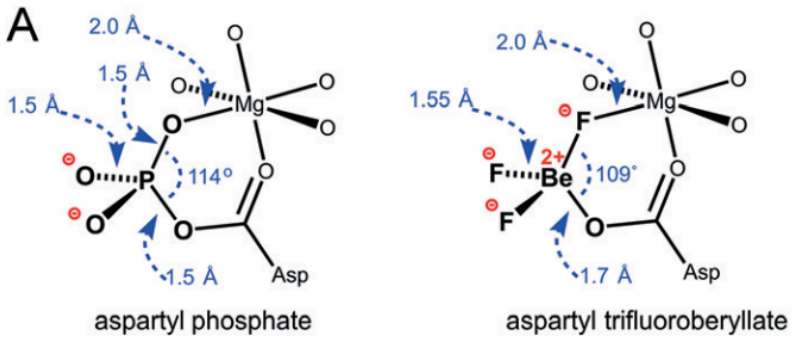


FIGURE 1.9: Comparison of aspartyl phosphate and aspartyl trifluoroberyllate adapted from (Jin et al., 2017b)

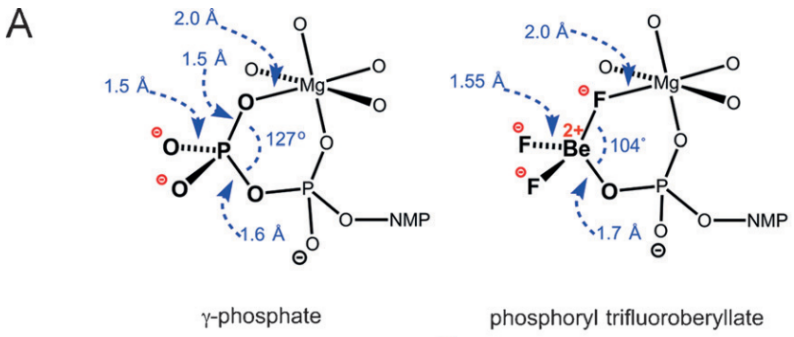


FIGURE 1.10: Comparison of gamma-phosphate and phosphoryl trifluoroberyllate adapted from (Jin et al., 2017b)

1.7.2 AlF_4^- complexes

Aluminium forms stable complexes with fluoride ions in solution, forming several species, including: $AlF_2^+ \cdot 4H_2O$, $AlF_3 \cdot 3H_2O$, $AlF_4^- \cdot 2H_2O$, and $AlF_5^{2-} \cdot H_2O$ with a fluoride concentration dependence (Jin, Molt, and Blackburn, 2017a). AlF_4^- adopts tetrahedral geometry, and has been used to investigate the transition state of phosphoryl transfer reactions, when complexed with ligand in the active site of enzymes. One of the first demonstrations of the use of AlF_4^- was the discovery that AlF_4 was leached out of glassware and activated small G proteins in the presence of GDP (Sternweis and Gilman, 1982). There are *ca.* 100 structures of AlF_4^- complexed with nucleotides in the PDB, and *ca.* 15 structures complexed with a carboxylate group (Jin et al., 2017b). The stable nature of this these complexes permitted structural investigation of the transition state of enzyme catalysis by trapping these metal-fluoride transition state analogues (TSAs) in the active site of enzymes. The AlF_4^- moiety is square planar, thus not isosteric with the transferring phosphoryl moiety (trigonal planar). Furthermore, the Al–F bond is slightly longer than the native P–O bond (Jin et al., 2017b), however, the AlF_4^- group is isoelectronic with the transferring phosphoryl moiety. This property was demonstrated to be of great importance when metal fluorides were binding to the active site of βPGM as pH > 7.5 causes a key transition between AlF_4^- and AlF_3, which is replaced by another metal fluoride moiety, magnesium trifluoride. Distinct ^{19}F chemical shift differences between AlF_4^- and MgF_3^- species permit differentiation, often with sharp, well-resolved, NMR signals that respond to SIIS investigation.

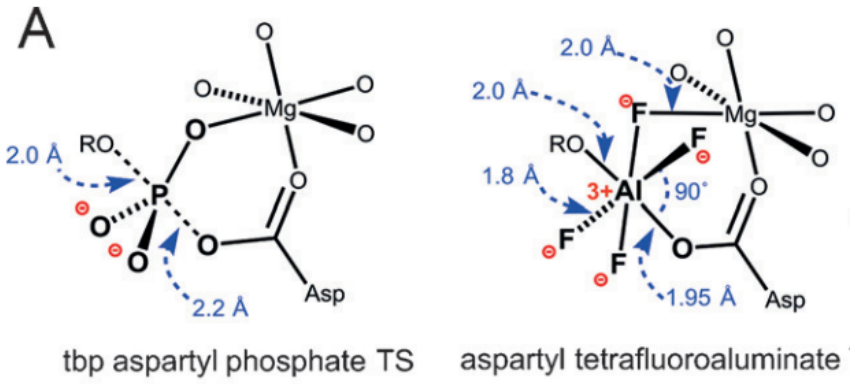


FIGURE 1.11: Comparison of trigonal bipyramidal aspartyl phosphate and aspartyl tetrafluoroaluminate adapted from (Jin et al., 2017b)

1.7.3 MgF_3^- complexes

Magnesium does not form stable complexes with fluoride ions in water. The presence of MgF_3^- was first predicted based on magnesium-dependent fluoride-inhibition studies, and the first observation crystallographically in a 1.8 Å structure of RhoA/RhoGAP (Graham et al., 2002). MgF_3^- is highly useful for the investigation of phosphoryl transfer as it is both isosteric and isoelectronic with a transferring phosphoryl moiety (Fig. 1.12). However, the MgF_3^- moiety has to assemble in the active site of the enzyme (a process that may take on the timescale of minutes (Golicnik et al., 2009)), which often leads to a reduced potency of inhibition compared to the pre-assembled AlF_4^- moiety.

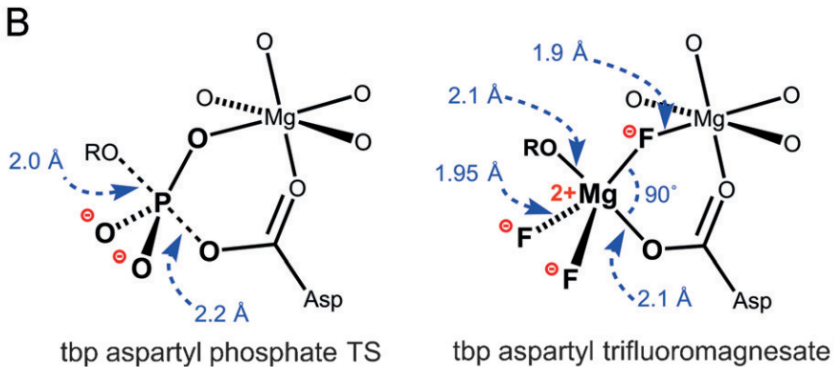


FIGURE 1.12: Comparison of trigonal bipyramidal aspartyl phosphate and aspartyl trifluoromagnesate adapted from (Jin et al., 2017b)

1.7.4 Other metal fluoride complexes

While anomalies such as AlF_3 have been reported in the PDB, often trigonal planar moieties are later demonstrated to be MgF_3^- groups by ^{19}F NMR (Jin et al., 2017b). In some cases higher (or lower) metal coordinations by fluoride have been reported, thorough discussions of these cases are presented in Jin et al., 2017b.

This protein isn't dead, it's pining for the Fjords . . .

Chapter 2

Theory

2.1 NMR spectroscopy

This following is based on thorough treatments of NMR theory and practice in several excellent texts by James Keeler (Keeler, 2011), Jon Cavanagh *et al.* (Cavanagh et al., 2007), Edward d'Auvergne (d'Auvergne, 2006), and Bertil Halle (Halle, 2009).

2.1.1 Nuclear spin and magnetic moment

The nuclear spin angular momentum **I** of a nucleus is a quantum mechanical property of the nucleus, and not the classical mechanical property of spin. **I** is a vector with magnitude given by:

$$|\,\mathbf{I}\,| = [\mathbf{I}\cdot\mathbf{I}]^{\frac{1}{2}} = \hbar[I(I+1)]^{\frac{1}{2}} \tag{2.1}$$

where I is the nuclear spin angular momentum quantum number and $\hbar$ is Planck's constant over 2π.

Nucleus	I	γ (T s)$^{-1}$	Natural abundance (%)
^{1}H	1/2	2.6752 x 10^{8}	99.99
^{2}H	1	4.107 x 10^{7}	0.012
^{13}C	1/2	6.728 x 10^{7}	1.07
^{15}N	1/2	-2.713 x 10^{7}	0.37
^{19}F	1/2	2.518 x 10^{8}	100.00
^{31}P	1/2	1.0839 x 10^{8}	100.00

TABLE 2.1: Properties of selected nuclei highly relevant to the study of biological systems. I is the nuclear spin angular momentum quantum number, γ is the magnetogyric ratio of the nucleus.

When an external magnetic field is applied along the z-axis (eg. via a very expensive NMR spectrometer), the Z- component of this angular momentum is given by:

$$I_z = \hbar m \tag{2.2}$$

where m is the magnetic quantum number that exists as m = (*-I, -I+1, ..., I-1, I*), thus I_z has $2I$+1 possible values. The nuclear magnetic moment μ is collinear with I_z as a consequence

of the Wigner-Eckart theorem (Cavanagh et al., 2007) and the Z- component is given by:

$$\mu_z = \gamma I_z = \gamma \hbar m \tag{2.3}$$

When μ is projected onto the magnetic field vector **B**, conventionally defined as being applied along the z-axis in the laboratory frame, the spin states become quantized with energies proportional to their projection onto **B**:

$$E^{\mathrm{m}} = \gamma I_z B_0 = -m\hbar\gamma B_0 \tag{2.4}$$

where B^0 is the field strength of the applied magnetic field. This results in 2I +1 equally spaced Zeeman levels in the z-axis of the laboratory frame. At equilibrium these states are unequally populated and the relative populations are given by the Boltzmann distribution:

$$\frac{N_m}{N} = \exp\left(\frac{-E_m}{k_B T}\right) \Big/ \sum_{m=-I}^{I} \exp\left(\frac{-E_m}{k_B T}\right) \approx \frac{1}{2I+1}\left(1 = \frac{m\hbar\gamma B_0}{k_B T}\right) \tag{2.5}$$

where N_m is the number of spins in state m and N is the total number of spins. Energy required to stimulate a transition between the m and m + 1 Zeeman states is given by:

$$\Delta E = \hbar\gamma B_0 \tag{2.6}$$

For an isolated spin 1/2 nucleus, only two nuclear spin states exist at two energy levels, α and β, separated by $\Delta E = \hbar\gamma B_0$. When m = +1/2 this is referred to as the α state, and m = -1/2 is referred to as the β state. If γ is positive (see Table 2.1) then the α state has lower energy than the β, if it is negative, then the relative energies are reversed.

The larmor frequency of a nucleus is given by:

$$\omega_0 = -\gamma B_0 \tag{2.7}$$

where B_0 is the strength of the static magnetic field. The angular frequency needed to induce a transition between Zeeman energy levels is (in rad s^{-1}):

$$\omega = \frac{\Delta E}{\hbar} = \gamma B_0 \tag{2.8}$$

2.1.2 Chemical shift

Resonance frequency (ω) at a given static field strength is given by:

$$\omega = -\gamma(1-\sigma)B_0 \tag{2.9}$$

where σ is the average isotropic shielding constant for the nucleus:

$$\sigma = \frac{\sigma_{11} + \sigma_{22} + \sigma_{33}}{3} \tag{2.10}$$

Chemical shift anisotropy (CSA) results from anisotropic shielding of a nucleus is given by:

$$\Delta\sigma = \frac{\sigma_{11} - (\sigma_{22} + \sigma_{33})}{2} \tag{2.11}$$

where the asymmetry (η) can be defined as:

$$\eta = \frac{3(\sigma_{22} - \sigma_{33})}{2\Delta\sigma} \tag{2.12}$$

The chemical shift of a nucleus is given by:

$$\delta = \frac{\Omega - \Omega_{ref}}{\omega_0} \times 10^6 = (\sigma_{ref} - \sigma) \tag{2.13}$$

where ω is the offset, and Ω_{ref} is the reference offset.

2.1.3 Linewidth

Equations for the absorbtion mode of a Lorentzian lineshape.

Lineshape (rad s^{-1})

$$A(\omega) = \frac{R_2}{R_2^2 + (\omega - \Omega)^2} \tag{2.14}$$

Peak height:

$$\frac{1}{R_2} \tag{2.15}$$

Peak width (rad s^{-1}):

$$W = 2R_2 \tag{2.16}$$

Peak width (Hz):

$$W = \frac{R_2}{\pi} = \frac{1}{\pi T_2} \tag{2.17}$$

Transverse relaxation, or R_2 rates are proportional to the overall correlation time of the protein, τ_c. This correlation time can be approximated using Stokes' law:

$$\tau_c = \frac{4\pi\eta_w r_H^3}{3k_B T} \tag{2.18}$$

Rotational correlation times in D_2O are reportedly *ca.* 25% greater than in H_2O due to a larger viscosity of D_2O (Cavanagh et al., 2007).

2.1.4 Chemical exchange

Another source of linebroadening of NMR peaks is exchange of the nucleus between two different chemical environments. This is termed chemical exchange and is typically modelled as a two site exchange process:

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B \tag{2.19}$$

for a two state process:

$$\begin{aligned} \frac{k_1}{k_{-1}} &= \frac{p_A}{p_B} \\ p_A + p_B &= 1 \\ k_{ex} &= k_1 + k_{-1} \end{aligned} \tag{2.20}$$

The average frequency will be population averaged according to:

$$\omega_{Avg} = p_A\omega_A + p_B\omega_B \tag{2.21}$$

If $k_{ex} << \Delta\omega$ then two distinct peaks will be observed (assuming $\Delta\omega >$ FID resolution). As k_{ex} tends towards $k_{ex} = \Delta\omega$ the two resonances will coalesce into a single broad peak and as k_{ex} tends towards $k_{ex} >> \Delta\omega$, a single sharp peak will emerge. While many processes conform to a two site exchange regime, higher order exchange processes can be modelled by modification of the Bloch equations (Keeler, 2011).

2.2 NMR spin relaxation

The Schrödinger equation describes the time dependent evolution of a quantum mechanical system:

$$\frac{\partial\Psi(t)}{\partial t} = -\frac{i}{\hbar}\mathscr{H}\Psi(t) \tag{2.22}$$

and the Hamiltonian operator of the system $\mathscr{H}$ incorporates the evolution of the system (not necessarily in a time dependent manner). These Hamiltonian operators are ideal for modelling the time evolution of the quantum mechanical spin interaction between two nuclei such as a ^{1}H and ^{15}N nuclei in a backbone amide, or between several ^{1}H nuclei separated in space.

2.2.1 Longitudinal dipolar relaxation of two spins

A full derivation of the Master equation for spin relaxation is beyond the scope of this text, however an excellent treatment is presented by Goldman (Goldman, 2001), and by Cavanagh (Cavanagh et al., 2007). For this text, a brief recap of the Solomon equations will be given that describes the longitudinal dipolar relaxation of two coupled spins as presented in (Keeler, 2011) if the rate constants $\Delta^{(1)}$ and $\Delta^{(2)} = 0$ in Fig 2.1.

$$\frac{dI_{1z}}{dt} = -R_z^{(1)}(I_{1z} - I_{1z}^0) - \sigma_{12}(I_{2z} - I_{2z}^0) \tag{2.23}$$

$$\frac{dI_{2z}}{dt} = -\sigma_{12}(I_{1z} - I_{1z}^0) - R_z^{(2)}(I_{2z} - I_{2z}^0) \tag{2.24}$$

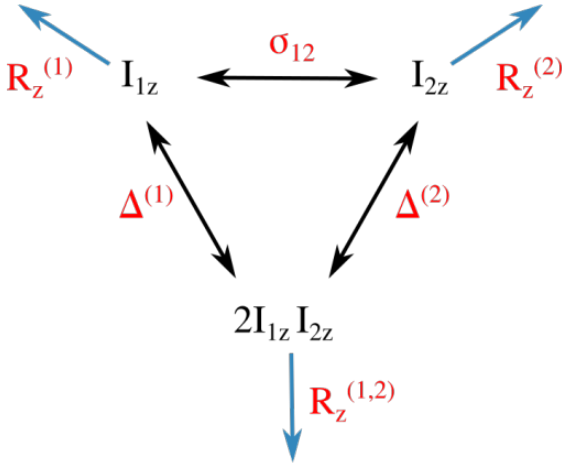


FIGURE 2.1: Dipolar relaxation of z-magnetization between different pathways of a two spin system. Blue arrows represent autorelaxation rates, whereas the black arrows denote terms connecting the two spins. The cross relaxation term (σ_{12}) transfers relaxation between the two nuclei.

$$\frac{d2I_{1z}dI_{2z}}{dt} = -R_z^{(1,2)}2I_{1z}I_{2z} \tag{2.25}$$

These rate constants simplify to:

$$R_z^{(1)} = 2W_1^{(1)} + W_2 + W_0 \tag{2.26}$$

$$R_z^{(2)} = 2W_1^{(2)} + W_2 + W_0 \tag{2.27}$$

$$\sigma_{12} = W_2 - W_0 \tag{2.28}$$

$$R_z^{(1,2)} = 2W_1^{(1)} + 2W_1^{(2)} \tag{2.29}$$

Where $W_1^{(i)}$ denotes a single quantum transition for nucleus i, W_2 indicates a double quantum transition, and W_0 denotes a zero quantum transition.

The reduced spectral density $j(\omega_{ij})$ at the frequency of the transition between the two energy levels is a measure of the amount of random motion at the Larmor frequency sufficient to induce transitions between energy levels (See section 2.4.1). For dipolar relaxation between two spins

$$W_1^{(1)} = \frac{3}{40}b^2 j(\omega_{0,1}) \tag{2.30}$$

$$W_1^{(2)} = \frac{3}{40}b^2 j(\omega_{0,2}) \tag{2.31}$$

$$W_2 = \frac{3}{10} b^2 j(\omega_{0,1} + \omega_{0,2}) \tag{2.32}$$

$$W_0 = \frac{1}{20} b^2 j(\omega_{0,1} - \omega_{0,2}) \tag{2.33}$$

$$b = \frac{\mu_0 \gamma_1 \gamma_2 \hbar}{4\pi r^3} \tag{2.34}$$

Using these expressions for the various rate constants, the equations 2.26, 2.27, 2.28, 2.29, can be rewritten:

$$R_z^{(1)} = b^2 \left[\frac{3}{20} j(\omega_{0,1}) + \frac{3}{10} j(\omega_{0,1} + \omega_{0,2}) + \frac{1}{20} j(\omega_{0,1} - \omega_{0,2}) \right] \tag{2.35}$$

$$R_z^{(2)} = b^2 \left[\frac{3}{20} j(\omega_{0,2}) + \frac{3}{10} j(\omega_{0,1} + \omega_{0,2}) + \frac{1}{20} j(\omega_{0,1} - \omega_{0,2}) \right] \tag{2.36}$$

$$\sigma_{12} = b^2 \left[\frac{3}{10} j(\omega_{0,1} + \omega_{0,2}) - \frac{1}{20} j(\omega_{0,1} - \omega_{0,2}) \right] \tag{2.37}$$

$$R_z^{(1,2)} = b^2 \left[\frac{3}{20} j(\omega_{0,1}) + \frac{3}{20} j(\omega_{0,2}) \right] \tag{2.38}$$

2.2.2 Transverse relaxation of two spins

There are many possible sources that contribute to the transverse relaxation rate (R_2). Typically, R_2 is reflected in the lineshape of the Fourier transformed signal for a population of spins according to Eq. 2.17. Linebroadening of a particular resonance often occurs if the spin is undergoing a chemical exchange process with rate k_{ex}. The nature of the chemical exchange is described as slow exchange if $k_{ex} << \Delta_{AB}$ or fast exchange if $k_{ex} >> \Delta_{AB}$ where Δ_{AB} is the frequency difference between the two exchanging peaks in Hz.

Fluctuating internal fields also provide a source of relaxation that contributes to R_2 as well as fluctuating external fields as a result of poor shimming of the magnet. Chemical shift anisotropy (CSA) also plays a key role in relaxation in a similar way to the fluctuating internal fields , although at a different timescale. It is often appropriate to treat the shielding tensor of NH and CH bond vectors as an axially symmetric tensor about the XH bond vector (ie. $\sigma_{||} > \sigma_{\perp}$).

The relaxation rate constant for an axially symmetric shielding tensor is given by:

$$R_{xy} = c^2 \left[\frac{4}{25} j(0) + \frac{1}{30} j(\omega_0) \right] \tag{2.39}$$

$$R_z = c^2 \frac{1}{15} j(\omega_0) \tag{2.40}$$

$$c = \gamma B_0(\sigma_{||} - \sigma_{\perp}) \tag{2.41}$$

2.2.3 Cross correlation

Cross correlation contributes to transverse relaxation of a backbone amide ^{15}N spin pair which has significant implications to protein NMR spectroscopy. In the slow motion limit (ie. when $R_{xy} \neq R_z$) the relaxation rate constants differ for each line of the ^{15}N doublet according to:

$$line1 : \frac{1}{10}b^2j(0) + \frac{2}{45}c_1^2j(0) + \frac{2}{15}c_1bP_2(cos\theta)j(0) \tag{2.42}$$

$$line2 : \frac{1}{10}b^2j(0) + \frac{2}{45}c_1^2j(0) - \frac{2}{15}c_1bP_2(cos\theta)j(0) \tag{2.43}$$

where b and c (for spin 1) are given by Eq. 2.34 and 2.41 respectively. $P_2(cos\theta)$ is the second order Legendre polynomial given by:

$$P_2(cos\theta) = \frac{1}{2}(3cos^2\theta - 1) \tag{2.44}$$

Notably, it is the different sign of the cross correlation term $\frac{2}{15}c_1bP_2(cos\theta)j(0)$ in *line2* that causes the cross relaxation to be subtracted from the resulting linewidth, rather than add to it as in *line1*. Physically, this means that in the case of *line2*, the random fields from the dipolar and CSA interactions are correlated in such a way that they cancel which gives rise to a phenomenon called the 'TROSY' effect. Transverse relaxation optimised spectroscopy (TROSY) takes advantage of this differential linewidth through line-selective transfer schemes, which allows significant resolution enhancements and the possibility of investigating both large and disordered proteins by NMR.

2.3 Spectral density mapping

In order to investigate high frequency dynamics in proteins (ps-ns), it is typical to perform a combination of relaxation and NOE experiments at multiple fields to characterise the spectral density function of backbone amide and sidechain methyl groups. Characterization of the spectral density functions using this approach permits the fitting of order parameters (S^2) which describe the rigidity of a given residue on the ps-ns timescale.

2.3.1 Backbone amide relaxation

^{15}N relaxation rate constants for ^{1}H-^{15}N spin pairs can be presented in terms of their respective larmor frequencies:

$$R_1 = \left(\frac{d^2}{4}\right)[3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N)] + c^2J(\omega_N) \tag{2.45}$$

$$R_2 = \left(\frac{d^2}{8}\right)[4J(0) + 3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \left(\frac{c^2}{6}\right)[4J(0) + 3J(\omega_N)] + R_{ex} \quad (2.46)$$

$$\sigma_{NH} = \left(\frac{d^2}{4}\right)(-J(\omega_H - \omega_N) + 6(\omega_H + \omega_N)) \quad (2.47)$$

$$NOE = 1 + \frac{\gamma_H}{\gamma_N}\frac{\sigma_{NH}}{R_1} \quad (2.48)$$

$$d = \frac{\mu_0 h \gamma_N \gamma_H}{8\pi^2 r_{NH}^3} \quad (2.49)$$

$$c = \frac{\Delta\sigma \gamma_N B_0}{\sqrt{3}} \quad (2.50)$$

where d is the dipolar coupling constant, and c is the CSA coupling constant. μ_0 is the permeability of free space, r_{NH} is the average amide bond length (1.02 Å), $\Delta\sigma$ is the amide CSA (held at -172 ppm in relax), ω_N and ω_H are the larmor frequencies at a given static magnetic field, and J(ω) is the spectral density function.

Using a reduced spectral density mapping approach, these expressions can be converted:

$$\Gamma_{auto} = R_2 - 0.5R_1 - 0.454\sigma_{NH} = J(0)(3d^2 + 4c^2)/6 \quad (2.51)$$

$$J(\omega_N) = \frac{R_1 - 1.249\sigma_{NH}}{(3d^2/4 + c^2)} \quad (2.52)$$

$$J(0.870\omega_H) = \frac{4\sigma_{NH}}{5d^2} \quad (2.53)$$

2.3.2 Sidechain deuterium relaxation

Relaxation rates for methyl groups in sidechains are given by the equations:

$$R_Q(D_z) = \frac{3}{40}\left(\frac{e^2qQ}{\hbar}\right)^2[J(\omega_D) + 4J(2\omega_D)] \quad (2.54)$$

$$R_Q(3D_z - 2) = \frac{3}{40}\left(\frac{e^2qQ}{\hbar}\right)^2[3J(\omega_D)] \quad (2.55)$$

$$R_Q(D_+) = \frac{1}{80}\left(\frac{e^2qQ}{\hbar}\right)^2[9J(0) + 15J(\omega_D) + 6J(2\omega_D)] \quad (2.56)$$

$$R_Q(D_+D_z + D_zD_+) = \frac{1}{80}\left(\frac{e^2qQ}{\hbar}\right)^2 [9J(0) + 3J(\omega_D) + 6J(2\omega_D)] \tag{2.57}$$

$$R_Q(D_+^2) = \frac{3}{40}\left(\frac{e^2qQ}{\hbar}\right)^2 [J(\omega_D) + 2J(2\omega_D)] \tag{2.58}$$

Where $\frac{e^2qQ}{h}$ is the quadrupolar coupling constant and ω_D is the deuterium resonance frequency.

Spectral density functions that describe the dynamics of methyl groups have been derived in a similar manner to those for backbone amides (Millet et al., 2002; Skrynnikov, Millet, and Kay, 2002):

$$J(\omega) = \frac{1}{9}\frac{S_f^2\tau_c}{(1+\omega^2\tau_c^2)} + \left(1 - \frac{1}{9}S_f^2\right)\frac{\tau_e}{(1+\omega^2\tau_e^2)} \tag{2.59}$$

More complex spectral density functions describing nanosecond motions similar to the transition between Equations 2.68 and 2.71 were derived:

$$J(\omega) = \frac{1}{9}\frac{S_f^2\tau_{c,eff}}{(1+\omega^2\tau_{c,eff}^2)} + \left(1 - \frac{1}{9}S_f^2\right)\frac{\tau}{(1+\omega^2\tau^2)} \tag{2.60}$$

2.4 Model Free Analysis

A robust explanation and derivation of both the original "Lipari-Szabo" modelfree (MF) equation (Lipari and Szabo, 1982a; Lipari and Szabo, 1982b) and the extended MF equation by Clore and coworkers (Clore et al., 1990) are presented by Halle (Halle, 2009). A brief description is presented here to illustrate how the internal and global tumbling of a protein relate to the larmor frequency, and thus affect the relaxation process.

2.4.1 Global correlation time

Some preliminaries:

The Fourier transform of global correlation function ($G(\tau)$) is the spectral density, $J(\omega)$. The amount of motion at the Larmor frequency (ω) is found by evaluating $J(\omega)$ when $\omega = \omega_0$. This global correlation function can be normalized thus,

$$g(\tau) = \frac{G(t) - G(\infty)}{G(0) - G(\infty)} \tag{2.61}$$

and the Fourier transform of this reduced correlation function ($g(\tau)$) is the reduced spectral density ($j(\omega)$).

The correlation function of a spherical top is typically defined:

$$G(\tau) = \frac{1}{5}\exp\left(\frac{-\tau}{\tau_c}\right) \tag{2.62}$$

The spectral density function of this spherical top is thus:

$$J(\omega) = \frac{2}{5}\frac{\tau_c}{(1+\omega^2\tau_c^2)} \tag{2.63}$$

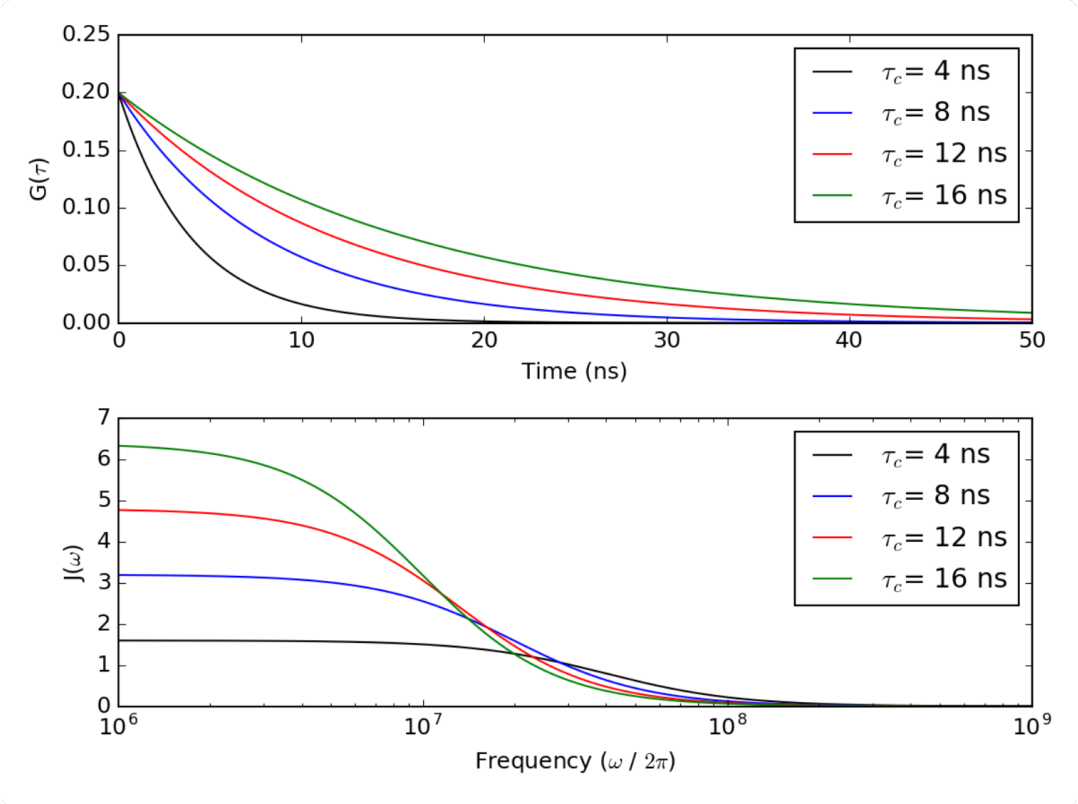


FIGURE 2.2: Time correlation function (top) and spectral density function (bottom) of a spherical top with varied correlation times.

If there is time scale separation between a global correlation and a local correlation, then the correlation function for a residue is described:

$$G(\tau) = G_O(\tau)G_L(\tau) \tag{2.64}$$

This G_O is described as in 2.62, whereas the local correalation function G_L is described as:

$$G_L(\tau) = \frac{4\pi}{5}\sum_{m,m'=-2}^{2}\langle Y_{2m}(\theta_a(0),\psi_a(0))Y^*_{2m'}(\theta_b(\tau),\psi_b(\tau))\rangle \tag{2.65}$$

where $Y_{2m}(\theta,\psi)$ correspond the the spherical harmonics which describe the angular fluctuations of the angular momentum. To derive limits of the correlation function when $\tau = 0$ and when $\tau = \infty$, $G_L(0)$ is defined as the Legendre polynomial $P_2(cos\theta_{ab})$, where θ_{ab} is the angle between the two interaction vectors:

$$G_L(\infty) = \frac{4\pi}{5}\sum_{m,m'=-2}^{2}\langle Y_{2m}(\theta_a,\psi_a)Y^*_{2m'}(\theta_b,\psi_b)\rangle = S^2_{ab} \tag{2.66}$$

Here, the modelfree parameter S^2_{ab} is introduced. The autocorrelation function combing both global and local correlations is thus:

$$G(\tau) = \frac{1}{5} \exp\left(-\frac{\tau}{\tau_c}\right)\left(S^2 + (1-S^2)\exp\left(-\frac{\tau}{\tau_e}\right)\right) \tag{2.67}$$

The Fourier transform of eq. 2.67 gives the classic Lipari-Szabo model-free spectral density function:

$$J(\omega) = \frac{2}{5}\left(\frac{S^2\tau_c}{(1+\omega^2\tau_e^2)} + \frac{(1-S^2)\tau_e}{(1+\omega^2\tau_e^2)}\right) \tag{2.68}$$

where:

$$\frac{1}{\tau_e} = \frac{1}{\tau_c} + \frac{1}{\tau_i} \tag{2.69}$$

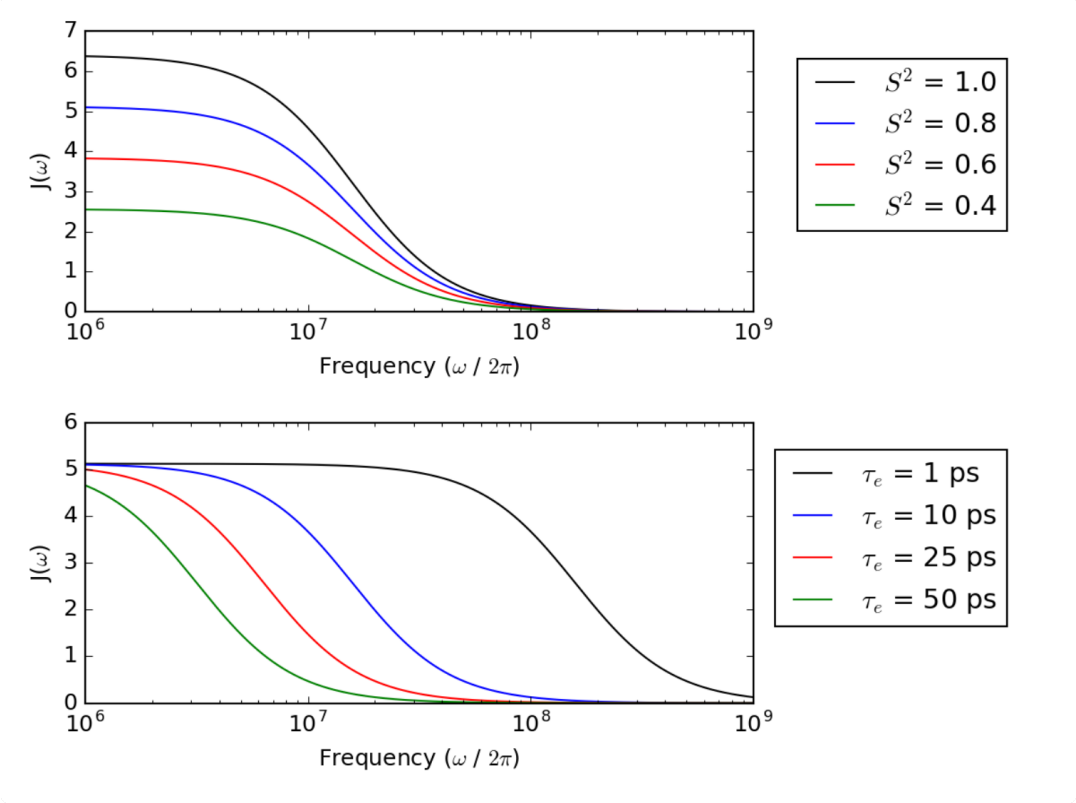


FIGURE 2.3: Spectral density function of a spherical protein using the modelfree equation Eq. 2.68 proposed in (Lipari and Szabo, 1982a) while varying the order parameter (top) and internal correlation time (τ_e ; bottom). τ_c was fixed at 16 ns in both cases. τ_e was fixed at 10 ns in the first figure while S^2 was varied, while S^2 was fixed at 0.8 in the second figure while τ_e was varied.

The original model-free approach has been expanded by Clore and co-workers (Clore et al., 1990) to include two separate internal motions. The correlation function is thus:

$$G(\tau) = \frac{1}{5} \exp\left(-\frac{\tau}{\tau_c}\right)\left(S^2 + (1-S_f^2)\exp\left(-\frac{\tau}{\tau_f}\right) + (S_s^2 - S^2)\exp\left(-\frac{\tau}{\tau_s}\right)\right) \tag{2.70}$$

The Fourier transform of this expanded correlation function gives:

$$J(\omega) = \frac{2}{5}\left(\frac{S^2\tau_c}{(1+\omega^2\tau_e^2)} + \frac{(1-S_f^2)\tau_e}{(1+\omega^2\tau_e^2)} + \frac{(S_f^2-S^2)\tau_Q}{(1+\omega^2\tau_Q^2)}\right) \tag{2.71}$$

With:

$$\frac{1}{\tau_Q} = \frac{1}{\tau_c} + \frac{1}{\tau_s}\frac{1}{\tau_e} \tag{2.72}$$

2.4.2 The diffusion tensor

A perfect spherical protein has the correlation time:

$$\tau_c = \frac{1}{6D^R} \tag{2.73}$$

where D^R is the overall tumbling coefficient. However most proteins deviate from isotropic tumbling.

2.4.3 Diffusion as an ellipsoid

This can be modelled using the generic Brownian diffusion NMR correlation function presented by d'Auvergne (d'Auvergne and Gooley, 2006).

$$C(\tau) = \frac{1}{5}\sum_{i=-k}^{k} c_i \cdot e^{-\tau/\tau_i} \tag{2.74}$$

where the summation index i ∈ -2, -1, 0, 1, 2. An ellipsoid diffusion tensor can be defined by the parameter set ($\mathscr{D}_{iso}$, $\mathscr{D}_a$, $\mathscr{D}_r$, α, β, γ) where the variable k in Eq. 2.74 is equal to 2. $\mathscr{D}_{iso}$, $\mathscr{D}_a$, $\mathscr{D}_r$ are geometric parameters, while α, β, γ, correspond to Euler angles that use the typical z-y-z rotation. The geometric parameters are defined as:

$$\mathscr{D}_{iso} = \frac{1}{3}(\mathscr{D}_x + \mathscr{D}_y + \mathscr{D}_z) \tag{2.75}$$

$$\mathscr{D}_a = \mathscr{D}_z - \frac{1}{2}(\mathscr{D}_x + \mathscr{D}_y) \tag{2.76}$$

$$\mathscr{D}_r = \frac{\mathscr{D}_y - \mathscr{D}_x}{2\mathscr{D}_a} \tag{2.77}$$

The five weights c_i are defined as:

$$c_{-2} = \frac{1}{4}(f-g) \tag{2.78}$$

$$c_{-1} = 3\delta_y^2\delta_z^2 \tag{2.79}$$

$$c_0 = 3\delta_x^2\delta_z^2 \tag{2.80}$$

$$c_1 = 3\delta_x^2\delta_y^2 \tag{2.81}$$

$$c_2 = \frac{1}{4}(f+g) \tag{2.82}$$

where:

$$f = 3(\delta_x^4\delta_y^4\delta_z^4) - 1 \tag{2.83}$$

$$g = \frac{1}{\mathscr{R}}\left[(1+3\mathscr{D}_r)(\delta_x^4 - 2\delta_y^2\delta_z^2) + (1-3\mathscr{D}_r)(\delta_y^4 - 2\delta_x^2\delta_z^2) - 2(\delta_z^4 - 2\delta_x^2\delta_y^2)\right] \tag{2.84}$$

$$\mathscr{R} = \sqrt{1+3\mathscr{D}_r^2} \tag{2.85}$$

The five correlation times are:

$$\frac{1}{\tau_{-2}} = 6\mathscr{D}_{iso} - 2\mathscr{D}_a\mathscr{R} \tag{2.86}$$

$$\frac{1}{\tau_{-1}} = 6\mathscr{D}_{iso} - \mathscr{D}_a(1+3\mathscr{D}_r) \tag{2.87}$$

$$\frac{1}{\tau_0} = 6\mathscr{D}_{iso} - \mathscr{D}_a(1-3\mathscr{D}_r) \tag{2.88}$$

$$\frac{1}{\tau_1} = 6\mathscr{D}_{iso} + 2\mathscr{D}_a \tag{2.89}$$

$$\frac{1}{\tau_2} = 6\mathscr{D}_{iso} + 2\mathscr{D}_a\mathscr{R} \tag{2.90}$$

2.4.4 Model optimization

Model optimization in modelfree analysis programs such as *relax* (Section 3.3.5) aims to reduce the chi-square target function:

$$\chi^2(\theta) = \sum_{i=1}^{n} \frac{(R_{exp} - R_{calc})^2}{\sigma^2} \tag{2.91}$$

in which i is the summation index, R_{exp} are the experimental data, R_{calc} are the back calculated data, and σ is the experimental error. Initially the defusion tensor is fixed while

different local models are optimized for each residue. First models tm0 to tm9 are optimized, failed models are then eliminated (d'Auvergne and Gooley, 2006) with AIC model selection (Akaike, 1998) used to select the best model for each residue. In d'Auvergne and Gooley, 2007, the authors demonstrate that in a standard grid search, the grid point with the lowest χ^2 value may be different from the local minimum. Therefore a more robust grid search method is implemented in relax.

Once initial parameters have been optimized, the local τ_m parameter is eliminated, MF parameters are held fixed, and the global diffusion parameters are optimized. Convergence is defined in relax as obtaining identical χ^2 values on sequential runs, although in practice, limiting the number of iterations to 30 was sufficient to obtain convergence to several decimal places. Sphere, oblate, prolate, and ellipsoid diffustion tensors are optimised using the same iterative procedure as for the model selection above, until a universal best fit of the data is achieved. This procedure is outlined in Fig. 2.4, adapted from the *relax* manual.

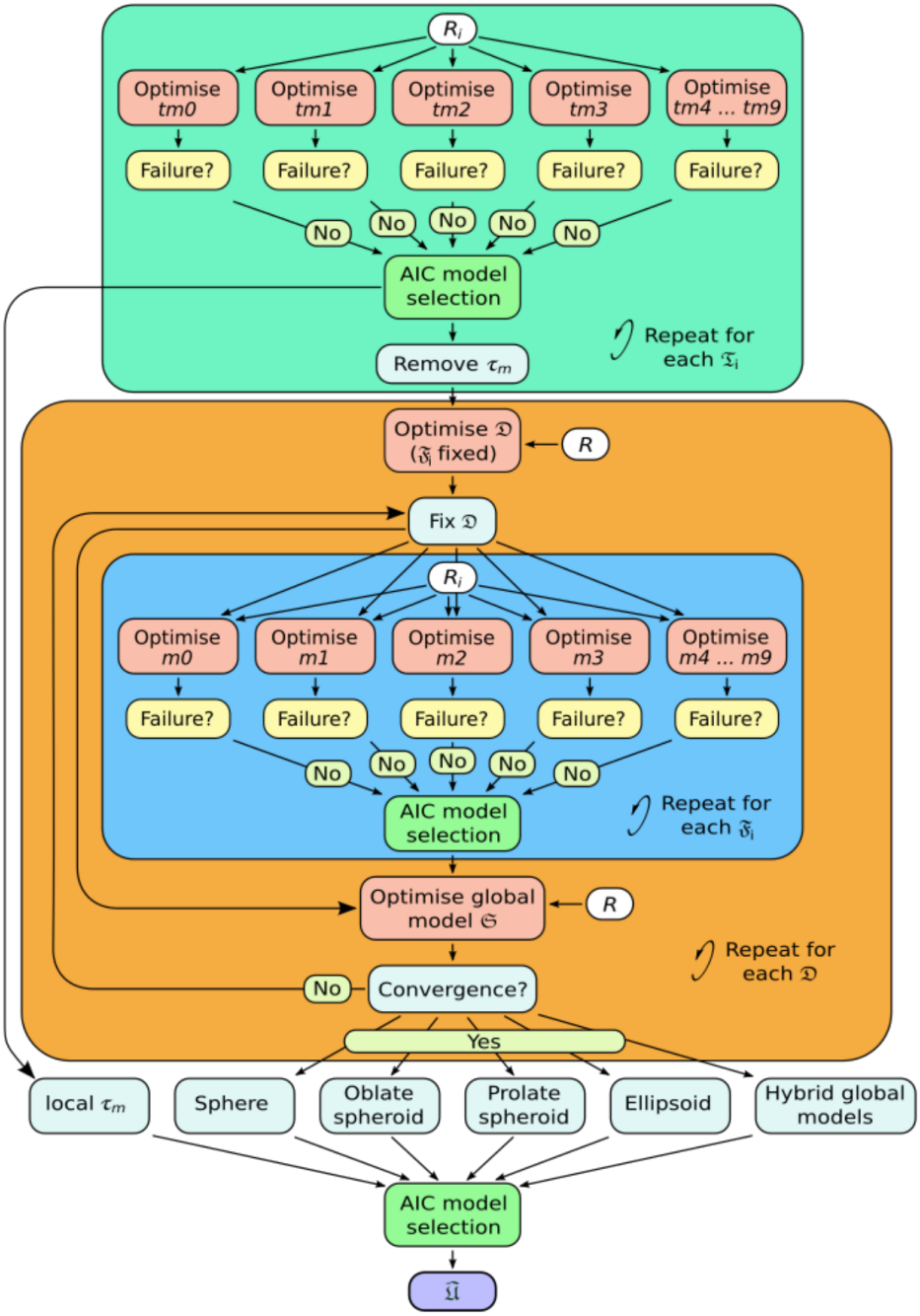


FIGURE 2.4: Modelfree optimization protocol implemented in relax v.4.0.3 adapted from the relax manual pp.102. First models tm0 to tm9 are optimized, failed models are then eliminated (d'Auvergne and Gooley, 2006) with AIC model selection (Akaike, 1998) used to select the best model. Once initial parameters have been optimized, the local τ_m parameter is eliminated, MF parameters held fixed, and the global diffusion parameters are optimized. Convergence is defined in relax as obtaining identical χ^2 values on sequential runs, although in practice, limiting the number of iterations to 30 was sufficient to obtain convergence to several decimal places. The universal solution at the bottom of the optimization protocol is determined using AIC model selection to select between m models for spins in each of the diffusion tensors optimized.

Chapter 3

Experimentation

3.1 Wet-lab practice

3.1.1 Brief overview of protein expression and purification as outlined in Johnson et al., 2018

The *pgmB* gene from *Lactococcus lactis* together with the *pgmB* gene containing a number of mutations were cloned in pET22b+ expression vectors and used to express βPGM proteins in *E. coli* strain BL21(DE3). One liter cell cultures were grown to log phase in either LB media or M9 media (with isotopic enrichments), induced with 1 mM IPTG and grown for a further 16 h at 25 °C. Cells were harvested by centrifugation at 10,000 rpm for 10 min at 4 °C, decanted and frozen at –80 °C. Cell pellets were resuspended in ice-cold standard native buffer (50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3) supplemented with one tablet of cOmplete^{TM} protease inhibitor cocktail (Roche). The cell suspension was lysed on ice by sonication for 6 cycles of pulsation for 20 s with 60 s cooling intervals. The cell lysate was then separated by ultracentrifugation (Beckman Coulter Avanti centrifuge) at 24,000 rpm for 35 min at 4 °C to remove insoluble matter. The cleared cell lysate was filtered using a 0.2 μm syringe filter and loaded onto a DEAE-Sepharose fast flow ion exchange column connected to an ÄKTA purification system that had been washed previously with 1 column volume of 6 M guanidine hydrochloride (GuHCl), 1 column volume of 1 M NaOH and equilibrated with $>$ 2 column volumes of standard native buffer. Following extensive washing, proteins bound to the DEAE-Sepharose column were eluted with a gradient of 0 to 100% standard native buffer containing 0.5 M NaCl. Fractions containing βPGM were checked for purity using SDS-PAGE, were pooled together and concentrated by Vivaspin (10 kDa MWCO). The protein sample was filtered using a 0.2 μm syringe filter and loaded onto a prepacked Hiload 26/60 Superdex 75 size-exclusion column connected to an ÄKTA purification system that had been pre-equilibrated with filtered and degassed standard native buffer containing 1 M NaCl. βPGM eluted as a single peak and fractions containing βPGM were checked for purity using SDS-PAGE, were pooled together, buffer exchanged into standard native buffer and concentrated to 1 mM by Vivaspin (10 kDa MWCO) for storage as 1 ml aliquots at –20 °C. The overall yield for βPGM was *ca.* 60 mg protein from 1 L of bacterial culture.

3.1.2 Reagent sourcing

The methods and reagents used thus far are outlined below. Any changes to either the protocol, recipe or reagent source are described in the relevant section of the text.

Reagent	Source
Mili-Q deionised water	Sartorius Arium 611VF Ultrapure Water System
Chemical reagents	Sigma-Aldrich
	Fisher Scientific
	Melford
Isotopically labelled chemical reagents	Goss Scientific Instruments Ltd.
	Cambridge Isotope Laboratories Inc.
	CortecNet
SDS-PAGE resin and markers	Bio-Rad
Primers for mutagenesis	Eurofins Scientific
Site directed mutagenesis kit	Agilent Technologies
Chromatography resins	GE Healthcare
Initial stocks of BL21(DE3), XL-blue, and XL-gold	Invitrogen or Novagen

3.1.3 Measuring pH

For large scale buffers (generally >100 ml) the pH of the solution was determined using a 'large probe' connected to a Russell RL150 pH meter at standard lab temperature (ca. 21 °C) unless otherwise stated. For small scale pH determination, (eg. for NMR samples) a smaller probe was used. The Sigma-Aldrich micro pH combination electrode was connected to a Russell RL150 pH meter at standard lab temperature (ca. 21 °C) unless otherwise stated. Both pH meters were calibrated to an R value ≥ 0.93 prior to usage using Scientific Laboroatory Supplies pH buffers at pH , 4.0, 7.0, and 10.0.

3.1.4 Common buffers used in protein preparation

The following buffers were made up in Mili-Q H_2O and the pH was adjusted to pH 7.2 using KOH or HCl prior to filter sterilisation.

TABLE 3.1: Standard purification buffer

Reagent	Concentration
K^+ HEPES (pH 7.2)	50 mM
$MgCl_2$	5 mM
NaN_3	2 mM

TABLE 3.2: Standard NMR buffer

Reagent	Concentration
K^+ HEPES (pH 7.2)	50 mM
$MgCl_2$	5 mM
NaN_3	2 mM
TSP	1 mM
D_2O	10%

3.1.5 Generation of Chemically Competent Cells ($CaCl_2$ method) for transformation

In order to isolate the single colony required to generate the chemically competent cell line, a sterile loop was scraped across an existing glycerol stock of the desired cell line and streaked across an LB Agar plate without antibiotics (Table 3.3). This plate was then incubated at 37 °C overnight. A single colony was picked and transferred to 5 ml of LB media in a 50 ml Falcon tube (Table 3.4) and this media was again incubated at 37 °C overnight, but at 200 rpm to optimise growth of cells. The following day, 1 ml of the starter culture was aseptically transferred to 9 ml of fresh LB media and incubated until the $OD_{600} \geq 0.4$. LB media was used as a reference absorbance. 1 ml aliquots of this culture were then transferred to eppendorfs and centrifuged at 13 krpm for 15 minutes in order to pellet the cells. The supernatant was discarded and the pellets were resuspended by pipette action in 1 ml of ice-cold 50 mM $CaCl_2$ solution and then left on ice for one hour, swirling every 20 min. The eppendorfs were again centrifuged to pellet the cells (13 krpm for 15 minutes). The supernatant was again discarded and the cells gently resuspended in 200 μl of ice-cold $CaCl_2$ solution containing 50% (w/v) Glycerol. These cells were then placed in the -80°C freezer until needed.

3.1.6 Transformation

The gene for the Lactococcus lactis βPGM was cloned into the pET-22b(+) vector (Figure 3.1), using the Nde1 and Xho1 restrictions sites by Dr. Nicola Baxter as described in Baxter et al., 2006. This construct does not have a periplasmic leader sequence, nor any N- or C-terminal tags. This construct conveys resistance to β-lactam containing antibiotics such as ampicillin and as such, ampicillin (100 μg/ml) is routinely used as a selection marker for transformed colonies.

1 μl of plasmid stock was routinely added (aseptically) to either BL21(DE3) or XL Blue chemically competent cells ($CaCl_2$ method) for either plasmid amplification of protein overexpression purposes respectively. Transformed cells were incubated on ice for 15-20 minutes, before a heat shock at 42 °C for 45 seconds, and were then returned to ice for two minutes. 1 ml of LB (without antibiotic, Table 3.4) was added to the culture and incubated at 37 °C for 1 hour without shaking. After the incubation period, typically volumes of 50-200 μl were

plated out on selective LB/ampicillin agar plates (Table 3.3) and incubated overnight at 37 °C to isolate successfully transformed bacteria.

TABLE 3.3: LB-agar

Reagent	Concentration (g/l)
Tryptone	10
Yeast extract	5
NaCl	10
Bacto-Agar	15

The pH of the solution was adjusted to approximately 7.0 prior to sterilisation by autoclave.

TABLE 3.4: LB-media

Reagent	Concentration (g/l)
Tryptone	10
Yeast extract	5
NaCl	10

The pH of the solution was adjusted to approximately 7.0 prior to sterilisation by autoclave.

3.1.7 Site Directed Mutagenesis

Site directed mutagenesis (SDM) was performed using the QuikChange II SDM kit (Agilent technologies) and the polymerase chain reaction (PCR) was carried out using a Progene Thermal Cycler (Techne). Primers (Eurofins) were designed using an online tool (PrimerX) to introduce point mutations in the βPGM gene and generate βPGM variants for structural and mechanistic investigation. These primers were then added to a reaction mixture containing as outlined in table 3.5 below. 22 cycles of annealing, extension and melting were carried out as illustrated in table 3.6 before 1 μl of Dpe1 was added as per the instructions to degrade the methylated parental DNA, selecting only for the PCR product. 2-4 μl was then transformed into chemically competent XL1-Blue cells. Latterly, the mutagenesis service provided by Genscript (Hong Kong) became a viable alternative for large mutagenesis studies.

TABLE 3.5: PCR reaction mixture

Reagent	Quantity
Forward primer	125 ng
Reverse primer	125 ng
Wild type plasmid	50 ng
10x reaction buffer	5 μl
dNTP mix	1 μl
PfuTurbo	2.5 U
ddH2O	add to 50 μl

The contents of the PCR reaction mixture, total volume = 50 μl .

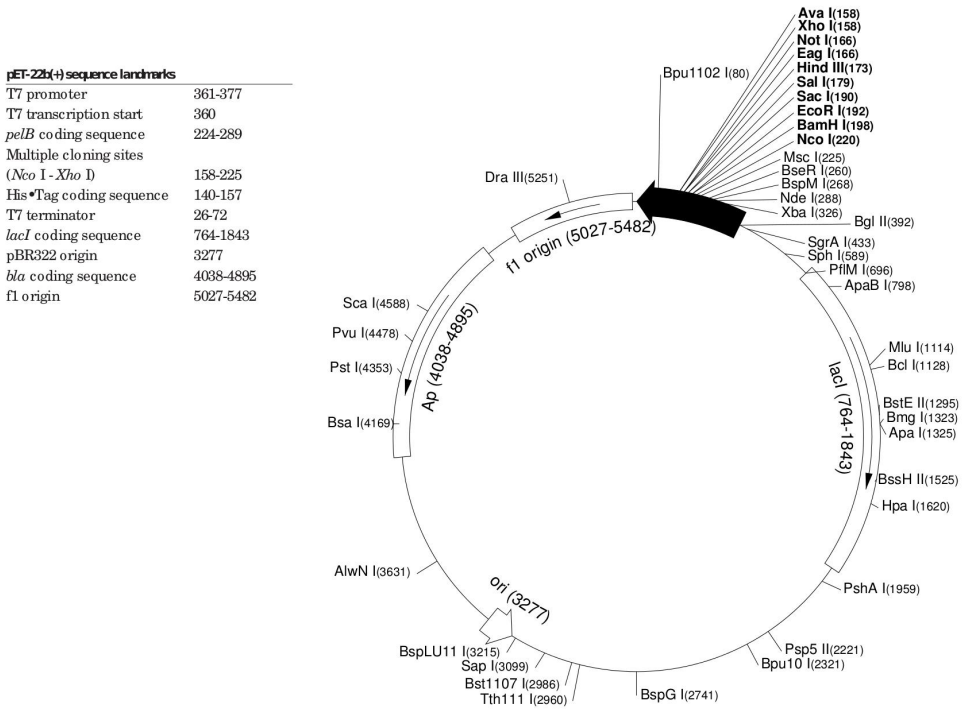


FIGURE 3.1: pET-22b(+) Vector Map (Novagen), illustrating the location of restriction sites, coding regions, and origins of replication and translation (adapted from: //www.helmholtz-muenchen.de/fileadmin/PEPF/pET_vectors/pET-22b_ map.pdf).

TABLE 3.6: PCR cycling parameters

Segment	Cycles	Temperature	Time
1	1	95°C	30 seconds
2	16-18	95°C	30 seconds
		55°C	1 minute
		68°C	6 min 30

The cycling parameters for the PCR reaction.

3.1.8 DNA sequencing and amplification

Plasmid stocks were were amplified prior to sequencing using the Miniprep protocol provided with the QIAprep Spin Miniprep Kit (Qiagen), but in the final step, the plasmid was eluted with water instead of the provided elution buffer. Ca. 30 μl of the resultant plasmid stock ($\geq$ 40 ng/μl) was then sent to GATC Biotech for sequencing. The resulting sequence data was visualised with FinchTV. Plasmid stocks were amplified in chemically competent *Escherichia coli* XL1-Blue using the Miniprep protocol provided with the QIAprep Spin Miniprep Kit (Qiagen).

3.1.9 Cell culture and labelling strategies

Following the successful transformation of the plasmid containing the βPGM gene into chemically competent BL21(DE3) *E. coli* cells, a single transformant colony was selected and inoculated aseptically into a starter culture. Depending on the labelling strategy, either 10 ml of LB (Table 3.4) or M9 (Table 3.7, 3.8) media was used containing 100 μg/ml Ampicillin in a 50 ml Falcon tube. This starter culture was then incubated at 37°C and 200 rpm overnight. This starter culture was then added to a full scale culture in many cases (see labelling strategies Table 3.10) supplemented with 10 /ml Ampicillin and then incubated at 37°C and 180 rpm until the OD_{600} reached 0.6. 1 mM IPTG was added to induce expression of βPGM and the flasks were then shaken at 200 rpm and 25°C for ca. 16 hours. The cells were then centrifuged in a Beckman Avanti J-25I centrifuge fitted with a JSP F500 rotor at 10 krpm for 10-15 minutes at 4°C. Cells were resuspended into a smaller volume of their growth buffer by pipette action and transferred into a 50 ml Falcon tube. The cells were then centrifuged at 9 krpm at room temperature (set to ca. 21 °C) for 20 mins using a Sigma 3–15 Centrifuge giving a single pellet. The supernatant was discarded, the pellet dried and then frozen at -80 °C.

TABLE 3.7: M9 minimal medium Step - 1

Reagent	Concentration (g/l)
Na_2HPO_4	6
KH_2PO_4	3
NaCl	0.5

After dilution up to the desired volume with Milli-Q water (99.7% D_2O), the pH of the solution was adjusted to 7.4 (7.4) prior to sterilisation by autoclave (filter sterilization). Following autoclaving the components outlined in Table 3.8 were added aseptically.

TABLE 3.8: M9 minimal medium Step - 2

Reagent	Amount added per litre
1 M $MgSO_4$ (autoclaved)	1000 μl
1 M $CaCl_2$ (autoclaved)	100 μl
Trace elements (autoclaved see Table 3.10)	650 μl
10 mg/ml Thiamine (filter-sterilised)	100 μl
^{15}N Nitrogen source ($(NH_4)_2SO_4$ or NH_4Cl)	1g (typically in $\leq$2 ml)
Carbon source (see Table 3.10)	10 - 15 ml (20% w/v glucose)

These components were all added aseptically to the M9 buffer from Table 3.7. For labelling schemes see table 3.10. Note that the $CaCl_2$ was always added last. When the growth was in *ca.* 100% D_2O, stock solutions of reagents were made up in D_2O to avoid the introduction of H_2O into the media.

3.1.10 Sonication

The frozen cell pellet was thawed on ice and resuspended in 25 ml of standard βPGM buffer (Table 3.1) by pipette action. One pellet of EDTA-free protease inhibitor cocktail (Protease cOmplete, Roche) was included to inhibit protease activity during the sonication procedure. The suspended cell culture in a 50 ml falcon tube was placed in a 200 ml beaker containing ice water and the suspension was then sonicated using a Soniprep 150 (MSE) for 5 x 20-second bursts at full power separated by 60 second intervals. The lysate was then transferred to a 50 ml centrifuge tube and transferred to a Beckman Avanti J-25I centrifuge fitted with a JA-25,50 rotor and spun at 24 krpm for 35 mins at 4°C. The supernatant containing soluble βPGM was aseptically transferred to a 50 ml Falcon tube and the pellet containing the insoluble lysate fraction discarded. This Falcon tube was kept on ice until it was transferred to the anion exchange column in the next step of the purification protocol.

3.1.11 Purification

The anion exchange column is a positively charged column at pH 7.2 that will bind and sufficiently retard negatively charged proteins and provide a selection method for their isolation. The theoretical pI of βPGM is 4.84, consequently at a neutral pH 7.2, βPGM will

TABLE 3.9: Trace elements

Reagent	mass added to 80 ml Milli-Q water (mg)
$CaCl_2.2H_2O$	550
$ZnSO_4.7H_2O$	220
$MnSO_4.H_2O$	140
$CoCl_2.6H_2O$	45
$CuSO_4.5H_2O$	40
H_3BO_3	40
$Na_2MoO_4.2H_2O$	26
KI	26
– pH adjusted to 8.0 using acetic acid	—
EDTA	500
– pH adjusted to 8.0 using acetic acid	—
$FeSO_4.7H_2O$	
– Solution made up to 100 ml using Milli-Q water and autoclaved	—

These components were all added in the order outlined above and the pH of the solution was corrected at the points indicated.

TABLE 3.10: Labelling schemes

Scheme	Carbon source (CS)	Vol (ml)	D_20
^{15}N	20% (w/v) unlabelled Glucose	15	0%
^{15}N and ^{13}C	20% (w/v) ^{13}C- Glucose	10	0%
^{15}N, ^{1}H and ^{13}C	20% (w/v) ^{13}C- ^{2}H- Glucose	10	100%

These components were all added aseptically.

have a negative charge while many of the other contaminating proteins in the supernatant post-centrifugation will not.

Cell lysate following sonication was added to a ca. 40 ml DEAE-sepharose column at room temperature that had previously been equilibrated with $\geq$ two volumes of standard purification buffer (Table 3.1). The column was then washed with 3-5 column volumes of standard buffer, or until the conductivity of the eluate dropped to zero (measured by the UV detector on the connected ÄKTAprime Plus). Bound proteins were eluted with a gradient of 0 to 100% standard βPGM buffer containing 0.5 M NaCl. The ÄKTAprime Plus collected 5 ml fractions and βPGM typically eluted at approx 0.2M NaCl across 50 ml and the purity was determined using SDS-PAGE (Section 3.1.12). Fractions that contained sufficiently high βPGM content ($\geq$ 95%) were pooled and concentrated using a 10 kDa MWCO Vivaspin to a volume of ca. 5 ml. The DEAE-sepharose column was then washed with 2 column volumes of 1 M NaOH, 6M GuHCl, and 20 % w/v Ethanol solutions.

A Superdex G75 column was used to separate the remaining proteins based on hydrodynamic radius. The column was equilibrated with approx. 2 column volumes of standard buffer with the addition of 1M NaCl which was used to prevent protein binding to the gel as

the resolution of size exclusion is reliant on steady flow of solution through the column. The concentrated impure βPGM sample from the anion exchange step was added to the column with care being taken not to introduce air bubbles and the column was run at 1.5-2 ml/min. 5 ml fractions were collected by the ÄKTAprime Plus and the putative βPGM containing fractions were assessed for purity using SDS-PAGE as outlined in section 3.1.12. The fractions containing pure βPGM were pooled and buffer exchanged ($\geq$ 800x dilution) back into standard buffer without salt using a 10 kDa MWCO Vivaspin at 4.5 krpm. Final enzyme concentrations for storage at -20 °C were 1-1.5 mM, and the Superdex G75 column was washed with two column volumes of 1M NaOH and 20% Ethanol solutions prior to storage.

3.1.12 SDS-PAGE

Polyacrylamide Gel Electrophoresis (PAGE) in the presence of the denaturant, Sodium Dodecyl Sulfate (SDS) was used routinely to assess expression levels, fraction protein content and fraction purity during the βPGM expression and purification procedures. Bio-Rad Mini-Protean II equipment was used to run the gels. The standard gel composition was ca. 5 cm resolving gel (Table 3.11, 3.13) to ca. 2 cm stacking gel (Table 3.12, 3.14). After setting, the gel was removed from the casting stand, any extraneous polyacrylamide removed, and wrapped in Mili-Q H_2O saturated lab roll which was in turn wrapped in aluminium foil. This was often refrigerated overnight to ensure complete crosslinking of the polyacrylamide prior to use.

Samples prior to SDS-PAGE were prepared by mixing 15 μl of sample with 5 μl of 4x SDS-PAGE loading buffer (Table 3.16). This mixture was then boiled at 95 °C for 5 minutes to ensure protein denaturation. A prepared gel was removed from the fridge and washed with Mili-Q H_2O prior to insertion into the electrode assembly and immersion in SDS-PAGE running buffer (see 3.15). Only 10 μl of sample was typically loaded into any well to increase to minimise gel distortion, while the two outermost lanes were never used and the first lane was reserved for 4 μl pre-stained marker (Bio-Rad).

The gels were run at 50 V for 10 min to load protein onto the stacking gel and then the gel was typically run at 180 V for 45-60 min, or until the layer of Bromophenol Blue from the loading buffer reached the end of the gel. The gel was then removed from the apparatus and the glass plates prior to immersion in CoomasieBlue instant stain. Protein bands were observable from 10 minutes after immersion in instant stain.

TABLE 3.11: 4X SDS-PAGE Resolving Gel Buffer

Reagent	Concentration
Tris	1.5 M
SDS	0.4 % (w/v)

The pH of the stacking gel buffer was adjusted to 8.8 with HCl prior to filter sterilisation.

TABLE 3.12: 4X SDS-PAGE Stacking Gel Buffer

Reagent	Concentration
Tris	0.5 M
SDS	0.4 % (w/v)

The pH of the stacking gel buffer was adjusted to 6.8 with HCl prior to filter sterilisation.

TABLE 3.13: SDS-PAGE Resolving Gel (16% BisAcrylamide)

Reagent	Volume added
4X SDS-PAGE Stacking Gel Buffer	2.5 ml
40% (w/v) Acrylamide/BisAcrylamide (37.5:1)	4 ml % (w/v)
Milli-Q water	3.5 ml
10% (w/v) Ammonium Persulphate (APS)	100 μl
Tetramethylethyldiamine (TEMED)	10 μl

The components of the gel were added in the order outlined above and the resultant solution was swirled to mix (avoiding excessive aeration) prior to pouring.

TABLE 3.14: SDS-PAGE Stacking Gel (4.5% BisAcrylamide)

Reagent	Volume added
4X SDS-PAGE Stacking Gel Buffer	2.5 ml
40% (w/v) BisAcrylamide (37.5:1)	1.125 ml % (w/v)
Milli-Q water	6.375 ml
10% (w/v) Ammonium Persulphate (APS)	110 μl
Tetramethylethyldiamine (TEMED)	11 μl

The components of the gel were added in the order outlined above and the resultant solution was swirled to mix (avoiding excessive aeration) prior to pouring.

3.1.13 Protein concentration determination

A Varian Cary 50 Bio UV/Vis spectrophotometer was used to determine the Optical Density at 280 nm (OD_{280}) for protein samples (βPGMϵ = 19940 L mol^{-1} cm^{-1}). This gave the concentration of protein when entered into a rearranged Beer-Lambert law as illustrated in Eq. 3.1,

$$A = \epsilon c l \tag{3.1}$$

where A is absorbance, ϵ is the molar attenuation coefficient, c is the molar concentration, and l is the path length of the cuvette in cm. *Ca.* 50 fold dilutions of final protein concentrations were made for samples following purification and concentration as the optimal

TABLE 3.15: SDS-PAGE Running Buffer

Reagent	Concentration
Tris	25 mM
SDS	0.1 % (w/v)
Glycine	250 mM

The pH of the stacking gel buffer was adjusted to 8.3 using HCl.

TABLE 3.16: 4x SDS-PAGE Loading Buffer

Reagent	Concentration
Tris	200 mM
Dithiothreitol (DTT)	400 mM
SDS	8% (w/v)
Bromophenol Blue	0.4% (w/v)
Glycerol	40% (w/v)

The pH of the SDS-PAGE loading buffer was adjusted to 6.8 with HCl.

sensitivity of this technique is A=0.6.

3.2 X-ray Crystallography

3.2.1 Overview of Crystallization techniques

Crystallization of native βPGM$_{WT}$ was achieved using the same conditions as described previously (Baxter et al., 2010). The βPGM$_{WT}$ protein solution was routinely mixed 1:1 with precipitants (26-30% (w/v) PEG 4000, 200 mM sodium acetate and 100 mM Tris-HCl (pH 7.5)) and crystals were grown at 290 K by hanging-drop vapor diffusion using a 2 μL drop suspended on a siliconized glass cover slip above a 700 μL well.

Both rod and plate shaped crystals formed after several days which were cryo-protected in their original mother liquor containing an additional 25% (v/v) ethylene glycol prior to plunging into liquid nitrogen. Both crystal morphologies often diffracted in the P $2_12_12_1$ spacegroup, or P 2_1 spacegroup with reduced symmetry. βPGM$_{WT}$ and several enzyme variants with single residue mutations all crystallized under the same conditions, producing either one monomer in the asymmetric unit (P $2_12_12_1$) or two (P 2_1).

For soaking experiments, native βPGM crystals were cryo-protected in their original mother liquor containing an additional 25% (v/v) ethylene glycol together with desired soaking molecules (eg. AcP or Pi), and were incubated for a range of timescales (30 - 180 s) prior to plunging into liquid nitrogen.

3.2.2 Data collection, processing, and refinement

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; Kabsch, 2010) with resolution cut-offs applied using CC-half values and the structures were determined by molecular replacement with MolRep (Vagin and Teplyakov, 1997) using previously modelled βPGM PDB structures 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 1.5Å) or anisotropic temperature factors (resolutions better than 1.5Å) was performed using REFMAC5 (Murshudov, Vagin, and Dodson, 1997) in the CCP4i suite (Winn et al., 2011). Ligands and protein modifications 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 1.8/2.0 Schrödinger, LLC), maps were generated using FFT (Read and Schierbeek, 1988) and domain movements were calculated using DynDom (Hayward and Berendsen, 1998).

3.2.3 Modelling partial occupancy ligands into electron density in the active site of βPGM: The case of βG16BP in PDB: 5OK0

Rod shaped crystals harvested after 1 week contained predominantly βG16BP in the βPGM$_{D10N}$ active site, with the 6-phosphate group located in the proximal site and the 1-phosphate group bound in the distal site (βPGM$_{D10N}$:P6G1P complex). After refinement, the ratio of 2Fo–Fc density present between the 1- and 6-phosphate groups (ca. 6σ and 5σ, respectively)

and did not correlate with a full βG16BP ligand occupancy in the βPGM$_{D10N}$:P6G1P complex. When modeled at a ligand occupancy of 0.8, B-factor convergence was attained between the βG16BP ligand and neighboring residues in the active site, confirming βG16BP as the dominant ligand. Remaining difference map peaks were consistent with the presence of a minor population of βG1P (with the 1-phosphate in the distal site) but, due to poor connectivity at this resolution, βG1P was not modeled into the structure. Crystals from the same drop with the same morphology harvested after 12 weeks contained only βG16BP bound in the alternate orientation with the 1-phosphate group located in the proximal site and the 6-phosphate group bound in the distal site (βPGM$_{D10N}$:P1G6P complex). Section adapted from Paper I.

3.3 NMR spectroscopy

Almost all experiments were acquired at 298 K using 0.5 – 1 mM βPGM with desired labelling strategies in standard NMR buffer (50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3 with 10% (v/v) D_2O and 1 mM trimethylsilylpropanoic acid (TSP), Table 3.2). Due to the availability of magnets at both Sheffield (Department of Molecular Biology and Biotechnology, Sheffield University (UoS)) and Manchester (Manchester Institute of Biotechnology; (MIB)) (and periodically access to the 950 MHz at the Crick institute, London), we were fortunate enough to have access to magnets at multiple fields fitted with cryoprobes, which proved to be invaluable for the investigation of backbone dynamics.

3.3.1 ^{1}H NMR

^{1}H 1D spectra were routinely acquired to assess sample condition, folding, pH and referencing. The pH was determined by correlating the frequency separation of two methylene HEPES peaks against a standard calibration curve determined by Dr. Clare Trevitt (in house). 1 mM TSP was almost always included in samples to aid in the referencing of both 1D and multidimensional NMR spectra. The _ *1dpecw1* pulseprogram was routinely used in Sheffield for acquisition of ^{1}H 1D NMR spectra (Fig. 3.2). This pulse program includes a water presaturation pulse during D1, as well as a Hahn-echo refocusing scheme prior to acquisition. Typically 64-256 scans were acquired with a spectral width of 25 ppm, centered on the water signal at 4.7 ppm. The inter scan delay (D1) was varied from 1s - 2.5s depending on the level of sample deuteration.

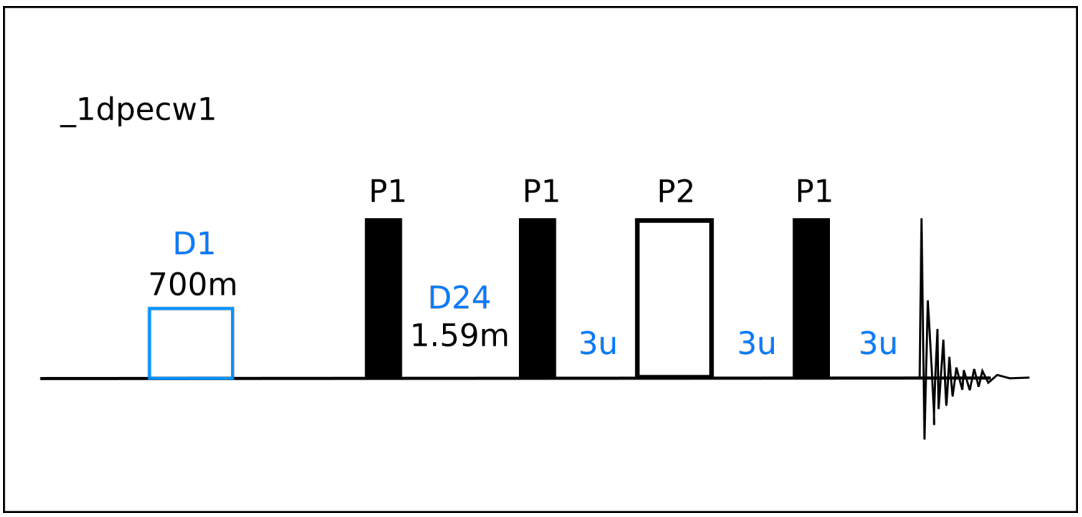


FIGURE 3.2: The pulseprogram _ 1dpecw was routinely used in Sheffield to record ^{1}H 1D NMR spectra. This pulse scheme features a water saturation pulse during the inter scan delay, followed by a Hahn-echo refocusing scheme. Typically the first 16 (600MHz) or 32 (800MHz) points were removed prior to processing of the FID.

The *zgesgp* pulseprogram was routinely used at the MIB for acquisition of ^{1}H 1D NMR spectra (Fig. 3.3). This pulse program uses water selective 180°pulses to eliminate signal from solvent. Typically 64-256 scans were acquired with a spectral width of 16 ppm, centered on

the water signal at 4.7 ppm. The inter scan delay (D1) was varied from 1s - 2.5s depending on the level of sample deuteration.

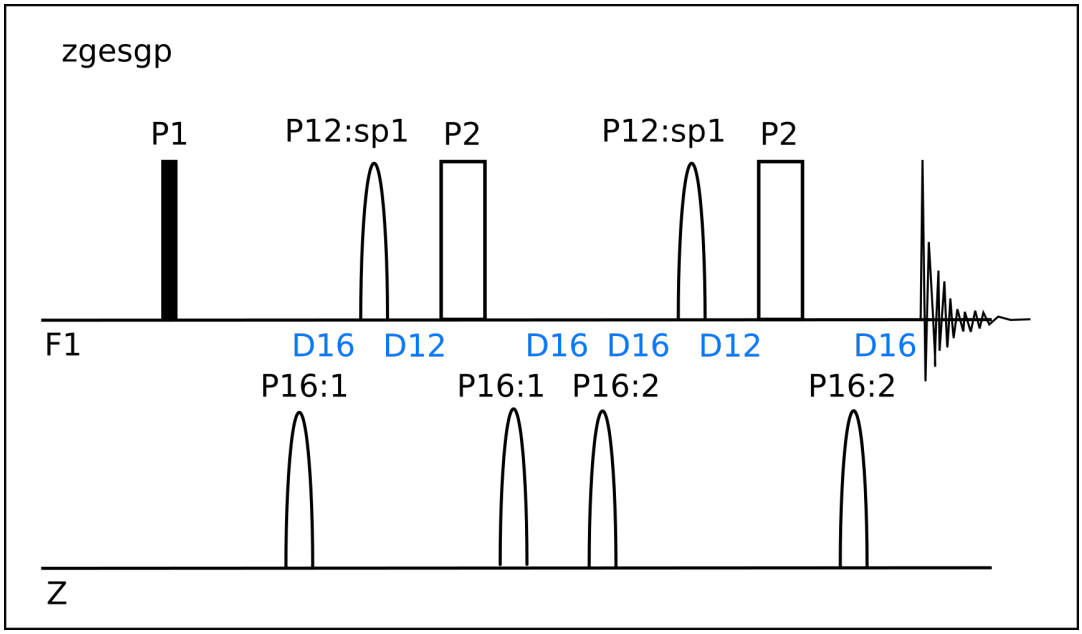


FIGURE 3.3: ^{1}H 1D pulseprogram zgesgp used at the MIB. P1 (90°) and P2 (180°) are hard ^{1}H pulses applied across the full spectral width, while P12:sp1 is a shaped pulse to selectively excite water signal. Gradients were applied along the z-axis as denoted by hollow semi-ellipses.

^{19}F NMR

The *zg* pulseprogram was routinely used at the MIB for acquisition of ^{19}F 1D NMR spectra (Fig. 3.4). This is a simple pulse-acquire scheme and due to the QCIF probe on the magnet in Manchester, extraordinary sensitivity was achievable compared to RT probes. Typically 64-1024 scans were acquired with a spectral width of 120 ppm, centred at -140 ppm. The inter scan delay (D1) was typically 1s - 1.5s.

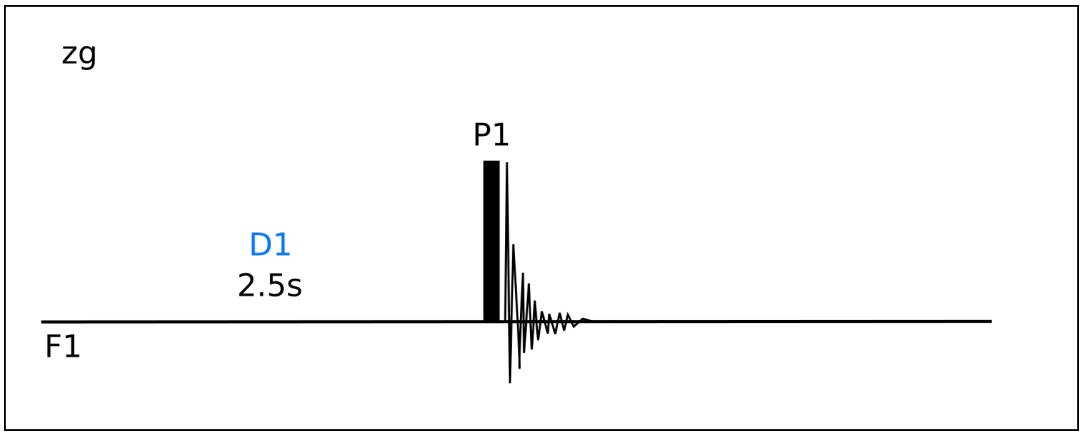


FIGURE 3.4: zg pulseprogram used to acquire ^{19}F 1D spectra at the MIB. All narrow rectangular pulses correspond to 90 °pulses about the x-axis unless otherwise indicated.

^{31}P NMR

Either *zg* or *zgig* pulseprogram used on the broadband probe at the UoS to acquire ^{31}P 1D NMR spectra. The *zgig* pulse scheme features decoupling of a second channel during acquisition, in this case that second nucleus was ^{1}H(Fig. 3.5). Typically 64-1024 scans were acquired with a spectral width of 40 ppm, centered at -10 ppm. The inter scan delay (D1) was typically 1s - 1.5s for semi quantitative spectra, or up to 10s for more quantitative spectra amenable to time dependent changes in concentration.

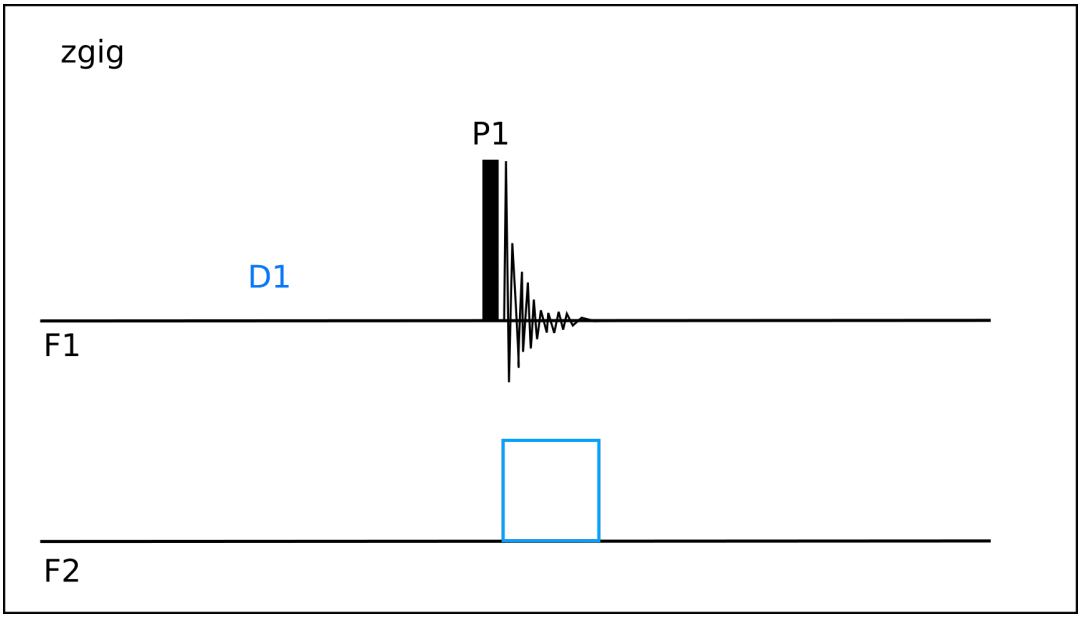


FIGURE 3.5: zgig pulseprogram used to acquire ^{31}P 1D spectra at the UoS with proton decoupling during acquisition.All narrow rectangular pulses correspond to 90°pulses about the x-axis unless otherwise indicated, the blue rectangular pulse denotes a decoupling scheme

3.3.2 2D NMR

^{1}H–^{15}N TROSY NMR

The *trosyetf3gpsi.2* pulseprogram was typically used for acquisition of ^{1}H^{15}N-TROSY (Transverse Relaxation Optimised SpectroscopY) 2D NMR spectra (Fig. 3.6) (Schulte-Herbrüggen and Sørensen, 2000; Czisch and Boelens, 1998; Pervushin, Wider, and Wüthrich, 1998; Meissner et al., 1998; Weigelt, 1998; Rance, Loria, and Palmer, 1999; Zhu, Kong, and Sze, 1999). This pulse scheme uses phase sensitive Echo/Antiecho gradient selection to isolate the ^{1}H^{15}N correlation peak with the narrowest linewidth in both ^{1}H and ^{15}N dimensions. Typically 8-32 scans of 256 complex points were acquired with a spectral width of 16 ppm and 36 ppm in proton and nitrogen dimensions respectively. Spectra were centred on the water signal at 4.7 ppm and the inter scan delay was varied from 1s - 2.5s depending on the level of sample deuteration.

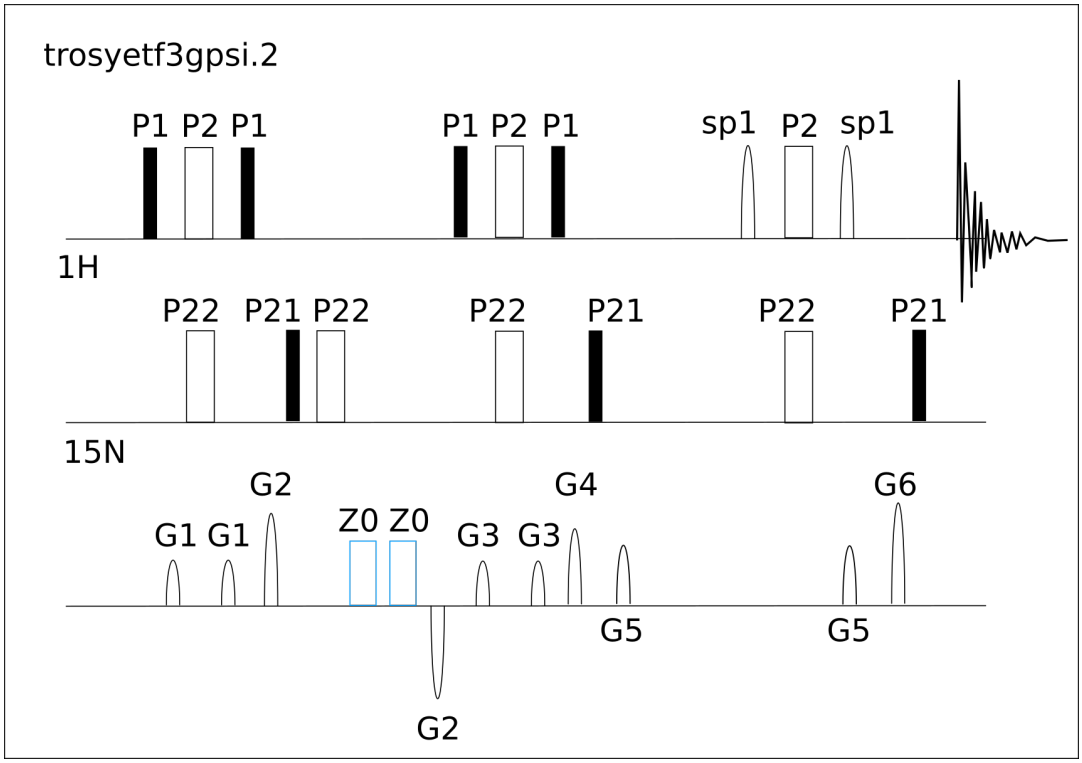


FIGURE 3.6: ^{1}H-^{15}N 2D pulseprogram *trosyetf3gpsi.2* used at the UoS. All narrow (wide hollow) rectangular pulses correspond to 90°(180° pulses about the x-axis unless otherwise indicated. Non-rectangular pulses indicate shaped pulses while blue rectangular pulses denote decoupling schemes. Gradients were applied (often in pairs) along the z-axis as denoted by the hollow semi-ellipses.

^{1}H–^{13}C HSQC NMR for stereoassignment

The _ *chsqcali3i* pulseprogram was used for acquisition of ^{1}H^{13}C-HSQC (Heteronuclear Single Quantum Coherence) 2D NMR spectra (Fig. 3.6). This pulse scheme uses a phase sensitive States-TPPI to isolate a ^{1}H-^{13}C correlation peak via a double inept method. This is a

constant time version of the pulse scheme with carbon decoupling during acquisition. Typically 8 scans of 736 complex points were acquired with a spectral width of 12.5 ppm and 69 ppm in proton and carbon dimensions respectively. Spectra were centred on the water signal at 4.7 ppm and the inter scan delay was typically 1s.

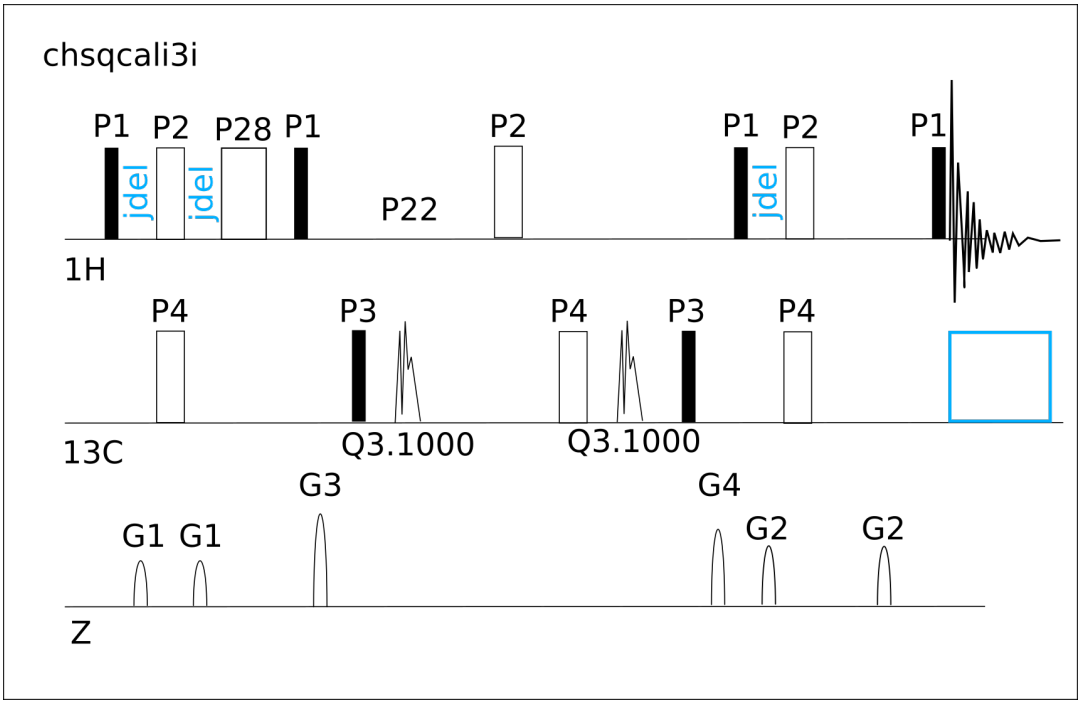


FIGURE 3.7: ^{1}H-^{13}C 2D pulseprogram chsqcali3i used in Sheffield. All narrow (wide hollow) rectangular pulses correspond to 90°(180° pulses about the x-axis unless otherwise indicated. Non-rectangular pulses indicate shaped pulses while blue rectangular pulses denote decoupling schemes. Gradients were applied (often in pairs) along the z-axis as denoted by the hollow semi-ellipses.

3.3.3 3D NMR

Acquisition of backbone assignment spectra

Backbone assignment spectra for triple labelled (^{2}H^{13}C^{15}N, back exchanged into H_2O) protein samples were either acquired without non-uniform sampling (NUS) in Sheffield or with NUS in Manchester on their respective 800 MHz magnets. Backbone assignment spectra without NUS were processed in Felix (Felix NMR, Inc) and assigned using the asstools suite of programs (Reed et al., 2003). Backbone assignment spectra with NUS were acquired using a multi dimensional Poisson Gap scheduling strategy with exponential weighting (Hyberts, Robson, and Wagner, 2013). NUS data were either reconstructed using TopSpin3 and multidimensional decomposition (Hyberts et al., 2012), or a command line based IST reconstruction (Hyberts et al., 2012) and processing in NMRPipe (Delaglio et al., 1995). The standard suite of 6 experiments were routinely recorded for protein backbone assignment, namely; HNCO, HN{CA}CO, HNCA, HN{CO}CA, HNCACB, HN{CO}CACB experiments. All programs were the TROSY based variants from the standard Bruker library, with minimal alteration.

3.3.4 pseudo-3D NMR

NMR Relaxation Measurements for 15N fast timescale relaxation were performed using $^{2}H^{15}N$–labelled βPGM enzyme in 5mm Shigemi D_2O matched tubes. Samples were recorded in standard NMR buffer with the addition of metal fluorides and ligands where necessary. Experiments were acquired using: (deep breath) a Bruker 600 MHz Avance DRX spectrometer equipped with a 5-mm TXI cryoprobe and z-axis gradients (Sheffield), a Bruker 600 MHz Avance III spectrometer equipped with a 5-mm TCI cryoprobe and z-axis gradients (Manchester), a Bruker 800 MHz Avance spectrometer equipped with a 5-mm TXI probe and z-axis gradients (Sheffield), a Bruker 800 MHz Avance III spectrometer equipped with a 5-mm TCI probe and z-axis gradients (Manchester) and a 950MHz Avance III spectrometer equipped with a TCI probe and z-axis gradients (Mill Hill).

^{1}H–^{15}N R_1, $R_{1\rho}$, HetNOE relaxation rates determination

^{15}N relaxation rates were determined in perdeuterated, amide-protonated proteins using a TROSY ^{1}H detection scheme (Fig 3.8). Fig 3.8**a** illustrates the scheme for determination of ^{15}N R_1 parameters. Substitution of the red bracketed block with that in 3.8**b** converts the experiment to a $R_{1\rho}$ experiment. Fig 3.8**c** illustrates the ^{15}N-{^{1}H} NOE pulseprogram. The shaped low power ^{1}H pulses correspond to the center lobe of a (sinx)/x function in order to return the water magnetization to z prior to detection. Gradients: G0, G1, G2, G3, and G9 are rectangular shaped, whereas G5, G4, G6, G7, and G8 are sine-bell shaped.

Quadrature detection is implemented using the Rance-Kay echo/anti-echo scheme (Kay, Keifer, and Saarinen, 1992). ^{13}C 180°pulses eliminate ^{15}N-^{13}C cross correlations in samples that include ^{13}C labelling. Temperature compensation is achieved by the inclusion of a ^{15}N temperature compensation pulse (Wang1993) that corresponds to the longest spin-lock time and RF power of the $R_{1\rho}$ experiment immediately following data acquisition. The 90°^{15}N pulse (red) preceding this temperature compensation pulse eliminates ^{15}N-z magnetization, transferred from ^{1}H to ^{15}N by the TROSY readout scheme (Favier and Brutscher, 2011). In **b**, the triangle shaped pulses immediately preceding and following the spin-lock period are adiabatic half passage (AHP) pulses (Mulder et al., 1998) of a tangent hyperbolic tangent (tanh/tan) adiabatic inversion pulse. In **c**, ^{1}H saturation is achieved by n repetitions of the symmetric (Δ–180°Δ) unit (Ferrage et al., 2010), with the ^{1}H carrier switched to 8.6 ppm, after saturation, the ^{1}H carrier is switched back to the H_2O resonance.

^{1}H Relaxation dispersion

^{1}H relaxation dispersion rates were determined in perdeuterated, amide-protonated proteins using a HSQC ^{1}H detection scheme (Fig 3.9) written by Kenji Sugase (Sugase et al., 2007), but with the inclusion of a square pulse during the CPMG period rather than a selective Reburp.1000. This pulse scheme observes a ^{1}H-^{13}C correlation peak via a double inept method separated by two separate CPMG trains (Carr-Purcell-Meiboom-Gill) about the y and x axes (Carr and Purcell, 1954; Meiboom and Gill, 1958). A watergate sequence is included at the end of the sequence with water flip-back pulses (P29) to suppress solvent signal. A constant time of 80 ms is used for the CPMG blocks, with the number of 180°pulses

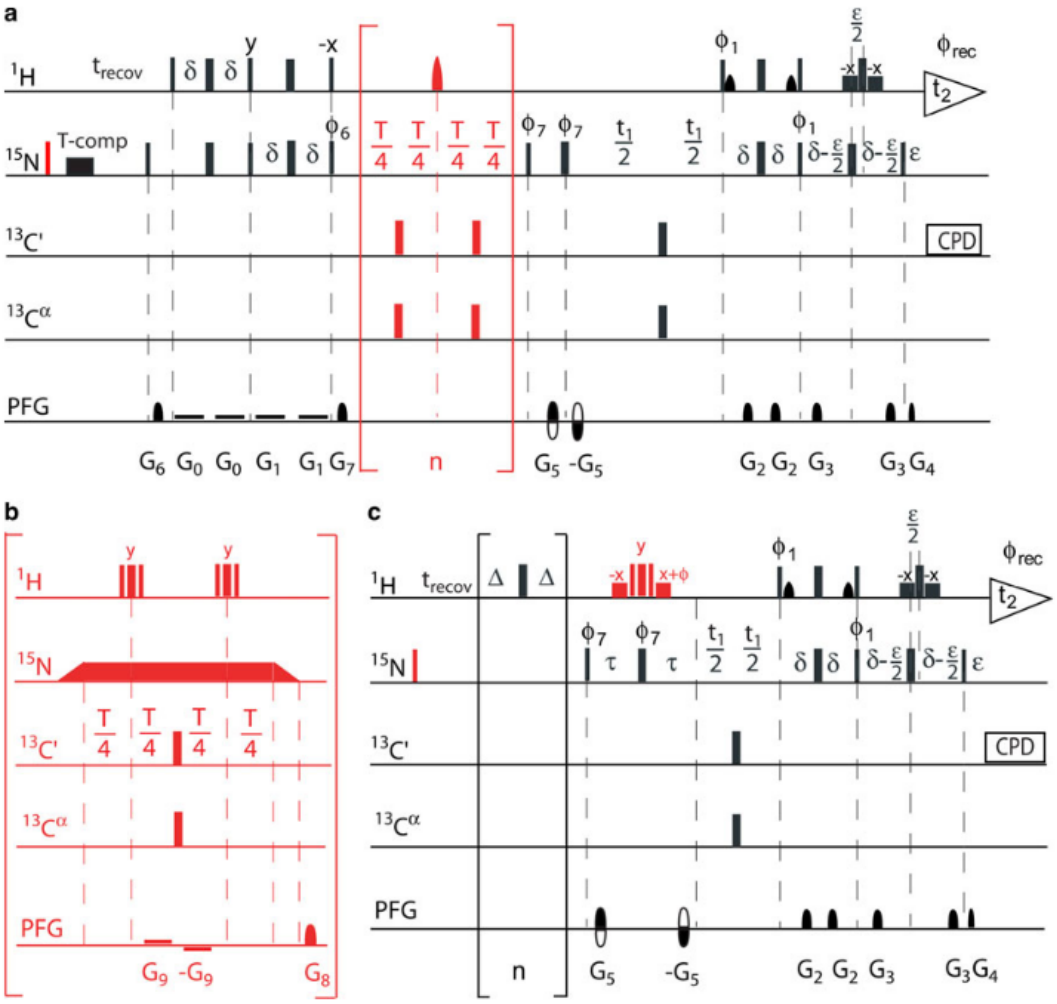


FIGURE 3.8: $^{1}H^{15}N$ Backbone relaxation pulseprograms used in Sheffield, adapted from (Lakomek, Ying, and Bax, 2012). **(a)** A R_1 experiment that following the replacement of the red bracketed region becomes a $R_{1\rho}$ experiment **(b)**. The HetNOE experiment is depicted in **(c)**. All narrow (wide) rectangular pulses correspond to 90°(180° pulses about the x-axis unless otherwise indicated. The rectangular low amplitude ^{1}H pulses correspond to low power square pulses, whereas shaped low-power pulses correspond to the center of a (sinx)/x function. In **b** the triangle shaped pulses immediately preceding and following the spin-lock period are adiabatic half-passage of a tangent hyperbolic tangent (tanh/tan) adiabatic inversion pulses.

given by the loop L4 within that period. Typically 16 scans of 256 complex points were acquired with a spectral width of 15 ppm and 34 ppm in proton and nitrogen dimensions respectively. Spectra were centred on the water signal at 4.7 ppm and the inter scan delay was typically 3.5s.

^{2}H–^{13}C R_1 and $R_{1\rho}$

Measurement of both $T1\rho(I_zC_zD_y)$ or $T1(I_zC_zD_z)$ values in $^{13}CH_2D$ spin systems was performed on Varian spectrometers in Lund using pulse schemes adapted from (Muhandiram et al., 1995) (3.10). These pulse schemes use a double inept HSQC method to isolate a ^{1}H-^{13}C correlation peak of residues [A,I,L,M,T,V] with quadrature detection achieved using a phase

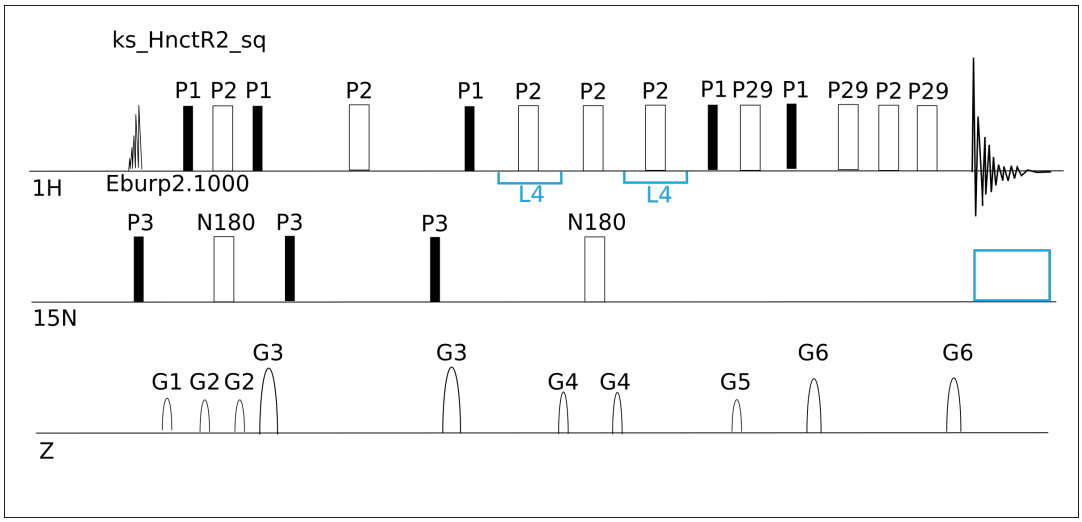


FIGURE 3.9: Backbone amide relaxation dispersion experiment ks_ HNctR2_ sq for ^{1}Hrelaxation used in Sheffield. All narrow (wide hollow) rectangular pulses correspond to 90°(180° pulses about the x-axis unless otherwise indicated. Non-rectangular pulses indicate shaped pulses while blue rectangular pulses denote decoupling schemes. Gradients were applied (often in pairs) along the z-axis as denoted by the hollow semi-ellipses.

sensitive States-TPPI. This ^{1}H-^{13}C correlation of a sidechain methyl group is encoded with either $T1$ or $T1\rho$ relaxation properties of the associated ^{2}H atom depending on the pulse scheme chosen. Carrier frequencies were typically centred at 4.7, 20.0, and 0.8 ppm for ^{1}H, ^{13}C, and ^{2}H, respectively and a 1.05-kHz 2 H spin-lock field (SLy) was typically used in the $T1\rho$ experiment. Solvent suppression was achieved using gradient dephasing (gradients g3, g4, g5, g6, and g7).

^{2}H–^{13}C Dz^{2} and ^{2}H–^{13}C DxDz

Measurement of both $R^{Q}(3D_z^2 - 2)$ and $R^{Q}(D_+D_z + D_zD_+)$ rates in $^{13}CH_2D$ spin systems was performed on Varian spectrometers at the Center for Molecular Protein Science (CMPS) at Lund university (Sweden) using pulse schemes adapted from (Millet et al., 2002) (Fig. 3.11).These pulse schemes use a double inept HSQC method to isolate a ^{1}H-^{13}C correlation peak of residues [A,I,L,M,T,V] with quadrature detection achieved using a phase sensitive States-TPPI. Blocks A and B are in serted into the scheme for measurement of $R^{Q}(3D_x^2)$ and $R^{Q}(3D_z^2 - 2)$, while either block C or D can be inserted for the measurement of $R^{Q}(D_+D_z + D_zD_+)$. Carrier frequencies were typically centred at 4.7, 20.0, and 0.8 ppm for ^{1}H, ^{13}C, and ^{2}H, respectively.

Acquisition of backbone relaxation data

Spin-lattice ^{15}N relaxation rates (R_1), rotating frame ^{15}N relaxation rates ($R_{1\rho}$) and heteronuclear steady-state ^{15}N-{^{1}H} NOE (HetNOE) values were obtained using interleaved TROSY-readout pulse sequences (Lakomek, Ying, and Bax, 2012). Temperature compensation was applied in the R_1 experiment by incorporating a spin-lock pulse placed off resonsnce in the inter-scan delay, equal to the longest spin-lock time and the RF power of the $R_{1\rho}$ experiment.

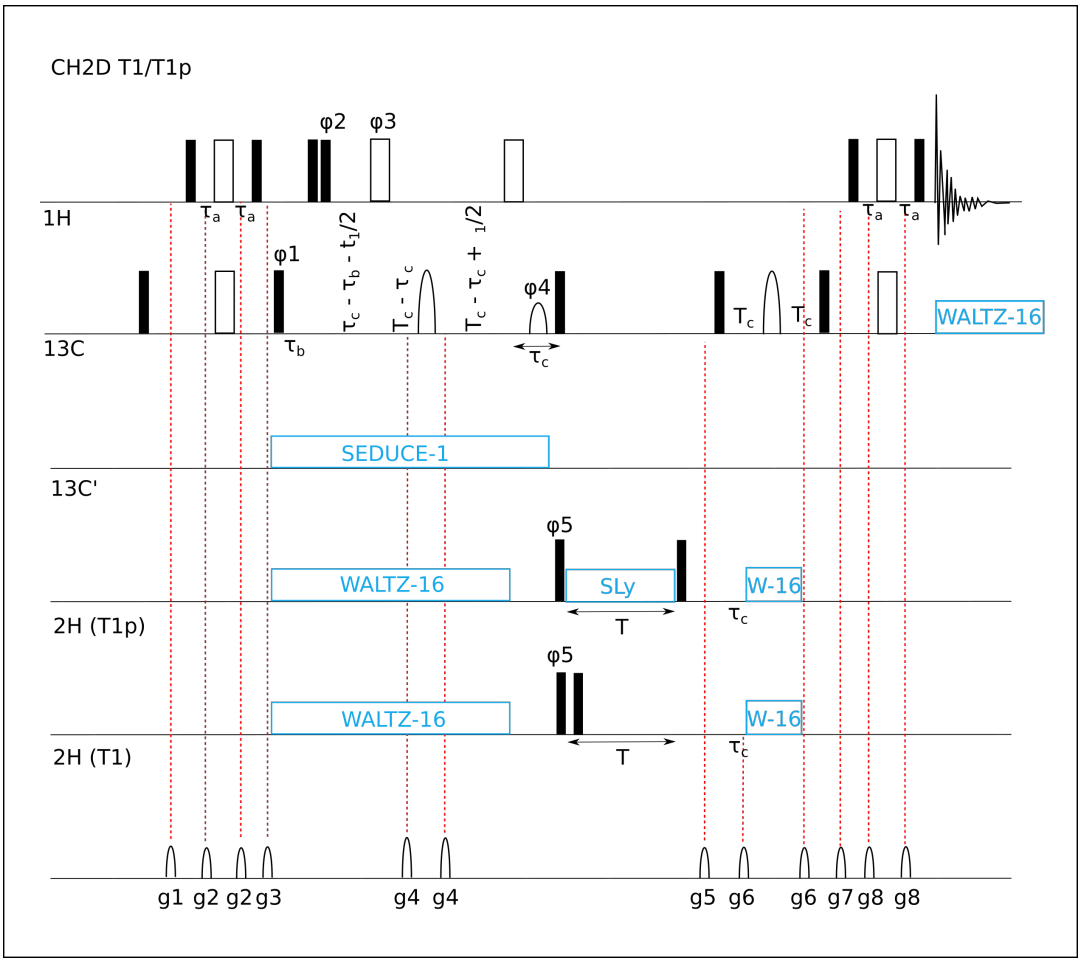


FIGURE 3.10: Pulseprograms for the measurment of sidechain $^{13}CH_2D$ R_1 and $R_{1\rho}$ relaxation rates used in Lund. All narrow (wide hollow) rectangular pulses correspond to 90°(180° pulses about the x-axis unless otherwise indicated. Non-rectangular pulses indicate shaped pulses while blue rectangular pulses denote decoupling schemes. Gradients were applied (often in pairs) along the z-axis as denoted by the hollow semi-ellipses.

Relaxation delays of 0, 80, 240, 400, 400, 640, 800, 1200, 1760, and 2400 ms were typically used to calculate R_1, and delays of 1, 20, 20, 30, 40, 60, 90, 110, 150, and 200 ms were used to calculate $R_{1\rho}$ at 600 MHz and 800 MHz for both complexes. Relaxation delays of 20, 40, 80, 240, 400, 640, 800, 1200, 1200, 1760, 2400, 3200, 4800, 6400, ms were used to calculate R_1, and delays of 1, 5, 5, 10, 15, 20, 20, 40, 60, 90, 110, 140, 160, 200 ms were used to calculate $R_{1\rho}$ were used at 950MHz. Delay times were guided by the optimal sampling of an exponential decay reported by Jones (Jones et al., 1996). The inter scan delay was 3.5 s and the strength of the RF spin-lock field during $R_{1\rho}$ measurement was 1400 Hz at 600MHz, 1866.7 at 800MHz, and 1500 HZ at 950MHz. For the HetNOE measurement, two interleaved experiments were acquired with relaxation delays of 10s.

Experiments were processed in NMRpipe (Delaglio et al., 1995) using a squared sine bell window function, without linear prediction in either dimension. R_1 and R_2 values were determined in PINT (Ahlner et al., 2013; Niklasson et al., 2017) by fitting the integral of the assigned peak to a decaying exponential function across the relaxation series. R_2 values

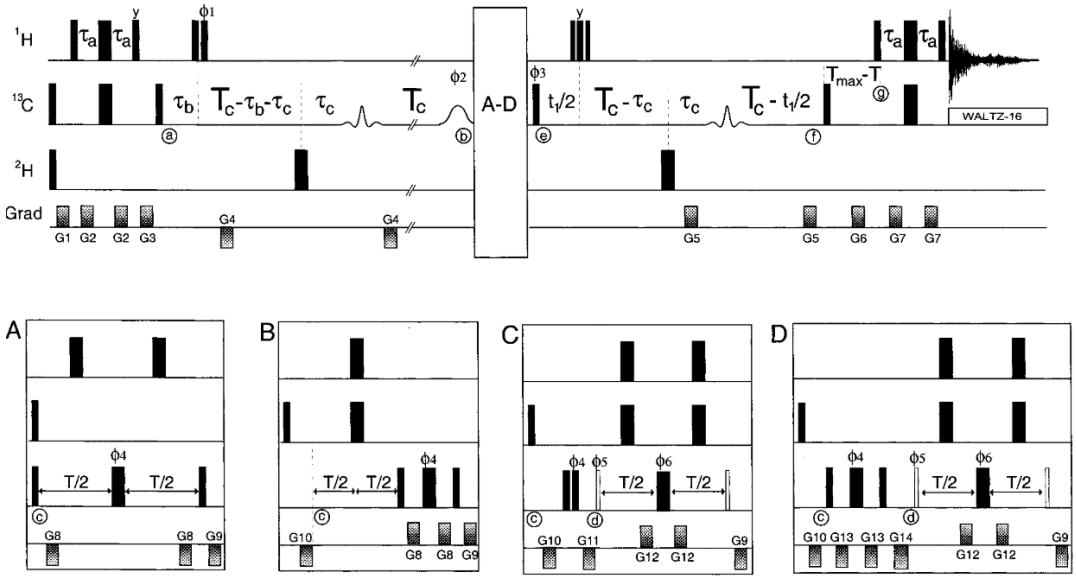


FIGURE 3.11: Pulseprograms for the measurment of sidechain $^{13}CH_2D$ $R^Q(3D_z^2-2)$ and $R^Q(D_+D_z+D_zD_+)$ relaxation rates used in Lund, adapted from Millet et al., 2002. All narrow (wide) rectangular pulses correspond to 90°(180° pulses about the x-axis unless otherwise indicated. Empty narrow rectangular pulses in blocks C and D have flip angles of 45° ^{1}H pulses are centered at 1 ppm prior to point *g* at chich point the carrier frequency is changed to 4.7 ppm. For a detailed description of the pulse scheme, see (Millet et al., 2002)

were calculated in PINT from fitted R_1 and $R_{1\rho}$ values using the relationship in Eq. 3.2. Het-NOE values were also fitted in PINT by calculating the difference in peak integral between saturated and unsaturated spectra, with noise determined from the spectral floor.

$$R2 = \frac{R_{1\rho}}{sin^2\theta} - \frac{R_1}{tan^2\theta} \tag{3.2}$$

Acquisition of sidechain relaxation data

Single and multiple quantum ^{2}H relaxation rates were recorded on 500MHz and 600MHz Varian spectrometers at Lund university. Linear sampling of decay profiles was used (Jones et al., 1996) with delay times of 0.0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.08, 0.1, and 0.03 s for both R_1 and $R^Q(3D_z^2-2)$ and experiments. Delay times of 0.0, 0.002, 0.004, 0.006, 0.008, 0.01, 0.013, 0.016, 0.02, and 0.004 s were used for R_2 and $R^Q(D_+D_z+D_zD_+)$ experiments. Experiments were processed in NMRpipe (Delaglio et al., 1995) using a squared sine bell window function, without linear prediction in either dimension. R_1 , R_2 , $R^Q(3D_z^2-2)$ and $R^Q(D_+D_z+D_zD_+)$ values were determined in PINT (Ahlner et al., 2013; Niklasson et al., 2017) by fitting the integral of the assigned peak to a decaying exponential function across the relaxation series.

Acquisition of backbone relaxation dispersion data

Experiments were processed in NMRpipe (Delaglio et al., 1995) using a squared sine bell window function, without linear prediction in either dimension. RD values were determined in PINT by fitting the integral of the assigned peak across the relaxation series to the Bloch - McConnell equations (Ahlner et al., 2013; Niklasson et al., 2017).

Relaxation dispersion (RD) experiments of backbone amide protons in perdeuterated, amide-protonated samples in standard NMR buffer were performed in Sheffield using the pulse program outlined in Fig. 3.9 (Sugase et al., 2007). The L2 parameter (L4 = L2 x 2) was varied across an experiment series: 0, 16, 1, 24, 5, 80, 2, 60, 3, 40, 8, 1, 80, 5, 40, 16, 30, 12, 70, 10, sampling exchange processes up to 2 kHz. Several duplicate experiments were recorded for jacknife error approximation in PINT. Interleaving of fids was not implemented during the data acquisition, nor was a temperature compensation block as the proton decoupling was likely to introduce minimal sample heating. However 2D spectra were roughly interleaved in that experiments with high and low L2 values (for the number of CPMG cycles) were immediately juxtaposed in an attempt to reduce sample heating over time. Experiments were processed in NMRpipe (Delaglio et al., 1995) using a squared sine bell window function, without linear prediction in either dimension. RD values were determined in PINT (Ahlner et al., 2013; Niklasson et al., 2017) by fitting the integral of the assigned peak to the standard Bloch - McConnell equations across the relaxation series

3.3.5 Modelfree analysis - Backbone amides

Modelfree (MF) analysis (Lipari and Szabo, 1982a; Lipari and Szabo, 1982b; Halle, 2009) was performed using *relax* (d'Auvergne and Gooley, 2003; d'Auvergne and Gooley, 2006; d'Auvergne and Gooley, 2007; d'Auvergne and Gooley, 2007; d'Auvergne and Gooley, 2008; Bieri, d'Auvergne, and Gooley, 2011). R_1, $R_{1\rho}$, with parallelization achieved using *mpi4py* (Dalcín, Paz, and Storti, 2005; Dalcín et al., 2008; Dalcin et al., 2011). HetNOE values were used with backbone amide coordinate geometry provided by corresponding crystal structures to the complex under investigation. Model free analysis was performed both with and without the extended MF formula presented by Clore and co-workers (Clore et al., 1990)), and the resulting fits were assessed using Akaike's Information Criterion (AIC) (Akaike, 1998). The resulting ellipsiodal diffusion tensors closely matched the geometry of the input crystal structures, and global correlation times reflected those obtained from single-field analysis using ROTDIF from the Fushman lab (Fushman, Xu, and Cowburn, 1999; Varadan et al., 2002; Fushman et al., 2004; Walker, Varadan, and Fushman, 2004). Consistency testing of multiple field data was performed using python scripts presented in Morin and M. Gagné, 2009 which utilize relax.

3.3.6 Modelfree analysis - Sidechain methyls

MF analysis of sidechain methyl groups was performed as outlined in Skrynnikov, Millet, and Kay, 2002, with consistency testing performed using in-house python scripts. CH vectors were extracted from the same PDB files as used for the backbone amide MF analysis. Model selection and parameter selection was performed using in-house (Lund) Matlab

scripts using the three LS models presented in (Skrynnikov, Millet, and Kay, 2002).

Now for something completely different . . .

Chapter 4

Results and Discussion

4.1 Paper I: van der Waals contact between nucleophile and transferring phosphorus is insufficient to achieve enzyme transition state architecture.

Just over a decade ago, a breakthrough paper (Lahiri et al., 2003) was published purportedly showing a phosphate group being "caught in the act" of being transferred between substrate and enzyme. The paper stimulated the late, great enzymologist, Jeremy Knowles, to write a Perspectives article entitled "Seeing is Believing" (Knowles, 2003), heralding that models of enzymes catalyzing these central processes of all life now had both theoretical and experimental support. Careful re-examination of the experiment by our group (Baxter et al., 2006; Baxter et al., 2010) established that the enzyme in question, βPGM, had in fact tricked the experimentalists by synthesizing in the active site what was at the time a highly unusual (Lahiri et al., 2003) transition state analogue (MgF_3) from the components of the crystallization buffer. However, the discovery of MgF_3 spawned a whole new field of examination of trifluoromagnesate complexes in the literature (recently reviewed in (Jin, Molt, and Blackburn, 2017b) and (Jin et al., 2017b)), but left the original breakthrough unsatisfied.

In this study a βPGM variant with a constitutively protonated general acid-base is shown to self-synthesise complexes that contain the native reaction intermediate but are arrested in the course of phosphoryl transfer. The complexes have everything in place for catalysis except that the enzyme will not release a proton to the substrate and, thereby, are truly "caught in the act" of phosphoryl transfer. These new complexes of βPGM detect two hitherto unidentified facets of enzyme catalysis. Firstly, the enzyme will not make the conformation it uses to stabilize the transition state until it forces the nucleophile (oxygen) well inside the van der Waals radius of the electrophile (phosphorus). This demonstrates the vital interplay between the preferences of the protein and the preferences of the reacting atoms. Secondly, the enzyme will not appropriately coordinate the essential Mg^{2+} ion, which is present in virtually all enzymes that transfer phosphoryl groups (kinases, phosphatases, ATPases, G-proteins, polymerases, nucleases, etc.) until, again, it forces the nucleophile well inside the van der Waals radius of the electrophile. This directly shows that the enzyme is using metal ion coordination to complement only the transition state, at the expense of ground states that are structurally very similar. While currently focusing on βPGM, the study opens up the opportunity to investigate a wide range of phosphoryl transfer enzymes in their native reaction complexes.

In summary, this study provides long-sought experimental insight into enzyme-catalysed phosphoryl transfer reactions, prematurely heralded in 2003. It illustrates to the field two primary mechanisms by which these enzymes discriminate between the transition state (binding it tightly enough to have a sufficiently fast chemical step) and reactants (binding it weakly enough that it does not dissociate too slowly). This is an especially difficult problem for phosphoryl transfer enzymes owing to the inherent stability of the substrates. The study also provides the field with the first experimental verification of the long-standing argument regarding the role of proton transfer in catalysis by this vitally important superfamily of enzymes. This in turn enables the field to determine the validity of the multitude of theoretical models purporting to rationalize the incredible levels of catalysis by these enzymes. Finally, the study brings into focus the credibility of conclusions derived from theoretical approaches to modelling any enzyme reaction, since the associated conformational changes observed here are not normally accommodated in the models.

Section adapted from the letter to the editor on ACS manuscript submission (Johnson et al., 2018).

4.2 Paper II: X-ray, NMR and QM approaches reveal the relationship between protein conformational change, proton transfer, and phosphoryl transfer in an archetypal enzyme

Molecular details for the timing and role of proton transfer in phosphoryl transfer reactions are poorly understood. Using a combination of NMR, X-ray crystallography and DFT approaches, we characterize pre- and post- proton transfer models of a phosphoryl transfer reaction in the archetypal phosphoryl transfer enzyme βPGM. We observe that the ionic nature of the AlF_4^- TSA may be highly useful in the investigation of proton transfer in phorphoryl transfer enzymes as out-of-plane distortion of the central Al^{3+} ion closely correlates with proton timing across the reaction coordinate. Backbone order parameters (ps-ns rigidity measurement) were used to guide QM model generation and residue truncation in pre- and post- proton transfer TSA models. The TS model displays a key contribution of this proton transfer to/from the GAB on the charge distribution within the transferring group, and consequently, the electrostatic interactions with surrounding residues in the active site. Given the free energy profile of the reaction, the GS of the reaction indicates that a high degree of proton transfer has already occurred to substrate which is closely reflected in ^{19}F and X-ray crystallographic observations which may further empower the use of ^{19}F NMR in the investigation of phosphoryl transfer reactions.

The negative charge on phosphate monoesters provides a strong repulsion to potential attacking nucleophiles, which contributes to these compounds being extremely inert in aqueous solution but labile in the active sites of phosphoryl transfer enzymes (Lad, Williams, and Wolfenden, 2003). Some phosphoryl transfer enzymes populate near attack complexes (NACs) in which the attacking hydroxyl group hydrogen bonds with the transferring phosphate group in a nonproductive orientation (Griffin et al., 2012; Jin et al., 2017b; Jin, Molt, and Blackburn, 2017a). The residue that provides general acid-base (GAB) catalysis is utilized not only to activate the hydroxyl group for nucleophilic attack of the target phosphate group, but also to stimulate the alignment of the nucleophilic oxygen atom with the phosphorus atom. Structural investigations of near transition state (TS) species have made use of both MgF_3^- and AlF_4^- as transition state analogues (TSAs) that mimic the transferring phosphate group, as they are both reportedly planar, and have a net single negative charge when complexed with substrate in the enzyme active site (Baxter et al., 2008; Cliff et al., 2010; Jin et al., 2014; Jin et al., 2017b; Jin, Molt, and Blackburn, 2017a). The TSA structures have indicated that the engagement of GAB residues is concurrent with phosphoryl group transfer (Dai et al., 2009; Baxter et al., 2010; Griffin et al., 2012; Johnson et al., 2018). However, controversy remains as to the timing of proton transfer associated with GAB catalysis meaning that any interpretation of the mechanism and the energy barrier of the chemical step is unreliable. An archetypal phosphoryl transfer enzyme that utilizes GAB catalysis for the attack of a phosphate by a hydroxyl group is β-phosphoglucomutase (βPGM) [EC 5.4.2.6], which catalyzes the reversible isomerization of β-glucose 1-phosphate (βG1P) and glucose 6-phosphate (G6P) via a β-glucose 1,6-bisphosphate (βG16BP) intermediate, and has been well characterized enzymatically and structurally. Previous computational studies on the phosphoryl transfer between βG16BP and residue D8 of βPGM (generating G6P), have

presented conflicting timings for the proton transfer associated with the GAB residue (D10). Analyzing the trajectories in the direction of phosphoryl group transfer from βG16BP to D8, these studies range in prediction from "early" (Webster, 2004; Marcos, Field, and Crehuet, 2010), through "concerted" (Barrozo et al., 2018), to "late" (Elsässer, Dohmeier-Fischer, and Fels, 2012) proton transfer events, with predicted barrier heights ranging from 14 to 64 kJ mol^{-1}.

Mutation of the GAB residue in βPGM to a constitutively protonated mimic (Asn) was found to trap a ground state analog (GSA) complex in which the phosphorus atom of the 1-phosphate group of βG16BP is at van der Waals contact distance from the nucleophilic carboxylate oxygen of D8 (Paper I; (Johnson et al., 2018)). This observation demonstrated that without proton transfer from the GAB to the bridging oxygen of βG16BP, the phosphate group prefers to remain associated with βG16BP. This is consistent with an "early" proton transfer step during phosphorylation of D8 by βG16BP, and suggests that the N10 variant provides a good model of the pre-proton-transfer state. However, in order to determine whether the GAB proton is likely to be transferred before the peak of the energy barrier in the native reaction, a post-proton-transfer model is also desirable. In this report, we establish that the metal fluoride complex of WT βPGM with G6P provides a suitable post-proton-transfer model and the equivalent complex of the D10N variant is a directly comparable model of the pre-proton-transfer state. While the metal fluorides act as surrogates for transferring phosphoryl groups in the transition state, they have reduced atomic charges and little covalency in their bonding (Griffin et al., 2012; Jin et al., 2017b; Jin, Molt, and Blackburn, 2017a). Correspondingly, they are shown to report on the electronic distribution within the active site pre- and post- proton-transfer, since they distort their geometry in line with the preferred positions of the phosphorus atom in each scenario. In parallel, solution NMR methods are used to calibrate DFT calculations to generate reliable models of the reaction trajectory for phosphoryl transfer. Collectively, these studies establish the timing of proton transfer in this reaction using a novel approach that is applicable to other phosphoryl transfer enzymes that rely on GAB catalysis.

The manuscript concludes that, the experimental X-ray structures and NMR measurements, in combination with QM models and their partitioning into atomic basins, all point towards a synergy between proton transfer from the GAB residue to the bridging oxygen of the phosphate group and the early stages of phosphoryl group dissociation. Both of these processes are assisted by the transition of the protein conformation between that poised by the Asp to Asn mutation of the GAB residue in the βG16BP complex and that adopted to stabilise the transition state in the metal fluoride TSA complexes. The data collectively also show that the distortion of the predominantly ionic metal fluoride TSA moieties can be used to report on the relative charges, in situ within the enzyme while in the near-TS conformation, of the axial oxygens that constitute the nucleophile and the leaving group for the reaction. ^{19}F NMR measurements, previously proposed as simple reporters of the electronic environment of the equatorial phosphate oxygens in phosphoryl transfer reactions(Jin et al., 2016; Jin et al., 2017a), can also be used in combination with QM models to corroborate the protonation state of the nucleophile and the leaving group oxygen atoms in the TSA complexes, thereby validating the reliability of the QM model. These data further establish that mutation of the GAB residue from Asp to Asn serves as a good model of the pre-proton transfer state (when considering phosphoryl transfer from βG16BP to βPGM) and that the corresponding

WT complexes serve as a good post-proton-transfer model in a GAB catalyzed phosphoryl transfer reaction.

4.3 Paper III: Arg - phosphate interaction in β-phosphoglucomutase improves substrate affinity, but introduces inhibition

Under biological conditions, phospho mono- (R-P-O-R) and di-ester (R-O-P-O-R) bonds have half-lives of millions of years (Lad, Williams, and Wolfenden, 2003). This inherent stability is essential to ensure the high-fidelity storage of our genetic information within DNA. However, many core biological processes, including DNA processing, metabolic cycles and cell signaling, depend on the efficient transfer of phosphate groups between metabolites, requiring an enzyme catalyst that can break these inert bonds. Phosphoryl transfer enzymes have evolved with some of the largest rate accelerations known to biology, with typical catalytic rate enhancements (k_{cat}/k_{non}) of 10^{21} (Lad, Williams, and Wolfenden, 2003; Lassila, Zalatan, and Herschlag, 2011; Kamerlin et al., 2013). Phosphoglucomutase enzymes are one class of phosphoryl transfer enzymes, found in both prokaryotes and eukaryotes, that reversibly produce glucose 6-phosphate, an important precursor for glycolysis and energy production in cells. β-phosphoglucomutase (βPGM) [EC. 5.4.2.6] from *Lactococcus lactis* is a well-characterized (Lahiri et al., 2004; Zhang et al., 2005; Dai et al., 2006; Baxter et al., 2006; Baxter et al., 2008; Baxter et al., 2010; Jin et al., 2014; Johnson et al., 2018) magnesium-dependent phosphoryl transfer enzyme, which catalyzes the reversible isomerization of β-glucose 1-phosphate (βG1P) to glucose 6-phosphate (G6P) via a β-glucose 1,6-bisphosphate (βG16BP) intermediate using a ping-pong bi-bi reaction mechanism (Dai et al., 2006). As part of this mutase reaction, the enzyme adopts two different catalytically competent states, the substrate-free state and the phospho-enzyme state (βPGM^P, phosphorylated at residue D8), which have different substrate specificities. The active site of βPGM is located at the interface between the helical cap domain (T16-V87) and the α/β core domain (M1-D15, S88-K216) and opening and closing of the cap domain relative to the core domain occurs during the catalytic cycle. The active site contains a magnesium(II) ion binding site and two phosphate binding sites. One phosphate binding site, termed the *proximal* site, is adjacent to the Mg^{2+} ion at the catalytic center, while the other phosphate binding site, termed the *distal* site, is *ca.* 10 Å removed from the catalytic center. The *distal* phosphate binding site has a role in anchoring ligands in the active site via interactions with several conserved residues (S116, K117, R49), together forming a positive electrostatic region in the substrate-free enzyme. The interplay between the two phosphate binding sites allows βPGM to bind substrates, intermediates and products in two orientations to facilitate mutase activity - the ability to both transfer and remove a phosphate from a substrate using the same active site residues.

Kinetic characterization of the βPGM-catalyzed conversion of βG1P to G6P using a spectrophotometric coupled assay has previously identified a lag-phase prior to steady-state catalysis resulting from two components (Golicnik et al., 2009). The first component is a chemical equilibration, where it was modelled that catalysis was retarded until sufficient βG16BP intermediate was generated to efficiently prime the enzyme for catalysis (by phosphorylation of residue D8). This is termed the βG1P-independent component of the lag-phase herein. The second component was modelled as βG1P binding to un-phosphorylated βPGM with a K_i of (122 ± 8 μM (Golicnik et al., 2009)) and is termed the βG1P-dependent component herein. In this model the formation of a non-catalytically active βPGM:βG1P complex could preclude βG16BP binding to the active site and prevent regeneration of

βPGM^{P} and further catalysis. However, there is currently no structural or mechanistic evidence to explain substrate inhibition by βG1P or how the lag phase can be alleviated.

Here we have structurally characterized βG1P bound to the βPGM enzyme in a non-catalytically competent, closed complex, which provides a structural basis for the βG1P inhibition of the substrate-free enzyme postulated previously in kinetic models (Golicnik et al., 2009). Furthermore, it is demonstrated that single mutations in the *distal* phosphate site can alleviate the βG1P-dependent component of the lag-phase prior to steady state catalysis, implicating a role for the specific bidentate hydrogen bonding interaction between phosphate in the *distal* site and the terminal guanidinium group of residue R49 in the cap domain. MgF_3^- and AlF_4^- transition state analogue complexes with G6P demonstrate minimal perturbation to the *proximal* phosphate binding site (at the point of phosphoryl transfer) in response to removal of the R49 guanidinium group in the *distal* phosphate binding site. This minimal communication between the two sites indicates that the role of the *distal* site is primarily to recruit ligand into the active site and induce domain closure prior to the chemical step, while playing minimal role in the chemical step itself.

Both MgF_3^- and AlF_4^- TSA crystal structures of the βPGM$_{R49K}$ and βPGM$_{R49A}$ variants complexed with G6P demonstrated a redundancy in the phosphate coordination in the *distal* phosphate binding site. In the βPGM$_{R49K}$ variant, the positive charge was maintained in the *distal* site, but with a reduced hydrogen bonding capacity. This manifests as a reduced stability for substrate bound complexes, but given that the charge is maintained in this conservative mutation, the small reduction in catalysis is readily rationalizable. In the βPGM$_{R49A}$ variant, though charge balance was removed from the *distal* site on the cap-domain side, there was still a substantial level of activity. In both of the TSA structures with G6P, K117 from the core-domain (which is solvent exposed in βPGM$_{WT}$ and βPGM$_{R49K}$ complexes), is repositioned in order to coordinate the *distal* phosphate group of G6P via its sidechain amine group. If this occurs in solution, which is consistent with the chemical shift changes in the $^{1}H^{15}N$-TROSY spectra, then this presents a redundancy in phosphate binding capacity in the *distal* phosphate binding site of βPGM. This conformer also indicates that a competitive binding interaction between K117 and R49 may exist in βPGM$_{WT}$. This alternative binding partner for phosphate groups in the *distal* site could present a pathway to ligand dissociation from the active site, prior to either reorientation of βG16BP or product release of G6P or βG1P.

Interestingly, the closed βPGM$_{D170N}$:βG1P complex closely resembles fully closed TSA structures of phosphoryl transfer, with key residues in the active site adopting catalytic orientations, but without a phosphoryl group to transfer and without a metal ion in the active site. The sidechain of residue N170 is rotated away from the *proximal* phosphate site, which potentially implicates D170 dissociation from the active site as a mechanism to release the Mg_{cat} ion in the WT enzyme. This may be important in the dissociation of the reaction intermediate βG16BP, since it has a high affinity to holo-βPGM (Mg_{cat} bound; Kd = 0.8 $\pm$ 0.2 μM (Golicnik et al., 2009)). However this protein complex displays a weak Mg_{cat} binding affinity (Johnson et al., 2018). Given that the βG16BP ligand has a higher binding affinity than βG1P or G6P ligands for the open-βPGM enzyme, it is important that this state does not become a kinetic trap. In order to avoid this, it is tempting to speculate that βPGM uses one (or both) of the above ligand dissociation pathways.

Guanidinium - phosphate interactions have been reported to provide substantial binding energies in the range 11-13 kcal/mol for glycerol 3-phosphate dehydrogenase (GPDH) (Tsang, Amyes, and Richard, 2008), for triose phosphate isomerase (TIM) (Amyes, OD́onoghue, and Richard, 2001), and for orotidine 5-monophosphate decarboxylase (Amyes, Richard, and Tait, 2005). This phosphodianion binding has been associated with a protein conformational change and active site assembly in other systems such as orotidine 5-monophosphate decarboxylase (Desai et al., 2012; Reyes, Amyes, and Richard, 2016) and for GPDH (Reyes et al., 2015). Furthermore, it has been demonstrated that the energetic cost of disconnecting groups of either substrate or enzyme GPDH (Tsang, Amyes, and Richard, 2008; Go, Amyes, and Richard, 2010; Reyes, Amyes, and Richard, 2016) and TIM (Go, Amyes, and Richard, 2010; Zhai, Amyes, and Richard, 2014) was directly reflected in the reduction in observed reaction rate. This suggests that the transition state of the reaction in the re-assembled complex closely reflected transition states of the native reaction (reviewed (Amyes and Richard, 2013)). In βPGM, binding of the phosphodianion to the *distal* site is insufficient to close the enzyme, furthermore, the inclusion of both phosphate (or analog) and glucose is insufficient to re-assemble the inhibited βG1P-bound complex. This is unsurprising as the Ki for the βG1P-dependent contribution to the lag phase is reportedly 122 $\pm$ 8 μM (Golicnik et al., 2009), however, the minimal impact of reduced hydrogen bonding capability in the *distal* site is more surprising. Given that the bidentate interaction between the sugar-associated phosphate group and the guanidinium group of R49 is well conserved substrate bound complexes in βPGM, it is surprising how well the reduction in observed rate correlates with energy associated with the loss of a hydrogen bond. This contrasts with the significant contribution of an Arg group binding to phosphate in GPDH, where mutation to alanine resulted in a 9.1 kcal mol^{-1} destabilization of the transition state for enzyme catalyzed reduction of DHAP (Reyes, Amyes, and Richard, 2016). This implies that the chemical step of the reaction in βPGM may not have been affected, which is highly consistent with the ^{19}F NMR of the TSA complexes where mutation of the R49 group had not affected chemical environment of the *proximal* site at the point of phosphoryl transfer to substrate.

Taken together, these observations illustrate some of the elegant mechanisms that enzymes employ in order to achieve the significant rate enhancements necessary for life. Here we see a *ca.* 10 fold rate enhancement through the use of a guanidinium group (WT) over an amine group (R49 variants) to coordinate the phosphate in the *distal* site (at approximately physiological βG1P concentrations). This rate enhancement, however, is at the expense of introducing a source of inhibition to catalysis – inhibition of the substrate-free enzyme by its initial substrate βG1P. Together this is an example of the elegant trade off present in numerous phosphoryl transfer enzymes, that of balancing substrate affinity with potential inhibitory consequences by introducing substrate inhibition.

4.4 Paper IV: Mechanisms of phosphatase activity in good and bad phosphatases of the HAD superfamily

Phosphoryl transfer enzymes play a key role in biology, with vital roles in metabolism, cell signaling, and manipulation of genetic material. These enzymes can be broadly categorized into phosphatases, phosphotransferases (eg. kinases), and mutases. While in phosphatase enzymes the phospho-enzyme state is destabilized such that spontaneous autodephosphorylation is promoted, mutase enzymes need to stabilize a phospho-enzyme state in order to efficiently perform a ping-pong bi-bi reaction mechanism. To investigate how specific phosphatase vs. mutase activity has diverged, two enzymes from the well characterized haloacid dehalogenase (HAD) superfamily were selected. Here we show that a specific phosphatase (Phosphoserine phosphatase (PSP) from *Methanococcus jannaschii*) employs several mechanisms that promote phosphatase activity compared to a mutase (β-phosphoglucomutase (βPGM) from *Lactococcus lactis*), which actively employs mechanisms to prevent such activity. These themes can be roughly partitioned into three areas; translation of catalytic machinery, dislocation of solvent from the transferring phosphate group, and rotation of the phosphate group on a catalytically relevant timescale. All three themes act to ensure that βPGM acts as a mutase not a phosphatase, and that PSP acts as a phosphatase, not a phosphotransferase and together present tools for the future design of enzymes in either class.

Here three themes (translation, dislocation, and rotation) have been described that distinguish the specific phosphatase activity of PSP from the specific mutase activity of βPGM in a key enzyme superfamily (Huang et al., 2015). Given the previous crystal structures of PSP using the GAB residue (D13) to align water for nucleophilic attack on the phospho-enzyme (Wang et al., 2002), there was little debate of the direct role that the GAB played. In βPGM where the same catalytic DXD motif could utilize the GAB residue (D10) for the same purpose, several crystal structures presented here indicate that it does not as was predicted previously (Griffin et al., 2012; Johnson et al., 2018). In these structures residue D10 not only adopts a rotamer that is both rotated *out* of the active site, but the residue is also translated away such that rotation from *out* to *in* is insufficient to align a water molecule for nucleophilic attack. In a structure where a partial *in* occupancy is observed, the GAB indirectly coordinates a nucleophilic water molecule (via a second water molecule), which may form the basis of a proton transfer network between nucleophilic water molecule to the GAB (Allen and Dunaway-Mariano, 2016). However, the minimal perturbation of the dephosphorylation rate when the GAB residue is mutated to a constitutively protonated mimic eliminates this possibility (Johnson et al., 2018). The comparison in PSP is that in both the PSP:BeF_3 and PSP:MgF_3 structures indicate that the GAB residue (D13) occupies an *in* rotamer that aligns water for nucleophilic attack on the phosphate group.

Key active site differences exist between PSP and βPGM beyond the GAB *in* to *out* transition. In PSP the positioning of polar groups in the active site is near identical when either L-Ser or water are accommodated. This presents a mechanism whereby the active site in PSP acts to specifically orient a shell of water molecules around an activated water nucleophile in the same manner that a ligand is usually coordinated in the active site of an enzyme. This water *as-substrate* model is relatively robust with local perturbation to affected fluoride groups as a result of no-observed rotation of the phosphate group mimic in the active site. By contrast,

open-βPGM structures with either phosphate (or transferring phosphate) surrogates coordinated by poorly defined water molecules in solution, or crystallographically. This indicates that in βPGM, one of the protection mechanisms of the high energy phosphate group is to expose it to bulk, unstructured, solvent, rather than preclude it from solvent.

Chemical exchange of the phosphate surrogate (BeF_3^-) is observed in βPGM but not in PSP, even when the number of coordinating groups to the phosphate is made equivalent in both enzymes through mutation. This chemical exchange is the result of rotation of the BeF_3^- moiety around the apartyl Oδ1 – bond and suggests a tightly controlled position of the BeF_3^- group in PSP which is not present in βPGM. The observation that removal of a coordinating positive charge (sidechain amine of K145) had only a moderate effect on this exchange process suggests that there are larger contributors to this exchange process. One possibility is that there is an underlying conformational dynamic in the active site of βPGM that manifests as a rotation of the BeF_3^- moiety. A second possibility is that a more charged species is dominating the electrostatic environment surrounding the BeF_3^- group. Given its proximity and reportedly poor affinity both catalytically (Golicnik et al., 2009) and structurally (Johnson et al., 2018), it is tempting to speculate that dissociation of the catalytic Mg^{2+} ion may be responsible.

It has been asserted that as enzymes evolve towards a specific function, they rigidify (although no specific timescale was given, and a catalytically relevant one is assumed) (Tokuriki and Tawfik, 2009; Dellus-Gur et al., 2015; Pabis, Duarte, and Kamerlin, 2016; Petrović et al., 2018). By both X-ray crystallography and solution NMR, it is observed that PSP coordinates both L-Ser and water stably in the active site, with well defined and robust water network. Contrastingly, the active site of βPGM exposes the transferring phosphate group to unstructured solvent which is typically more plastic in nature than protein residues. Herein lies a potential distinction between the two enzymes, PSP binds water *as-substrate* in a stable manner conducive to specific phosphatase activity, whereas βPGM employs several conformational and geometric measures to prevent this from happening.

Chapter 5

Discussion and future directions

This thesis has been primarily concerned with how we can use this archetypal phosphoryl transfer enzyme, βPGM, to investigate the phosphoryl transfer process in general. In order to corroborate observations and justify claims, a multidisciplinary approach was adopted that utilized NMR spectroscopy, X-ray crystallography, kinetic assays, as well as DFT QM calculations. As a general rule, convergence between all disciplines was sought to validate each phenomenon. This cross disciplinary approach was only made possible through the concerted effort of several people (often in different institutions), to which the author owes a great deal of thanks. The principal outcomes of this thesis are outlined in the following sections, which will hopefully tie-in to several of the key themes outlined in the Introduction.

5.1 The D10N mutation

The mutation of general acid-base (GAB) aspartate residues in enzymes is an immensely powerful tool for structural biology. In the case of Paper I, we managed to trap native substrate in the active site at Van der Waals contact distance from the nucleophilic carboxylate oxygen. Not only did this work allow the investigation of native substrate approaching a catalytic conformation, but it directly demonstrated that the pathway from open to closed transition state protein architecture was a non-linear event. This dog-leg possibility is often omitted from computational calculations of reaction trajectories and could lead to significant errors in final predictions. One implication of this is parallel catalytic pathways, with these structural investigations presenting some mechanistic evidence towards supporting an entropic contribution towards an activation energy barrier. This study was initially designed to address the question of the role of the GAB outlined in section 1.6.3 in a more direct manner than had previously been adopted. The implication was that additional protein conformational changes from the GS may also be necessary in order to stabilize the TS in enzyme catalyzed phosphoryl transfer. This work led onto the investigation of the D10N variant as a pre-proton-transfer model when the G6P ligand is complexed with AlF_4^- which formed the starting point of Paper II.

5.2 Investigations of the implications of proton transfer

Following from Paper I, Paper II continues to address the role of the GAB in the catalytic mechanism of βPGM. This paper advocates for an additional use of AlF_4^- groups, namely, the investigation of proton transfer in phosphoryl transfer enzymes. Here we establish that the Al^{3+} atom of the AlF_4^- moiety displays a distortion towards the side of the reaction where the proton transfer event occurs. While the readout of this is a 0.15 Å migration of the Al^{3+} atom and subtle distortion of the AlF_4^- moiety when observed crystallographically, ^{19}F 1D NMR shows a more pronounced effect. The ^{19}F NMR shifts of the AlF_4^- group transition from typical 6 coordinate Al^{3+} to 5 coordinate AlF_x ^{19}F shifts as the Al^{3+} atom moves towards the sugar 1-oxygen atom resulting in an average downfield shift of 4 ppm. These observations are corroborated by DFT calculations, which allows further investigation of principal interaction partners across the phosphoryl transfer event. While this paper initially sought to solve a controversy over the timing of proton transfer in βPGM, one of the principal outcomes is that AlF_4^- may report on more than was previously thought. Previous investigations demonstrated that AlF_4^- provided a highly sensitive reporter on the electrostatic environment of the active site (Jin et al., 2016; Jin et al., 2017a). This study demonstrates that it can also be a highly sensitive tool for the investigation of proton transfer timing in enzymes that utilize GAB catalysis. Furthermore, the protein conformational change from the βPGM_{D10N}:βG16BP complex in Paper I to the TSA protein architecture described in Paper II, is associated with partial bond cleavage of both P-O and O-H bonds. The corollary of this argument is that this conformational change may be necessary in order to stabilize the product state.

5.3 A single hydrogen bond results in a catalytic lag phase in βPGM

Moving away from direct mechanistic investigations of the chemical step, Paper III is an example of the elegant trade off present in many enzymes, that of balancing substrate affinity with potential inhibitory consequences. As an added bonus, this work also permitted the investigation of through-substrate communication in catalysis, or lack thereof. We discern a disconnect between transferring (*proximal*) and non-transferring (*distal*) phosphate sites at the point of phosphoryl transfer, with key roles for the two sites becoming more defined. Namely, the principal role of the *distal* site is to recruit ligand, with minimal role in the chemical transfer, while the *proximal* site may both recruit ligand (*via* the side-chain amine of K145) and perform catalysis. This work presents a very elegant trade-off between substrate affinity and inhibition with weak inhibition in a non-essential enzyme (Levander, Andersson, and Rådström, 2001) the result of a 10 fold increase in observed rate at approximately physiological βG1P conditions.

5.4 Do you want to build a phosphatase? Come on, let's find a way!

Paper IV was initially designed to address the native phospho-enzyme controversy (outlined in section 1.6.2). Once the native phospho-enzyme structures were obtained, the project became more about how the βPGM enzyme dephosphorylated, than how it became phosphorylated. Once these states were obtained the comparison of βPGM (a poor phosphatase) with PSP (a good phosphatase) became the obvious next step. Given the work done previously in the Kim and Wemmer labs (Wang et al., 2001; Wang et al., 2002), alongside work done by previous PhD students in the Waltho lab, three key themes could be teased out of the combined effort. These themes each indicate mechanisms of how βPGM had evolved to promote mutase activity over phosphatase activity.

Firstly, translation of catalytic machinery away from the phospho-enzyme site plays a role in preventing alignment of water for nucleophilic attack by a general acid-base. Secondly, rotation of the phosphate group and local millisecond conformational dynamic prevents water stabilization around the high energy phosphate group and this dynamic was not observed in the specific phosphatase PSP. Finally, and most surprisingly, the exposure of phospho-enzyme to bulk water reduces hydrolysis rate. This final observation is borne out of the unstructured water observed in βPGM TSAs compared to the well ordered water molecules observed in PSP (both hydrolysis TSA and ligand polar groups for the phosphatase reaction). Together these observations point towards a more rigid, pre-organized active site in (compared to βPGM^{P}) when the active site is fully solvated. While mechanisms of orbital steering are not being invoked, the structure and alignment of reactive groups points towards a key mechanism of specifically increasing phosphatase activity.

5.5 Future directions

Building from the work in Paper III, investigation of ligand binding to the open enzyme could go a long way towards explaining the observed "lag-phase" phenomena. To this end, a characterization of the open enzyme is in process, with a view to dynamic landscape characterization that was so productive for DHFR, and Abl/Src kinases. Furthermore, given the high sensitivity of ^{19}F probes to the electrostatic environments surrounding MFx TSAs (Paper II and references therein), a quantifiable contribution of the active site to the energy barriers associated with catalysis is now achievable. Building from Paper II, the precise role of electrostatic interactions in the active site of βPGM are under way, using ^{19}F chemical shifts to validate computational models of phosphoryl transfer under non-ideal reaction conditions.

> *We shall not cease from exploration, and the end of all our exploring will be to arrive where we started and know the place for the first time.*
>
> TS Eliot

Bibliography

Ahlner, Alexandra et al. (2013). "PINT: a software for integration of peak volumes and extraction of relaxation rates". In: *Journal of Biomolecular NMR* 56.3, pp. 191–202. ISSN: 1573-5001. DOI: 10.1007/s10858-013-9737-7. URL: https://doi.org/10.1007/s10858-013-9737-7.

Akaike, Hirotogu (1998). "Information Theory and an Extension of the Maximum Likelihood Principle". In: *Selected Papers of Hirotugu Akaike*. Ed. by Emanuel Parzen, Kunio Tanabe, and Genshiro Kitagawa. New York, NY: Springer New York, pp. 199–213. ISBN: 978-1-4612-1694-0. DOI: 10.1007/978-1-4612-1694-0_15. URL: https://doi.org/10.1007/978-1-4612-1694-0_15.

Allen, Karen N. and Debra Dunaway-Mariano (2003). "Response to Comment on "The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction"". In: *Science* 301.5637, pp. 1184–1184. ISSN: 0036-8075. DOI: 10.1126/science.1087309. eprint: http://science.sciencemag.org/content/301/5637/1184.4.full.pdf. URL: http://science.sciencemag.org/content/301/5637/1184.4.

Allen, Karen N and Debra Dunaway-Mariano (2016). "Catalytic scaffolds for phosphoryl group transfer". In: *Current Opinion in Structural Biology* 41. Multi-protein assemblies in signaling • Catalysis and regulation, pp. 172 –179. ISSN: 0959-440X. DOI: https://doi.org/10.1016/j.sbi.2016.07.017. URL: http://www.sciencedirect.com/science/article/pii/S0959440X16300926.

Ampaw, Anna et al. (2017). "Observing enzyme ternary transition state analogue complexes by 19F NMR spectroscopy". In: *Chem. Sci.* 8 (12), pp. 8427–8434. DOI: 10.1039/C7SC04204C. URL: http://dx.doi.org/10.1039/C7SC04204C.

Amyes, Tina L., AnnMarie C. OĎonoghue, and John P. Richard (2001). "Contribution of Phosphate Intrinsic Binding Energy to the Enzymatic Rate Acceleration for Triosephosphate Isomerase". In: *Journal of the American Chemical Society* 123.45. PMID: 11697989, pp. 11325–11326. DOI: 10.1021/ja016754a. eprint: https://doi.org/10.1021/ja016754a. URL: https://doi.org/10.1021/ja016754a.

Amyes, Tina L. and John P. Richard (2013). "Specificity in Transition State Binding: The Pauling Model Revisited". In: *Biochemistry* 52.12. PMID: 23327224, pp. 2021–2035. DOI: 10.1021/bi301491r. eprint: https://doi.org/10.1021/bi301491r. URL: https://doi.org/10.1021/bi301491r.

Amyes, Tina L., John P. Richard, and James J. Tait (2005). "Activation of Orotidine 5'-Monophosphate Decarboxylase by Phosphite Dianion: The Whole Substrate is the Sum of Two Parts". In: *Journal of the American Chemical Society* 127.45. PMID: 16277505, pp. 15708–15709. DOI: 10.1021/ja055493s. eprint: https://doi.org/10.1021/ja055493s. URL: https://doi.org/10.1021/ja055493s.

Bae, Euiyoung and George N Phillips (2006). "Roles of static and dynamic domains in stability and catalysis of adenylate kinase." In: *Proceedings of the National Academy of Sciences*

of the United States of America 103.7, pp. 2132–2137. ISSN: 0027-8424. DOI: 10.1073/pnas.0507527103.

Bah, Alaji et al. (2014). "Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch". In: *Nature* 519, pp. 106–109. URL: http://dx.doi.org/10.1038/nature13999.

Barrozo, Alexandre et al. (2018). "Computer simulations of the catalytic mechanism of wild-type and mutant β-phosphoglucomutase". In: *Org. Biomol. Chem.* 16 (12), pp. 2060–2073. DOI: 10.1039/C8OB00312B. URL: http://dx.doi.org/10.1039/C8OB00312B.

Baxter, N J et al. (2010). "Atomic details of near-transition state conformers for enzyme phosphoryl transfer revealed by MgF_3^- rather than by phosphoranes". In: *Proceedings of the National Academy of Sciences (USA)* 107, pp. 4555–4560. DOI: 10.1073/pnas.0910333106. URL: http://dx.doi.org/10.1073/pnas.0910333106.

Baxter, Nicola J et al. (2006). "A Trojan horse transition state analogue generated by MgF_3^- formation in an enzyme active site". In: *Proceedings of the National Academy of Sciences (USA)* 103.40, pp. 14732–14737. DOI: 10.1073/pnas.0604448103.

Baxter, Nicola J. et al. (2008). "Anionic charge is prioritized over geometry in aluminum and magnesium fluoride transition state analogs of phosphoryl transfer enzymes". In: *Journal of the American Chemical Society* 130.12, pp. 3952–3958. ISSN: 00027863. DOI: 10.1021/ja078000n.

Baxter, Nicola J. et al. (2009). "MgF_3^- and α-Galactose 1-Phosphate in the Active Site of β-Phosphoglucomutase Form a Transition State Analogue of Phosphoryl Transfer". In: *Journal of the American Chemical Society* 131.45. PMID: 19852484, pp. 16334–16335. DOI: 10.1021/ja905972m. eprint: https://doi.org/10.1021/ja905972m. URL: https://doi.org/10.1021/ja905972m.

Benkovic, Stephen J. and Sharon Hammes-Schiffer (2003). "A Perspective on Enzyme Catalysis". In: *Science* 301.5637, pp. 1196–1202. DOI: 10.1126/science.1085515. eprint: http://www.sciencemag.org/content/301/5637/1196.full.pdf. URL: http://www.sciencemag.org/content/301/5637/1196.abstract.

Berg, Jeremy M., John L. Tymoczko, and Lubert Stryer (2012). *Biochemistry*. W.H. Freeman, pp. 1–1050. ISBN: 9781429276351.

Bieri, Michael, Edward J. d'Auvergne, and Paul R. Gooley (2011). "relaxGUI: a new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and μs motion of proteins". In: *Journal of Biomolecular NMR* 50.2, pp. 147–155. ISSN: 1573-5001. DOI: 10.1007/s10858-011-9509-1. URL: https://doi.org/10.1007/s10858-011-9509-1.

Blackburn, G. Michael et al. (2003). "Comment on "The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction"". In: *Science* 301.5637, pp. 1184–1184. ISSN: 0036–8075. DOI: 10.1126/science.1085796. eprint: http://science.sciencemag.org/content/301/5637/1184.3.full.pdf. URL: http://science.sciencemag.org/content/301/5637/1184.3.

Carr, H. Y. and E. M. Purcell (1954). "Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments". In: *Phys. Rev.* 94 (3), pp. 630–638. DOI: 10.1103/PhysRev.94.630. URL: https://link.aps.org/doi/10.1103/PhysRev.94.630.

Cavanagh, John et al. (2007). *Protein NMR Spectroscopy (Second Edition)*. Second Edition. Burlington: Academic Press, pp. 1 –885. ISBN: 978-0-12-164491-8. DOI: https://doi.org/10.1016/B978-012164491-8/50016-6. URL: http://www.sciencedirect.com/science/article/pii/B9780121644918500166.

Charlier, Cyril et al. (2018). "Study of protein folding under native conditions by rapidly switching the hydrostatic pressure inside an NMR sample cell". In: *Proceedings of the National Academy of Sciences*. ISSN: 0027-8424. DOI: 10.1073/pnas.1803642115. eprint: http://www.pnas.org/content/early/2018/04/10/1803642115.full.pdf. URL: http://www.pnas.org/content/early/2018/04/10/1803642115.

Chen, Vincent B. et al. (2010). "*MolProbity*: all-atom structure validation for macromolecular crystallography". In: *Acta Crystallographica Section D* 66.1, pp. 12–21. DOI: 10.1107/S0907444909042073. URL: https://doi.org/10.1107/S0907444909042073.

Cliff, Matthew J. et al. (2010). "Transition State Analogue Structures of Human Phosphoglycerate Kinase Establish the Importance of Charge Balance in Catalysis". In: *Journal of the American Chemical Society* 132.18, pp. 6507–6516. DOI: 10.1021/ja100974t. eprint: https://doi.org/10.1021/ja100974t. URL: https://doi.org/10.1021/ja100974t.

Clore, G. Marius et al. (1990). "Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins". In: *Journal of the American Chemical Society* 112.12, pp. 4989–4991. DOI: 10.1021/ja00168a070. eprint: https://doi.org/10.1021/ja00168a070. URL: https://doi.org/10.1021/ja00168a070.

Cornish-Bowden, Athel (2012). *Fundamentals of Enzyme Kinetics*. Butterworth-Heinemann, pp. 1–510. ISBN: 978-3-527-33074-4. URL: https://www.wiley.com/en-gb/Fundamentals+of+Enzyme+Kinetics%2C+4th+Edition-p-9783527330744.

Czisch, M. and R. Boelens (1998). "Sensitivity Enhancement in the TROSY Experiment". In: *Journal of Magnetic Resonance* 134.1, pp. 158 –160. ISSN: 1090-7807. DOI: https://doi.org/10.1006/jmre.1998.1483. URL: http://www.sciencedirect.com/science/article/pii/S1090780798914835.

Dai, Jianying et al. (2006). "Conformational cycling in β-phosphoglucomutase catalysis: Reorientation of the β-D-glucose 1,6-(bis)phosphate intermediate". In: *Biochemistry* 45.25, pp. 7818–7824. ISSN: 00062960. DOI: 10.1021/bi060136v.

Dai, Jianying et al. (2009). "Analysis of the Structural Determinants Underlying Discrimination between Substrate and Solvent in β-Phosphoglucomutase Catalysis". In: *Biochemistry* 48.9, pp. 1984–1995. DOI: 10.1021/bi801653r. eprint: https://doi.org/10.1021/bi801653r. URL: https://doi.org/10.1021/bi801653r.

Dalcin, Lisandro D. et al. (2011). "Parallel distributed computing using Python". In: *Advances in Water Resources* 34.9. New Computational Methods and Software Tools, pp. 1124 –1139. ISSN: 0309-1708. DOI: https://doi.org/10.1016/j.advwatres.2011.04.013. URL: http://www.sciencedirect.com/science/article/pii/S0309170811000777.

Dalcín, Lisandro, Rodrigo Paz, and Mario Storti (2005). "MPI for Python". In: *Journal of Parallel and Distributed Computing* 65.9, pp. 1108 –1115. ISSN: 0743-7315. DOI: https://doi.org/10.1016/j.jpdc.2005.03.010. URL: http://www.sciencedirect.com/science/article/pii/S0743731505000560.

Dalcín, Lisandro et al. (2008). "MPI for Python: Performance improvements and MPI-2 extensions". In: *Journal of Parallel and Distributed Computing* 68.5, pp. 655 –662. ISSN: 0743-7315. DOI: https://doi.org/10.1016/j.jpdc.2007.09.005. URL: http://www.sciencedirect.com/science/article/pii/S0743731507001712.

d'Auvergne, Edward J. (2006). "Protein dynamics: a study of the model-free analysis of NMR relaxation data". PhD thesis. University of Melbourne.

d'Auvergne, Edward J. and Paul R. Gooley (2003). "The use of model selection in the model-free analysis of protein dynamics". In: *Journal of Biomolecular NMR* 25.1, pp. 25–39. ISSN: 1573-5001. DOI: 10.1023/A:1021902006114. URL: https://doi.org/10.1023/A:1021902006114.

– (2006). "Model-free model elimination: A new step in the model-free dynamic analysis of NMR relaxation data". In: *Journal of Biomolecular NMR* 35.2, p. 117. ISSN: 1573-5001. DOI: 10.1007/s10858-006-9007-z. URL: https://doi.org/10.1007/s10858-006-9007-z.

– (2007). "Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces". In: *Journal of Biomolecular NMR* 40.2, p. 107. ISSN: 1573-5001. DOI: 10.1007/s10858-007-9214-2. URL: https://doi.org/10.1007/s10858-007-9214-2.

d'Auvergne, Edward J. and Paul R. Gooley (2007). "Set theory formulation of the model-free problem and the diffusion seeded model-free paradigm". In: *Mol. BioSyst.* 3 (7), pp. 483–494. DOI: 10.1039/B702202F. URL: http://dx.doi.org/10.1039/B702202F.

d'Auvergne, Edward J. and Paul R. Gooley (2008). "Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor". In: *Journal of Biomolecular NMR* 40.2, pp. 121–133. ISSN: 1573-5001. DOI: 10.1007/s10858-007-9213-3. URL: https://doi.org/10.1007/s10858-007-9213-3.

Delaglio, Frank et al. (1995). "NMRPipe: A multidimensional spectral processing system based on UNIX pipes". In: *Journal of Biomolecular NMR* 6.3, pp. 277–293. ISSN: 1573-5001. DOI: 10.1007/BF00197809. URL: https://doi.org/10.1007/BF00197809.

Dellus-Gur, Eynat et al. (2015). "Negative Epistasis and Evolvability in TEM-1 β-Lactamase - The Thin Line between an Enzyme's Conformational Freedom and Disorder". In: *Journal of Molecular Biology* 427.14, pp. 2396 –2409. ISSN: 0022-2836. DOI: https://doi.org/10.1016/j.jmb.2015.05.011. URL: http://www.sciencedirect.com/science/article/pii/S0022283615002958.

Desai, Bijoy J. et al. (2012). "Conformational Changes in Orotidine 5'-Monophosphate Decarboxylase: A Structure-Based Explanation for How the 5'-Phosphate Group Activates the Enzyme". In: *Biochemistry* 51.43. PMID: 23030629, pp. 8665–8678. DOI: 10.1021/bi301188k. eprint: https://doi.org/10.1021/bi301188k. URL: https://doi.org/10.1021/bi301188k.

Di Sabato, G. and W.P. Jencks (1961). "Mechanism and Catalysis of Reactions of Acyl Phosphates". In: *Journal of the American Chemical Society* 83, pp. 4400–4405.

Diehl, Carl et al. (2010). "Protein Flexibility and Conformational Entropy in Ligand Design Targeting the Carbohydrate Recognition Domain of Galectin-3". In: *Journal of the American Chemical Society* 132.41. PMID: 20873837, pp. 14577–14589. DOI: 10.1021/ja105852y. eprint: https://doi.org/10.1021/ja105852y. URL: https://doi.org/10.1021/ja105852y.

Dill, K A and H S Chan (1997). "From Levinthal to pathways to funnels." In: *Nature Structural Biology* 4.1, pp. 10–19.

Elsässer, Brigitta, Silvia Dohmeier-Fischer, and Gregor Fels (2012). "Theoretical investigation of the enzymatic phosphoryl transfer of β-phosphoglucomutase: revisiting both steps of the catalytic cycle". In: *Journal of Molecular Modeling* 18.7, pp. 3169–3179. ISSN: 0948-5023. DOI: 10.1007/s00894-011-1344-5. URL: https://doi.org/10.1007/s00894-011-1344-5.

Emsley, P. et al. (2010). "Features and development of *Coot*". In: *Acta Crystallographica Section D* 66.4, pp. 486–501. DOI: 10.1107/S0907444910007493. URL: https://doi.org/10.1107/S0907444910007493.

Favier, Adrien and Bernhard Brutscher (2011). "Recovering lost magnetization: polarization enhancement in biomolecular NMR". In: *Journal of Biomolecular NMR* 49.1, pp. 9–15. ISSN: 1573-5001. DOI: 10.1007/s10858-010-9461-5. URL: https://doi.org/10.1007/s10858-010-9461-5.

Ferrage, Fabien et al. (2010). "On the measurement of 15N–1H nuclear Overhauser effects. 2. Effects of the saturation scheme and water signal suppression". In: *Journal of Magnetic Resonance* 207.2, pp. 294 –303. ISSN: 1090-7807. DOI: https://doi.org/10.1016/j.jmr.2010.09.014. URL: http://www.sciencedirect.com/science/article/pii/S1090780710002934.

Fischer, Emil (1909). "Einfluß der Konfiguration auf die Wirkung der Enzyme. I". German. In: *Untersuchungen Über Kohlenhydrate und Fermente (1884–1908)*. Springer Berlin Heidelberg, pp. 836–844. ISBN: 978-3-642-98686-4. DOI: 10.1007/978-3-642-99501-9_101. URL: http://dx.doi.org/10.1007/978-3-642-99501-9_101.

Fushman, David, Rong Xu, and David Cowburn (1999). "Direct Determination of Changes of Interdomain Orientation on Ligation: Use of the Orientational Dependence of ^{15}N NMR Relaxation in Abl SH(32)". In: *Biochemistry* 38.32. PMID: 10441115, pp. 10225–10230. DOI: 10.1021/bi990897g. eprint: https://doi.org/10.1021/bi990897g. URL: https://doi.org/10.1021/bi990897g.

Fushman, David et al. (2004). "Determining domain orientation in macromolecules by using spin-relaxation and residual dipolar coupling measurements". In: *Progress in Nuclear Magnetic Resonance Spectroscopy* 44.3, pp. 189 –214. ISSN: 0079-6565. DOI: https://doi.org/10.1016/j.pnmrs.2004.02.001. URL: http://www.sciencedirect.com/science/article/pii/S0079656504000032.

Go, Maybelle K., Tina L. Amyes, and John P. Richard (2010). "Rescue of K12G Triosephosphate Isomerase by Ammonium Cations: The Reaction of an Enzyme in Pieces". In: *Journal of the American Chemical Society* 132.38. PMID: 20822141, pp. 13525–13532. DOI: 10.1021/ja106104h. eprint: https://doi.org/10.1021/ja106104h. URL: https://doi.org/10.1021/ja106104h.

Goldman, Maurice (2001). "Formal Theory of Spin–Lattice Relaxation". In: *Journal of Magnetic Resonance* 149.2, pp. 160 –187. ISSN: 1090-7807. DOI: https://doi.org/10.1006/jmre.2000.2239. URL: http://www.sciencedirect.com/science/article/pii/S1090780700922390.

Golicnik, Marko et al. (2009). "Kinetic Analysis of β-Phosphoglucomutase and Its Inhibition by Magnesium Fluoride." In: *Journal of the American Chemical Society* 131.4, pp. 1575–88. ISSN: 1520-5126. DOI: 10.1021/ja806421f. URL: http://dx.doi.org/10.1021/ja806421f.

Gong, C. X. et al. (2006). "Dysregulation of protein phosphorylation/dephosphorylation in Alzheimer's disease: a therapeutic target". In: *J. Biomed. Biotechnol.* 2006.3, p. 31825.

Graauw, Marjo de, Paul Hensbergen, and Bob van de Water (2006). "Phospho-proteomic analysis of cellular signaling". In: *ELECTROPHORESIS* 27.13, pp. 2676–2686. ISSN: 1522-2683. DOI: 10.1002/elps.200600018. URL: http://dx.doi.org/10.1002/elps.200600018.

Graham, Debbie L et al. (2002). "MgF_3^- as a Transition State Analog of Phosphoryl Transfer". In: *Chemistry Biology* 9.3, pp. 375 –381. ISSN: 1074-5521. DOI: https://doi.org/10.1016/S1074-5521(02)00112-6. URL: http://www.sciencedirect.com/science/article/pii/S1074552102001126.

Griffin, J. L. et al. (2012). "Near attack conformers dominate β-phosphoglucomutase complexes where geometry and charge distribution reflect those of substrate". In: *Proceedings of the National Academy of Sciences (USA)* 109.18, pp. 6910–6915. ISSN: 0027-8424. DOI: 10.1073/pnas.1116855109.

Griffin, Joanna L (2011). "Investigations of the Metal Flouride Transition state and Ground State Analogue Complexes of HAD superfamily Proteins by Nuclear Magnetic Resonance Spectroscopy". PhD thesis. University of Sheffield.

Halle, Bertil (2009). "The physical basis of model-free analysis of NMR relaxation data from proteins and complex fluids". In: *The Journal of Chemical Physics* 131.22, p. 224507. DOI: 10.1063/1.3269991. eprint: https://doi.org/10.1063/1.3269991. URL: https://doi.org/10.1063/1.3269991.

Hay, Sam and Nigel S Scrutton (Mar. 2012). "Good vibrations in enzyme-catalysed reactions." In: *Nature chemistry* 4.3, pp. 161–8. ISSN: 1755-4349. DOI: 10.1038/nchem.1223. URL: http://dx.doi.org/10.1038/nchem.1223.

Hayward, Steven and Herman J.C. Berendsen (1998). "Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T4 lysozyme". In: *Proteins: Structure, Function, and Bioinformatics* 30.2, pp. 144–154. ISSN: 1097-0134. DOI: 10.1002/(SICI)1097-0134(19980201)30:2<144::AID-PROT4>3.0.CO;2-N. URL: http://dx.doi.org/10.1002/(SICI)1097-0134(19980201)30:2<144::AID-PROT4>3.0.CO;2-N.

Hilvert, D. (2000). "Critical analysis of antibody catalysis". In: *Annu Rev Biochem* 69, pp. 751–793. DOI: 10.1146/annurev.biochem.69.1.751. URL: http://dx.doi.org/10.1146/annurev.biochem.69.1.751.

Huang, Hua et al. (2015). "Panoramic view of a superfamily of phosphatases through substrate profiling". In: *Proceedings of the National Academy of Sciences* 112.16, E1974–E1983. ISSN: 0027-8424. DOI: 10.1073/pnas.1423570112. eprint: http://www.pnas.org/content/112/16/E1974.full.pdf. URL: http://www.pnas.org/content/112/16/E1974.

Hyberts, Sven G., Scott A. Robson, and Gerhard Wagner (2013). "Exploring signal-to-noise ratio and sensitivity in non-uniformly sampled multi-dimensional NMR spectra". In: *Journal of Biomolecular NMR* 55.2, pp. 167–178. ISSN: 1573-5001. DOI: 10.1007/s10858-012-9698-2. URL: https://doi.org/10.1007/s10858-012-9698-2.

Hyberts, Sven G. et al. (2012). "Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling". In: *Journal of Biomolecular NMR* 52.4, pp. 315–327. ISSN: 1573-5001. DOI: 10.1007/s10858-012-9611-z. URL: https://doi.org/10.1007/s10858-012-9611-z.

Jencks, William P. (1969). *Catalysis in Chemistry and Enzymology*. McGraw-Hill. New York.

Jin, Yi, Robert W. Molt, and G. Michael Blackburn (2017a). "Metal Fluorides: Tools for Structural and Computational Analysis of Phosphoryl Transfer Enzymes". In: *Topics in Current Chemistry* 375.2, p. 36. ISSN: 2364-8961. DOI: 10.1007/s41061-017-0130-y. URL: https://doi.org/10.1007/s41061-017-0130-y.

– (2017b). "Metal Fluorides: Tools for Structural and Computational Analysis of Phosphoryl Transfer Enzymes". In: *Topics in Current Chemistry* 375.2, pp. 1–31. ISSN: 03401022. DOI: 10.1007/s41061-017-0130-y.

Jin, Yi et al. (2014). "α-Fluorophosphonates reveal how a phosphomutase conserves transition state conformation over hexose recognition in its two-step reaction". In: *Proceedings of the National Academy of Sciences (USA)* 111.34, pp. 12384–12389. ISSN: 0027-8424. DOI: 10.1073/pnas.1402850111. eprint: http://www.pnas.org/content/111/34/12384.full.pdf. URL: http://www.pnas.org/content/111/34/12384.

Jin, Yi et al. (2016). "^{19}F NMR and DFT Analysis Reveal Structural and Electronic Transition State Features for RhoA-Catalyzed GTP Hydrolysis". In: *Angewandte Chemie International Edition* 55.10, pp. 3318–3322. DOI: 10.1002/anie.201509477. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.201509477. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201509477.

Jin, Yi et al. (2017a). "Assessing the Influence of Mutation on GTPase Transition States by Using X-ray Crystallography, 19F NMR, and DFT Approaches". In: *Angewandte Chemie International Edition* 56.33, pp. 9732–9735. DOI: 10.1002/anie.201703074. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.201703074. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201703074.

Jin, Yi et al. (2017b). "Metal Fluorides as Analogues for Studies on Phosphoryl Transfer Enzymes". In: *Angewandte Chemie - International Edition* 56.15, pp. 4110–4128. ISSN: 15213773. DOI: 10.1002/anie.201606474.

Johnson, Luke A. et al. (2018). "van der Waals Contact between Nucleophile and Transferring Phosphorus Is Insufficient To Achieve Enzyme Transition-State Architecture". In: *ACS Catalysis* 8.9, pp. 8140–8153. DOI: 10.1021/acscatal.8b01612. eprint: https://doi.org/10.1021/acscatal.8b01612. URL: https://doi.org/10.1021/acscatal.8b01612.

Johnson, Luke J (2015). "Structural and mechanistic investigation of enzyme-catalysed phosphoryl transfer in two HAD superfamily proteins". PhD thesis. University of Sheffield.

Jones, J.A. et al. (1996). "Optimal Sampling Strategies for the Measurement of Spin–Spin Relaxation Times". In: *Journal of Magnetic Resonance, Series B* 113.1, pp. 25 –34. ISSN: 1064-1866. DOI: https://doi.org/10.1006/jmrb.1996.0151. URL: http://www.sciencedirect.com/science/article/pii/S106418669690151X.

Kabsch, Wolfgang (2010). "*XDS*". In: *Acta Crystallographica Section D* 66.2, pp. 125–132. DOI: 10.1107/S0907444909047337. URL: http://dx.doi.org/10.1107/S0907444909047337.

Kamerlin, Shina C. L. et al. (Feb. 2013). "Why nature really chose phosphate". In: *Quarterly Reviews of Biophysics* 46 (01), pp. 1–132. ISSN: 1469-8994. DOI: 10.1017/S0033583512000157. URL: http://journals.cambridge.org/article_S0033583512000157.

Kay, Lewis, Paul Keifer, and Tim Saarinen (1992). "Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity". In: *Journal of the American Chemical Society* 114.26, pp. 10663–10665. DOI: 10.1021/ja00052a088. eprint: https://doi.org/10.1021/ja00052a088. URL: https://doi.org/10.1021/ja00052a088.

Keeler, James (2011). *Understanding NMR Spectroscopy*. Wiley. ISBN: 9781119964933. URL: https://www.wiley.com/en-gb/Understanding+NMR+Spectroscopy,+2nd+Edition-p-9780470746097.

Knowles, Jeremy (2003). "Seeing Is Believing". In: *Science* 299.5615, pp. 2002–2003. ISSN: 0036-8075. DOI: 10.1126/science.1084036. eprint: http://science.sciencemag.org/

content/299/5615/2002.full.pdf. URL: http://science.sciencemag.org/content/299/5615/2002.

Lad, C, N H Williams, and R Wolfenden (2003). "The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases". In: *Proc Natl Acad Sci U S A* 100, pp. 5607–5610.

Lahiri, Sushmita D. et al. (2002a). "Caught in the Act: The Structure of Phosphorylated β-Phosphoglucomutase from *Lactococcus lactis*," in: *Biochemistry* 41.26. PMID: 12081483, pp. 8351–8359. DOI: 10.1021/bi0202373. eprint: https://doi.org/10.1021/bi0202373. URL: https://doi.org/10.1021/bi0202373.

Lahiri, Sushmita D. et al. (2002b). "Crystallization and preliminary X-ray diffraction studies of β-phosphoglucomutase from *Lactococcus lactus*". In: *Acta Crystallographica Section D* 58.2, pp. 324–326. DOI: 10.1107/S0907444901019989. URL: https://doi.org/10.1107/S0907444901019989.

Lahiri, Sushmita D. et al. (2003). "The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction". In: *Science* 299.5615, pp. 2067–2071. ISSN: 0036-8075. DOI: 10.1126/science.1082710. eprint: http://science.sciencemag.org/content/299/5615/2067.full.pdf. URL: http://science.sciencemag.org/content/299/5615/2067.

Lahiri, Sushmita D. et al. (2004). "Analysis of the Substrate Specificity Loop of the HAD Superfamily Cap Domain". In: *Biochemistry* 43.10, pp. 2812–2820. ISSN: 00062960. DOI: 10.1021/bi0356810.

Lakomek, Nils-Alexander, Jinfa Ying, and Ad Bax (2012). "Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods". In: *Journal of Biomolecular NMR* 53.3, pp. 209–221. ISSN: 1573-5001. DOI: 10.1007/s10858-012-9626-5. URL: https://doi.org/10.1007/s10858-012-9626-5.

Lassila, Jonathan K., Jesse G. Zalatan, and Daniel Herschlag (2011). "Biological Phosphoryl-Transfer Reactions: Understanding Mechanism and Catalysis". In: *Annual Review of Biochemistry* 80.1. PMID: 21513457, pp. 669–702. DOI: 10.1146/annurev-biochem-060409-092741. eprint: http://dx.doi.org/10.1146/annurev-biochem-060409-092741. URL: http://dx.doi.org/10.1146/annurev-biochem-060409-092741.

Levander, F, U Andersson, and P Rådström (2001). "Physiological role of beta-phosphoglucomutase in *Lactococcus lactis*." In: *Applied and environmental microbiology* 67.10, pp. 4546–4553. ISSN: 0099–2240 (Print). DOI: 10.1128/AEM.67.10.4546. eprint: http://aem.asm.org/content/67/10/4546.full.pdf+html. URL: http://aem.asm.org/content/67/10/4546.abstract.

Lipari, Giovanni and Attila Szabo (1982a). "Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity". In: *Journal of the American Chemical Society* 104.17, pp. 4546–4559. DOI: 10.1021/ja00381a009. eprint: https://doi.org/10.1021/ja00381a009. URL: https://doi.org/10.1021/ja00381a009.

– (1982b). "Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results". In: *Journal of the American Chemical Society* 104.17, pp. 4559–4570. DOI: 10.1021/ja00381a010. eprint: https://doi.org/10.1021/ja00381a010. URL: https://doi.org/10.1021/ja00381a010.

Lodish, H. et al. (2007). *Molecular Cell Biology / Iclicker*. Macmillan Higher Education. ISBN: 9781429220521. URL: https://books.google.co.uk/books?id=0V8VPQAACAAJ.

Manning, G. et al. (2002). "The Protein Kinase Complement of the Human Genome". In: *Science* 298.5600, pp. 1912–1934. ISSN: 0036-8075. DOI: 10.1126/science.1075762. eprint: http://science.sciencemag.org/content/298/5600/1912.full.pdf. URL: http://science.sciencemag.org/content/298/5600/1912.

Marcos, Enrique, Martin J. Field, and Ramon Crehuet (2010). "Pentacoordinated phosphorus revisited by high-level QM/MM calculations". In: *Proteins: Structure, Function, and Bioinformatics* 78.11, pp. 2405–2411. DOI: 10.1002/prot.22758. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/prot.22758. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.22758.

Meiboom, S. and D. Gill (1958). "Modified Spin-Echo Method for Measuring Nuclear Relaxation Times". In: *Review of Scientific Instruments* 29.8, pp. 688–691. DOI: 10.1063/1.1716296. eprint: https://doi.org/10.1063/1.1716296. URL: https://doi.org/10.1063/1.1716296.

Meissner, Axel et al. (1998). "Double spin-state-selective coherence transfer. Application for two-dimensional selection of multiplet components with long transverse relaxation times". In: *Molecular Physics* 95.6, pp. 1137–1142. DOI: 10.1080/00268979809483245. eprint: https://doi.org/10.1080/00268979809483245. URL: https://doi.org/10.1080/00268979809483245.

Millet, Oscar et al. (2002). "Deuterium Spin Probes of Side-Chain Dynamics in Proteins. 1. Measurement of Five Relaxation Rates per Deuteron in ^{13}C-Labeled and Fractionally ^{2}H-Enriched Proteins in Solution". In: *Journal of the American Chemical Society* 124.22. PMID: 12033875, pp. 6439–6448. DOI: 10.1021/ja012497y. eprint: https://doi.org/10.1021/ja012497y. URL: https://doi.org/10.1021/ja012497y.

Morin, Sébastien and Stéphane M. Gagné (2009). "Simple tests for the validation of multiple field spin relaxation data". In: *Journal of Biomolecular NMR* 45.4, p. 361. ISSN: 1573-5001. DOI: 10.1007/s10858-009-9381-4. URL: https://doi.org/10.1007/s10858-009-9381-4.

Muhandiram, D. R. et al. (1995). "Measurement of 2H T1 and T1ρ Relaxation Times in Uniformly ^{13}C-Labeled and Fractionally ^{2}H-Labeled Proteins in Solution". In: *Journal of the American Chemical Society* 117.46, pp. 11536–11544. DOI: 10.1021/ja00151a018. eprint: https://doi.org/10.1021/ja00151a018. URL: https://doi.org/10.1021/ja00151a018.

Mulder, Frans A.A. et al. (1998). "An Off-resonance Rotating Frame Relaxation Experiment for the Investigation of Macromolecular Dynamics Using Adiabatic Rotations". In: *Journal of Magnetic Resonance* 131.2, pp. 351 –357. ISSN: 1090-7807. DOI: https://doi.org/10.1006/jmre.1998.1380. URL: http://www.sciencedirect.com/science/article/pii/S1090780798913805.

Murshudov, G. N., A. A. Vagin, and E. J. Dodson (1997). "Refinement of Macromolecular Structures by the Maximum-Likelihood Method". In: *Acta Crystallographica Section D* 53.3, pp. 240–255. DOI: 10.1107/S0907444996012255. URL: https://doi.org/10.1107/S0907444996012255.

Niklasson, Markus et al. (2017). "Comprehensive analysis of NMR data using advanced line shape fitting". In: *Journal of Biomolecular NMR* 69.2, pp. 93–99. ISSN: 1573-5001. DOI: 10.1007/s10858-017-0141-6. URL: https://doi.org/10.1007/s10858-017-0141-6.

Okazaki, Kei-Ichi and Shoji Takada (2008). "Dynamic energy landscape view of coupled binding and protein conformational change: induced-fit versus population-shift mechanisms." In: *Proceedings of the National Academy of Sciences of the United States of America* 105.32, pp. 11182–11187. ISSN: 0027-8424. DOI: 10.1073/pnas.0802524105.

Pabis, Anna, Fernanda Duarte, and Shina C. L. Kamerlin (2016). "Promiscuity in the Enzymatic Catalysis of Phosphate and Sulfate Transfer". In: *Biochemistry* 55.22. PMID: 27187273, pp. 3061–3081. DOI: 10.1021/acs.biochem.6b00297. eprint: https://doi.org/10.1021/acs.biochem.6b00297. URL: https://doi.org/10.1021/acs.biochem.6b00297.

Pauling, Linus (1948). "Nature of Forces between Large Molecules of Biological Interest". In: *Nature* 161, pp. 707–709.

Pervushin, Konstantin V., Gerhard Wider, and Kurt Wüthrich (1998). "Single Transition-to-single Transition Polarization Transfer (ST2-PT) in [15N,1H]-TROSY". In: *Journal of Biomolecular NMR* 12.2, pp. 345–348. ISSN: 1573-5001. DOI: 10.1023/A:1008268930690. URL: https://doi.org/10.1023/A:1008268930690.

Petrović, Dušan et al. (2018). "Conformational dynamics and enzyme evolution". In: *Journal of The Royal Society Interface* 15.144. ISSN: 1742-5689. DOI: 10.1098/rsif.2018.0330. eprint: http://rsif.royalsocietypublishing.org/content/15/144/20180330.full.pdf. URL: http://rsif.royalsocietypublishing.org/content/15/144/20180330.

Pisliakov, Andrei V et al. (Oct. 2009). "Enzyme millisecond conformational dynamics do not catalyze the chemical step." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.41, pp. 17359–64. ISSN: 1091-6490. DOI: 10.1073/pnas.0909150106. URL: http://www.pnas.org/content/106/41/17359.short.

Qian, Ny et al. (1994). "Purification and characterization of two phosphoglucomutases from *Lactococcus lactis* subsp. lactis and their regulation in maltose- and glucose-utilizing cells". In: *Journal of Bacteriology* 176.17, pp. 5304–5311. ISSN: 00219193. DOI: 10.1128/jb.176.17.5304-5311.1994. eprint: http://jb.asm.org/content/176/17/5304.full.pdf+html. URL: http://jb.asm.org/content/176/17/5304.abstract.

Qian, Ny et al. (1997). "Product formation and phosphoglucomutase activities in *Lactococcus lactis*: cloning and characterization of a novel phosphoglucomutase gene". In: *Microbiology* 143.3, pp. 855–865. URL: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-143-3-855.

Rance, Mark, J.Patrick Loria, and Arthur G. Palmer (1999). "Sensitivity Improvement of Transverse Relaxation-Optimized Spectroscopy". In: *Journal of Magnetic Resonance* 136.1, pp. 92 –101. ISSN: 1090-7807. DOI: https://doi.org/10.1006/jmre.1998.1626. URL: http://www.sciencedirect.com/science/article/pii/S1090780798916263.

Read, R. J. and A. J. Schierbeek (1988). "A phased translation function". In: *Journal of Applied Crystallography* 21.5, pp. 490–495. DOI: 10.1107/S002188988800562X. URL: https://doi.org/10.1107/S002188988800562X.

Reed, Michelle A.C. et al. (2003). "Effects of Domain Dissection on the Folding and Stability of the 43 kDa Protein PGK Probed by NMR". In: *Journal of Molecular Biology* 330.5, pp. 1189 –1201. ISSN: 0022-2836. DOI: https://doi.org/10.1016/S0022-2836(03)00625-9. URL: http://www.sciencedirect.com/science/article/pii/S0022283603006259.

Reyes, Archie C., Tina L. Amyes, and John P. Richard (2016). "Enzyme Architecture: Self-Assembly of Enzyme and Substrate Pieces of Glycerol-3-Phosphate Dehydrogenase into a Robust Catalyst of Hydride Transfer". In: *Journal of the American Chemical Society* 138.46. PMID: 27792325, pp. 15251–15259. DOI: 10.1021/jacs.6b09936. eprint: https://doi.org/10.1021/jacs.6b09936. URL: https://doi.org/10.1021/jacs.6b09936.

Reyes, Archie C. et al. (2015). "Enzyme Architecture: Optimization of Transition State Stabilization from a Cation–Phosphodianion Pair". In: *Journal of the American Chemical Society*

137.16. PMID: 25884759, pp. 5312–5315. DOI: 10.1021/jacs.5b02202. eprint: https://doi.org/10.1021/jacs.5b02202. URL: https://doi.org/10.1021/jacs.5b02202.

Schnell, Jason R., H. Jane Dyson, and Peter E. Wright (2004). "Structure, Dynamics, and Catalytic Function of Dihydrofolate Reductase". In: *Annual Review of Biophysics and Biomolecular Structure* 33.1. PMID: 15139807, pp. 119–140. DOI: 10.1146/annurev.biophys.33.110502.133613. eprint: https://doi.org/10.1146/annurev.biophys.33.110502.133613. URL: https://doi.org/10.1146/annurev.biophys.33.110502.133613.

Schulte-Herbrüggen, Thomas and Ole Winneche Sørensen (2000). "Clean TROSY: Compensation for Relaxation-Induced Artifacts". In: *Journal of Magnetic Resonance* 144.1, pp. 123–128. ISSN: 1090-7807. DOI: https://doi.org/10.1006/jmre.2000.2020. URL: http://www.sciencedirect.com/science/article/pii/S1090780700920202.

Sheldrick, George M. and Thomas R. Schneider (1997). "SHELXL: High-resolution refinement". In: *Macromolecular Crystallography Part B*. Vol. 277. Methods in Enzymology. Academic Press, pp. 319–343. DOI: https://doi.org/10.1016/S0076-6879(97)77018-6. URL: http://www.sciencedirect.com/science/article/pii/S0076687997770186.

Skrynnikov, Nikolai R., Oscar Millet, and Lewis E. Kay (2002). "Deuterium Spin Probes of Side-Chain Dynamics in Proteins. 2. Spectral Density Mapping and Identification of Nanosecond Time-Scale Side-Chain Motions". In: *Journal of the American Chemical Society* 124.22. PMID: 12033876, pp. 6449–6460. DOI: 10.1021/ja012498q. eprint: https://doi.org/10.1021/ja012498q. URL: https://doi.org/10.1021/ja012498q.

Sternweis, P C and A G Gilman (1982). "Aluminum: a requirement for activation of the regulatory component of adenylate cyclase by fluoride". In: *Proceedings of the National Academy of Sciences* 79.16, pp. 4888–4891. ISSN: 0027-8424. DOI: 10.1073/pnas.79.16.4888. eprint: http://www.pnas.org/content/79/16/4888.full.pdf. URL: http://www.pnas.org/content/79/16/4888.

Sugase, Kenji et al. (2007). "Tailoring Relaxation Dispersion Experiments for Fast-Associating Protein Complexes". In: *Journal of the American Chemical Society* 129.44. PMID: 17935336, pp. 13406–13407. DOI: 10.1021/ja0762238. eprint: https://doi.org/10.1021/ja0762238. URL: https://doi.org/10.1021/ja0762238.

Todd, Sir Alexander (1959). "SOME ASPECTS OF PHOSPHATE CHEMISTRY". In: *Proceedings of the National Academy of Sciences* 45.9, pp. 1389–1397. eprint: http://www.pnas.org/content/45/9/1389.full.pdf. URL: http://www.pnas.org/content/45/9/1389.short.

Tokuriki, Nobuhiko and Dan S Tawfik (2009). "Stability effects of mutations and protein evolvability". In: *Current Opinion in Structural Biology* 19.5. Carbohydradtes and glycoconjugates / Biophysical methods, pp. 596–604. ISSN: 0959-440X. DOI: https://doi.org/10.1016/j.sbi.2009.08.003. URL: http://www.sciencedirect.com/science/article/pii/S0959440X09001249.

Tramontano, A, KD Janda, and RA Lerner (1986). "Catalytic antibodies". In: *Science* 234.4783, pp. 1566–1570. DOI: 10.1126/science.3787261. eprint: http://www.sciencemag.org/content/234/4783/1566.full.pdf. URL: http://www.sciencemag.org/content/234/4783/1566.abstract.

Tremblay, Lee W. et al. (2005). "Chemical Confirmation of a Pentavalent Phosphorane in Complex with β-Phosphoglucomutase". In: *Journal of the American Chemical Society* 127.15. PMID: 15826149, pp. 5298–5299. DOI: 10.1021/ja0509073. eprint: https://doi.org/10.1021/ja0509073. URL: https://doi.org/10.1021/ja0509073.

Tsang, Wing-Yin, Tina L. Amyes, and John P. Richard (2008). "A Substrate in Pieces: Allosteric Activation of Glycerol 3-Phosphate Dehydrogenase (NAD+) by Phosphite Dianion". In: *Biochemistry* 47.16. PMID: 18376850, pp. 4575–4582. DOI: 10.1021/bi8001743. eprint: https://doi.org/10.1021/bi8001743. URL: https://doi.org/10.1021/bi8001743.

Vagin, A. and A. Teplyakov (1997). "*MOLREP*: an Automated Program for Molecular Replacement". In: *Journal of Applied Crystallography* 30.6, pp. 1022–1025. DOI: 10.1107/S0021889897006766. URL: https://doi.org/10.1107/S0021889897006766.

Varadan, Ranjani et al. (2002). "Structural Properties of Polyubiquitin Chains in Solution". In: *Journal of Molecular Biology* 324.4, pp. 637 –647. ISSN: 0022-2836. DOI: https://doi.org/10.1016/S0022-2836(02)01198-1. URL: http://www.sciencedirect.com/science/article/pii/S0022283602011981.

Voet, D. and J.G. Voet (2010). *Biochemistry, 4th Edition*. John Wiley & Sons. ISBN: 9781118139936. URL: https://books.google.co.uk/books?id=ne0bAAAAQBAJ.

Walker, Olivier, Ranjani Varadan, and David Fushman (2004). "Efficient and accurate determination of the overall rotational diffusion tensor of a molecule from ^{15}N relaxation data using computer program ROTDIF". In: *Journal of Magnetic Resonance* 168.2, pp. 336 –345. ISSN: 1090-7807. DOI: https://doi.org/10.1016/j.jmr.2004.03.019. URL: http://www.sciencedirect.com/science/article/pii/S109078070400076X.

Wang, Weiru et al. (2001). "Crystal Structure of Phosphoserine Phosphatase from *Methanococcus jannaschii*, a Hyperthermophile, at 1.8 Å Resolution". In: *Structure* 9.1, pp. 65 –71. ISSN: 0969-2126. DOI: https://doi.org/10.1016/S0969-2126(00)00558-X. URL: http://www.sciencedirect.com/science/article/pii/S096921260000558X.

Wang, Weiru et al. (2002). "Structural Characterization of the Reaction Pathway in Phosphoserine Phosphatase: Crystallographic "snapshots" of Intermediate States". In: *Journal of Molecular Biology* 319.2, pp. 421 –431. ISSN: 0022-2836. DOI: https://doi.org/10.1016/S0022-2836(02)00324-8. URL: http://www.sciencedirect.com/science/article/pii/S0022283602003248.

Webster, Charles Edwin (2004). "High-Energy Intermediate or Stable Transition State Analogue: Theoretical Perspective of the Active Site and Mechanism of β– Phosphoglucomutase". In: *Journal of the American Chemical Society* 126.22. PMID: 15174833, pp. 6840–6841. DOI: 10.1021/ja049232e. eprint: https://doi.org/10.1021/ja049232e. URL: https://doi.org/10.1021/ja049232e.

Weigelt, Johan (1998). "Single Scan, Sensitivity- and Gradient-Enhanced TROSY for Multidimensional NMR Experiments". In: *Journal of the American Chemical Society* 120.41, pp. 10778–10779. DOI: 10.1021/ja982649y. eprint: https://doi.org/10.1021/ja982649y. URL: https://doi.org/10.1021/ja982649y.

Westheimer, FH (1987). "Why nature chose phosphates". In: *Science* 235.4793, pp. 1173–1178. ISSN: 0036-8075. DOI: 10.1126/science.2434996. eprint: http://science.sciencemag.org/content/235/4793/1173.full.pdf. URL: http://science.sciencemag.org/content/235/4793/1173.

Williamson, M. (2012). *How Proteins Work*. CRC Press. ISBN: 9781136665493. URL: https://books.google.co.uk/books?id=TSsWBAAAQBAJ.

Wilson, C. et al. (2015). "Using ancient protein kinases to unravel a modern cancer drug's mechanism". In: *Science* 347.6224, pp. 882–886. ISSN: 0036-8075. DOI: 10.1126/science.

aaa1823. eprint: http://science.sciencemag.org/content/347/6224/882.full.pdf. URL: http://science.sciencemag.org/content/347/6224/882.

Winn, Martyn D. et al. (2011). "Overview of the *CCP*4 suite and current developments". In: *Acta Crystallographica Section D* 67.4, pp. 235–242. DOI: 10.1107/S0907444910045749. URL: https://doi.org/10.1107/S0907444910045749.

Winter, G. (2010). "*xia2*: an expert system for macromolecular crystallography data reduction". In: *Journal of Applied Crystallography* 43.1, pp. 186–190. DOI: 10.1107/S0021889809045701. URL: https://doi.org/10.1107/S0021889809045701.

Wittinghofer, Alfred (1997). "Signaling mechanistics: Aluminum fluoride for molecule of the year". In: *Current Biology* 7.11, R682 –R685. ISSN: 0960-9822. DOI: https://doi.org/10.1016/S0960-9822(06)00355-1. URL: http://www.sciencedirect.com/science/article/pii/S0960982206003551.

Zhai, Xiang, Tina L. Amyes, and John P. Richard (2014). "Enzyme Architecture: Remarkably Similar Transition States for Triosephosphate Isomerase-Catalyzed Reactions of the Whole Substrate and the Substrate in Pieces". In: *Journal of the American Chemical Society* 136.11. PMID: 24588650, pp. 4145–4148. DOI: 10.1021/ja501103b. eprint: https://doi.org/10.1021/ja501103b. URL: https://doi.org/10.1021/ja501103b.

Zhang, Guofeng et al. (2005). "Catalytic cycling in β-phosphoglucomutase: A kinetic and structural analysis". In: *Biochemistry* 44.27, pp. 9404–9416. ISSN: 00062960. DOI: 10.1021/bi050558p.

Zhang, Hui et al. (2002). "Phosphoprotein Analysis Using Antibodies Broadly Reactive against Phosphorylated Motifs". In: *Journal of Biological Chemistry* 277.42, pp. 39379–39387. DOI: 10.1074/jbc.M206399200. eprint: http://www.jbc.org/content/277/42/39379.full.pdf+html. URL: http://www.jbc.org/content/277/42/39379.abstract.

Zhu, Guang, Xiang Ming Kong, and Kong Hung Sze (1999). "Gradient and sensitivity enhancement of 2D TROSY with water flip-back, 3D NOESY-TROSY and TOCSY-TROSY experiments". In: *Journal of Biomolecular NMR* 13.1, pp. 77–81. ISSN: 1573-5001. DOI: 10.1023/A:1008398227519. URL: https://doi.org/10.1023/A:1008398227519.

Appendix A

Papers

A.1 Paper I: van der Waals contact between nucleophile and transferring phosphorus is insufficient to achieve enzyme transition state architecture.

Contribution: Paper I: I expressed and purified protein, I performed most of the crystallography and some of the NMR, I analysed and interpreted the data and designed further experiments, I wrote processing scripts for analysis of kinetic data, I took part in writing of the manuscript alongside NJB, CRT, and JPW.

ACS Catalysis

Research Article

Cite This: ACS Catal. 2018, 8, 8140−8153

pubs.acs.org/acscatalysis

van der Waals Contact between Nucleophile and Transferring Phosphorus Is Insufficient To Achieve Enzyme Transition-State Architecture

Luke A. Johnson,[†,¶,#] Angus J. Robertson,[†,#] Nicola J. Baxter,[†,‡] Clare R. Trevitt,[†] Claudine Bisson,[†,§] Yi Jin,[†,¶] Henry P. Wood,[†] Andrea M. Hounslow,[†] Matthew J. Cliff,[‡] G. Michael Blackburn,[†] Matthew W. Bowler,[∥] and Jonathan P. Waltho*[,†,‡]

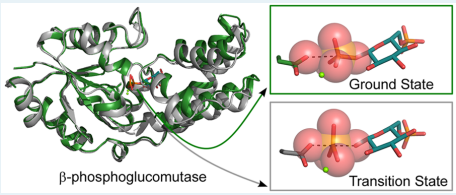
[†]Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield S10 2TN, United Kingdom

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

[∥]European Molecular Biology Laboratory, Grenoble Outstation, 71 Avenue des Martyrs, CS 90181, F-38042 Grenoble, France

Supporting Information

ABSTRACT: Phosphate plays a crucial role in biology because of the stability of the phosphate ester bond. To overcome this inherent stability, enzymes that catalyze phosphoryl transfer reactions achieve enormous rate accelerations to operate on biologically relevant time scales, and the mechanisms that underpin catalysis have been the subject of extensive debate. In an archetypal system, β-phosphoglucomutase catalyzes the reversible isomerization of β-glucose 1-phosphate and glucose 6-phosphate via two phosphoryl transfer steps using a β-glucose 1,6-bisphosphate intermediate and a catalytic Mg^{II} ion. In the present work, a variant of β-phosphoglucomutase, where the aspartate residue that acts as a general acid−base is replaced with asparagine, traps highly stable complexes containing the β-glucose 1,6-bisphosphate intermediate in the active site. Crystal structures of these complexes show that, when the enzyme is unable to transfer a proton, the intermediate is arrested in catalysis at an initial stage of phosphoryl transfer. The nucleophilic oxygen and transferring phosphorus atoms are aligned and in van der Waals contact, yet the enzyme is less closed than in transition-state (analogue) complexes, and binding of the catalytic Mg^{II} ion is compromised. Together, these observations indicate that optimal closure and optimal Mg^{II} binding occur only at higher energy positions on the reaction trajectory, allowing the enzyme to balance efficient catalysis with product dissociation. It is also confirmed that the general acid−base ensures that mutase activity is ∼10^3 fold greater than phosphatase activity in β-phosphoglucomutase.



KEYWORDS: *phosphoryl transfer enzyme, general acid−base catalysis, near attack conformation, magnesium ion affinity, X-ray crystallography*

INTRODUCTION

The efficiency of phosphoryl transfer enzymes in overcoming the stability of phosphate mono- and diesters under physiological conditions has enabled biology to perform a vast array of functions, spanning transient cell signaling cascades, energy storage and consumption, protein regulation, and the manipulation of genetic material.[1] Phosphoryl transfer enzymes can achieve catalytic rate constants (k_{cat}) of greater than 100 s^{-1}, even when spontaneous rate constants are as low as 10^{-20} s^{-1}. As such, they possess some of the largest enzymatic accelerations identified, with catalytic enhancements approaching 10^{21}.[2] Part of these accelerations has often been ascribed to general acid−base catalysis that both augments phosphorylation rates by assisting deprotonation of the nucleophilic hydroxyl oxygen and enhances dephosphorylation rates by aiding protonation of the same oxygen atom (now the bridging oxygen of the phosphate group). Residues that satisfy the assignment of the general acid−base (commonly aspartate, glutamate, or histidine residues) are repeatedly conserved in the active sites of multiple superfamilies of phosphoryl transfer enzymes and are consistently identified by mutation studies as key elements of enzyme activity.[3−8] While structural studies reveal the close proximity of the general acid−base to reacting groups in near-transition-state complexes, the precise relationship of proton transfer to the mechanism of the phosphoryl transfer reaction remains uncertain. Density-functional-theory (DFT) models of the phosphoryl transfer step in some enzymes predict that proton transfer occurs only when there is

Received: April 25, 2018
Revised: July 17, 2018
Published: July 24, 2018

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

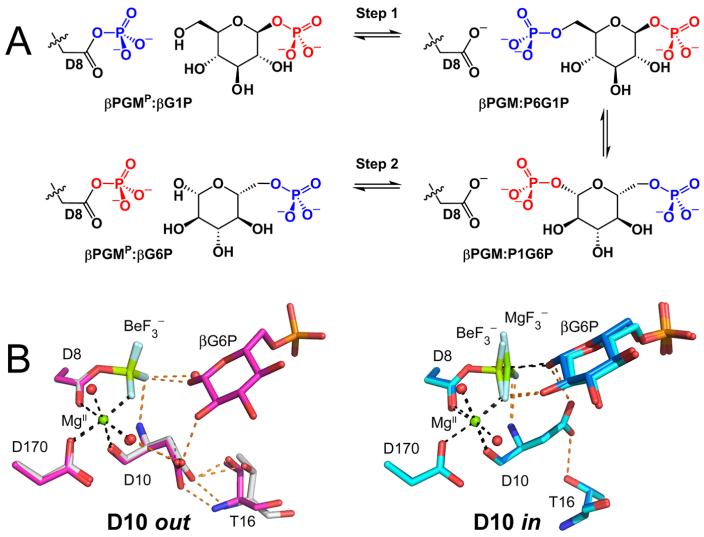


Figure 1. *β*PGM reaction scheme and change in orientation of residue D10, the assigned general acid−base. (A) *β*PGM reaction scheme for the enzymatic conversion of *β*G1P to G6P via a *β*G16BP intermediate. The phosphoryl transfer reaction between the phospho-enzyme (*β*PGM^P, phosphorylated at residue D8) and *β*G1P is termed Step 1 and is illustrated with the transferring phosphate (blue) in the *proximal* site and the 1-phosphate (red) of *β*G1P in the *distal* site. The equivalent reaction between *β*PGM^P and G6P is termed Step 2 and is shown with the transferring phosphate (red) in the *proximal* site and the 6-phosphate (blue) of G6P in the *distal* site. The two intermediate complexes are labeled *β*PGM:P6G1P and *β*PGM:P1G6P to explicitly denote the orientation of *β*G16BP bound in the active site. (B) The carboxylate group of residue D10 is in the *out* position in both the open substrate-free *β*PGM^P analogue structure (*β*PGM:BeF_3^- complex; PDB 2WFA;[23] gray carbon atoms) and in the hydrogen bonded NAC (*β*PGM:BeF_3^-:G6P complex; PDB 2WF9;[23] magenta carbon atoms). In contrast, the carboxylate group of residue D10 is in the *in* position in both the transition-state-analogue (TSA) structure (*β*PGM:MgF_3^-:G6P TSA complex; PDB 2WF5;[20] blue carbon atoms) and in the aligned NAC (*β*PGM:BeF_3^-:G6P complex; PDB 2WF8;[23] cyan carbon atoms). Selected active site residues and ligand are shown as sticks in standard CPK colors, with beryllium (light green), magnesium (green), and fluorine (light blue). Structural waters (red) and the catalytic Mg^{II} ion (green) are drawn as spheres. Orange dashes indicate hydrogen bonds and black dashes show metal ion coordination.

substantial bond formation between the hydroxyl nucleophile and the phosphorus atom,[9−13] but conclusions based on DFT models depend on how closely the protein conformation reflects that in which proton transfer takes place. However, solvent deuterium isotope effect measurements and the pH dependence of presteady-state kinetic analyses often support the DFT models in that the rate of phosphoryl transfer is interpreted to be independent of hydroxyl nucleophile deprotonation.[14−16] A resolution of the uncertainty over how the proton transfer step contributes to the catalytic cycle requires direct structural evidence of the protein conformation in which proton transfer occurs.

β-Phosphoglucomutase (*β*PGM) from *Lactococcus lactis* is a well-studied magnesium-dependent phosphoryl transfer enzyme of the haloacid dehalogenase (HAD) superfamily,[8,17−23] which catalyzes the reversible isomerization of *β*-glucose 1-phosphate (*β*G1P) and glucose 6-phosphate (G6P) (Figure 1A). The active site is located in the cleft formed between the helical cap domain (T16−V87) and the *α*/*β* core domain (M1−D15, S88−K216), with closure of the cleft through domain reorientation occurring during catalysis. The active site binds two phosphate groups, one in the *proximal* site adjacent to D8 and the catalytic Mg^{II} ion, and one in the *distal* site (∼8 Å away in the closed enzyme). *β*PGM transfers a phosphate group from the phospho-enzyme (*β*PGM^P, phosphorylated on the carboxylate side chain of residue D8) to the physiological substrate, *β*G1P, (Step 1)[19] forming an enzyme-bound *β*-glucose 1,6-bisphosphate (*β*G16BP) intermediate.[18] Subsequent release of *β*G16BP to solution permits its binding in the alternate orientation, leading to dephosphorylation of *β*G16BP (Step 2)[20] and the generation of G6P and *β*PGM^P as products (Figure 1A). In the Step 1 complexes, *β*PGM hydrogen bonds to the substrate directly, whereas in the Step 2 complexes, two water molecules mediate hydrogen bonding with substrate.[19] Structural investigations of species along the reaction coordinate have made extensive use of metal fluoride-based ground and transition-state-analogue complexes,[24,25] and have experimentally corroborated the in-line nucleophilic attack of phosphoryl transfer, the trigonal bipyramidal nature of the chemical transition state (TS), and the requirement for charge balance in the active site.[20−22] Moreover, these studies have highlighted how the carboxylate group of the assigned general acid−base (residue D10) can adopt different orientations.[8] In substrate-free *β*PGM and *β*PGM^P analogue structures,[20,23] the active site cleft is open and the D10 carboxylate is in the *out* position (Figure 1B). In transition-state-analogue (TSA) structures,[20] domain reorientation has closed the active site cleft and the D10 carboxylate is in the *in* position, where it is positioned to facilitate general acid−base catalysis. In the substrate-bound *β*PGM^P analogue structures containing BeF_3^-[23] two conformations are observed, in both of which the active site cleft is closed. One has the same conformation as the TSA structures, while in the other the cap and core domains have a relative rotation of 17° and the D10

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

carboxylate is in the *out* position. Both of the substrate-bound βPGM^{P} analogue structures conform to the criteria of near attack conformations (NACs).[26] The TSA-like conformation is termed an aligned NAC as the nucleophile is aligned to attack the BeF_3^- moiety, whereas the rotated conformation is termed a hydrogen-bonded NAC as the nucleophilic hydroxyl group is hydrogen bonded to the BeF_3^- moiety.[23] The observation of both NACs supports a model where the conformational change between the two closed forms is correlated with the *out* to *in* transition of D10 and the alignment of the substrate for nucleophilic attack.

The models above require extrapolation from the behavior of metal fluoride analogues in the active site to that of the substrates. While there is growing computational evidence for a close relationship between metal fluoride TSA complexes and the corresponding phosphoryl species,[27,28] there are few experimental systems where the properties of both species can be examined in detail. In order to address this, we sought to establish a stable enzyme:substrate complex using an aspartate to asparagine substitution, in a system for which the behavior of metal fluoride analogue complexes is well determined.[20,23] Here, we report the properties of several complexes involving the βPGM D10N variant (βPGM$_{D10N}$), which serves as a model of wild-type βPGM (βPGM$_{WT}$) with the general acid−base in its protonated form. This variant has previously been reported to be inactive,[8] and was expected to offer the opportunity to study βPGM^{P}:βG1P, βPGM^{P}:G6P, and βPGM:βG16BP complexes independently. Here we show that the βPGM$_{D10N}$ variant purifies as βPGM$_{D10N}$:βG16BP complexes. Low-level mutase activity was observed, which was enhanced once the noncovalently bound intermediate is removed by denaturation-refolding. Subsequently, exposure to substrate leads to the reformation of βPGM$_{D10N}$:βG16BP complexes in solution, and the trapping of two distinct βPGM$_{D10N}$:βG16BP complexes in crystallo, with either the 1- or the 6-phosphate group in the *proximal* site. In both of these complexes, the nucleophilic carboxylate oxygen and the phosphorus atoms are aligned and in van der Waals contact, but phosphoryl transfer is arrested by the failure of N10 to release a proton to βG16BP. However, the βPGM$_{D10N}$:βG16BP complexes do not adopt the fully closed conformation of the TSA complexes, indicating that such close proximity between reacting groups is insufficient to achieve the architecture used by the enzyme to bind the TS. Remarkably, the binding affinity of the catalytic Mg^{II} ion in the βPGM$_{D10N}$:βG16BP complexes is reduced compared with the phospho-enzyme analogue and the TSA complexes, which implies that antagonism within the coordination of the Mg^{II} ion facilitates the release of the high-affinity βG16BP intermediate.

EXPERIMENTAL METHODS

β-Phosphoglucomutase (βPGM) Expression, Purification, and Refolding. Site-directed mutagenesis (QuikChange II kit, Agilent Technologies) of the βPGM gene from *Lactococcus lactis* (EC 5.4.2.6) cloned in a pET22b+ vector was employed to generate the D10N variant (βPGM$_{D10N}$) and the D8N variant (βPGM$_{D8N}$) using primers with single-site base changes and mutagenesis of the βPGM gene was confirmed by DNA sequencing. Wild-type βPGM (βPGM$_{WT}$), βPGM$_{D10N}$ and βPGM$_{D8N}$ proteins were expressed using natural abundance, ^{15}N or $^{2}H^{15}N^{13}C$ isotopic enrichment[21,29] and purified using the following methodology which minimized the presence of contaminating phosphoryl transfer enzymes (e.g., phosphoglucose isomerase and βPGM from *E. coli*). The cell pellet was resuspended in ice-cold standard native buffer (50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3) supplemented with one tablet of cOmplete protease inhibitor cocktail (Roche). The cell suspension was lysed on ice by sonication for 5 cycles of pulsation for 20 s with 60 s cooling intervals. The cell lysate was then separated by ultracentrifugation (Beckman Coulter Avanti centrifuge) at 24 000 rpm for 35 min at 4 °C. The cleared cell lysate was filtered using a 0.2 μM syringe filter and loaded onto a DEAE-Sepharose fast flow ion-exchange column connected to an ÄKTA purification system that had been washed previously with 1 column volume of 6 M guanidine hydrochloride, 1 column volume of 1 M NaOH and equilibrated with 5 column volumes of standard native buffer. Following extensive washing, proteins bound to the DEAE-Sepharose column were eluted with a gradient of 0 to 100% standard native buffer containing 0.5 M NaCl. Fractions containing βPGM were checked for purity using SDS-PAGE, were pooled together, and concentrated by Vivaspin (10 kDa MWCO). The protein sample was filtered using a 0.2 μM syringe filter and loaded onto a prepacked Hiload 26/60 Superdex 75 size-exclusion column connected to an ÄKTA purification system that had been washed previously with 1 column volume of 1 M NaOH and equilibrated with 5 column volumes of filtered and degassed standard native buffer containing 1 M NaCl. Fractions containing βPGM were checked for purity using SDS-PAGE, were pooled together, buffer exchanged into standard native buffer, and concentrated to 2 mM by Vivaspin (10 kDa MWCO) for storage as 1 mL aliquots at −20 °C.

In contrast to βPGM$_{WT}$ and βPGM$_{D8N}$, βPGM$_{D10N}$ copurified with βG16BP as tight, noncovalently bound βPGM$_{D10N}$:βG16BP complexes. Substrate-free βPGM$_{D10N}$ was prepared from the copurified βPGM$_{D10N}$:βG16BP complexes using an unfolding-dilution-refolding strategy to remove βG16BP. Samples of the copurified βPGM$_{D10N}$:βG16BP complexes were diluted into unfolding buffer (4 M guanidine hydrochloride, 50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3), buffer exchanged by Vivaspin (10 kDa MWCO) in unfolding buffer to dilute βG16BP by 200-fold, and the retained βPGM$_{D10N}$ was refolded by pulse renaturation or dialysis into standard native buffer. A final buffer exchange to remove any remaining denaturant was performed using a Vivaspin (3 kDa MWCO), and the protein was concentrated to 2 mM for storage as 1 mL aliquots at −20 °C. Removal of βG16BP from βPGM$_{D10N}$ was confirmed by ^{31}P NMR spectroscopy in standard NMR buffer (50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3, 10% (v/v) $^{2}H_2O$, and 1 mM trimethylsilyl propanoic acid (TSP)).

The reconstituted βPGM$_{D10N}$:βG16BP complexes were formed by the addition of 20 mM acetyl phosphate (AcP) and 10 mM glucose 6-phosphate (G6P) or 10 mM β-glucose 1-phosphate (βG1P) to 1 mM substrate-free βPGM$_{D10N}$ in 200 mM K^+ HEPES buffer (pH 7.2), 5 mM $MgCl_2$, and 2 mM NaN_3. Unbound ligands in the sample (excess G6P, βG1P, and AcP) were removed by buffer exchange into standard NMR buffer.

Reagents. Unless otherwise stated, reagents were purchased from Sigma-Aldrich, GE Healthcare, Melford Laboratories, or CortecNet.

βG16BP was isolated from the copurified βPGM$_{D10N}$:βG16BP complexes in standard NMR buffer by

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

heat denaturation of βPGM$_{D10N}$ (2 min at 80 °C), centrifugation at 13 000 rpm to remove denatured βPGM$_{D10N}$, and filtration of the supernatant containing βG16BP using a Vivaspin (3 kDa MWCO). Resonance assignments of βG16BP were confirmed by ^{31}P and natural abundance ^{1}H^{13}C HSQC NMR spectra following the addition of 6 mM EDTA to the sample.

βG1P was synthesized enzymatically from maltose using maltose phosphorylase (EC 2.4.1.8). Maltose (1 M) was incubated overnight at 30 °C with 1.5 units mL^{-1} maltose phosphorylase in 0.5 M phosphate buffer (pH 7.0). βG1P production was confirmed using ^{31}P NMR spectroscopy. Maltose phosphorylase (90 kDa) was removed using a Vivaspin (5 kDa MWCO), and the resulting flow-through solution containing βG1P was used without further purification. The concentration of βG1P was measured to be 150 mM by quantitative ^{31}P NMR spectroscopy (recycle time 60 s) against a known concentration of G6P. The concentrations of other components in the solution were estimated as follows: 150 mM glucose, 850 mM maltose, and 350 mM inorganic phosphate.

Uniformly ^{13}C-labeled G6P was synthesized enzymatically from 45 mM uniformly ^{13}C-labeled D-glucose by incubation for 90 min at 37 °C with 14 units mL^{-1} hexokinase (EC 2.7.1.1) and 50 mM ATP in 100 mM Tris-HCl (pH 8.0), 50 mM $MgCl_2$, and 2 mM EDTA. G6P production was confirmed using ^{31}P NMR spectroscopy. Hexokinase (110 kDa) was removed by denaturation at 80 °C followed by filtration using a Vivaspin (3 kDa MWCO). The flow-through containing uniformly ^{13}C-labeled G6P was used without further purification together with AcP and substrate-free βPGM$_{D10N}$ for the formation of uniformly ^{13}C-labeled βG16BP in the reconstituted βPGM$_{D10N}$:βG16BP complexes.

Chemically synthesized βG16BP was a gift from Prof. Nicholas Williams, Department of Chemistry, The University of Sheffield.[30]

NMR Spectroscopy. *Instruments and Data Processing.* NMR experiments were acquired at 298 K using Bruker spectrometers located at the following institutions: Department of Molecular Biology and Biotechnology (MBB), The University of Sheffield; School of Chemistry (SC), The University of Manchester; Manchester Institute of Biotechnology (MIB), The University of Manchester. Experiments were processed using TopSpin (Bruker) or FELIX (Felix NMR, Inc.), and figures were prepared using either FELIX or CcpNmr Analysis.[31] ^{1}H chemical shifts were referenced relative to the internal TSP signal resonating at 0.0 ppm and ^{13}C, ^{15}N, and ^{31}P chemical shifts were referenced indirectly using nuclei-specific gyromagnetic ratios.

^{1}H^{15}N TROSY Spectra. ^{1}H^{15}N TROSY spectra of βPGM$_{WT}$ and substrate-free βPGM$_{D10N}$ were acquired 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) $^{2}H_2O$ and 2 mM TSP) containing 50 mM $MgCl_2$. ^{1}H^{15}N TROSY spectra of the βPGM$_{WT}$:BeF_3 and βPGM$_{D10N}$:BeF_3 complexes were acquired using 0.5−1 mM ^{15}N-βPGM$_{WT}$ or ^{15}N-substrate-free βPGM$_{D10N}$ in standard NMR buffer containing 5 mM $BeCl_2$ and 10 mM NH_4F. Experiments were recorded using a Bruker 600 MHz Avance DRX spectrometer equipped with a TXI cryoprobe and z-axis gradients (MBB) or a Bruker 800 MHz Avance I spectrometer equipped with a TXI probe and z-axis gradients (MBB).

^{31}P Spectra. One-dimensional ^{31}P spectra to characterize βG16BP and the βPGM$_{D10N}$:βG16BP complexes were acquired using a Bruker 500 MHz Avance DRX spectrometer (operating at 202.456 MHz for ^{31}P) equipped with a broadband probe (MBB). A spectral width of 50 ppm centered at −10 ppm enabled the observation of the relevant phosphorus signals. Typically, accumulations of 10 000 transients without proton-phosphorus decoupling were necessary to achieve a sufficient signal-to-noise ratio with sample concentrations in the 0.5−1 mM range. Spectra were processed with baseline correction and 10 Hz Lorentzian apodization.

^{31}P Spectra for Kinetic Measurements. Reaction kinetics for βPGM-catalyzed reactions were followed using a Bruker 500 MHz Avance III HD spectrometer (operating at 202.48 MHz for ^{31}P) equipped with a Prodigy BBO cryoprobe (SC), which offered significant improvements in signal sensitivity. One-dimensional ^{31}P spectra without proton-phosphorus decoupling were recorded 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 equilibrations of 10 mM βG1P with G6P by 0.1−1 μM βPGM$_{WT}$, 5−50 μM substrate-free βPGM$_{D10N}$, and 10 μM βPGM$_{D8N}$ were measured in standard kinetic buffer (200 mM K$^+$ HEPES buffer (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3, 10% $^{2}H_2O$, 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 spectra. The equilibration of 10 mM βG1P with G6P by 5 μM substrate-free βPGM$_{D10N}$ using βG16BP extracted from the copurified βPGM$_{D10N}$:βG16BP complexes as a priming agent was measured in standard kinetic buffer monitored by one-dimensional ^{31}P spectra recorded without proton-phosphorus decoupling with 256 transients and a 1 s recycle delay using a Bruker 500 MHz Avance DRX spectrometer (MBB). Normalized integral values of both the βG1P and G6P peaks following baseline correction and 2 Hz Lorentzian apodization 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 catalytic rate constant, k_{cat}. The hydrolysis kinetics of 50 mM AcP to inorganic phosphate and acetate by 250 μM βPGM was measured in standard kinetic buffer containing 50 mM $MgCl_2$ and 1 mM EDTA. The reaction was timed from the addition of AcP and monitored by the acquisition of consecutive ^{31}P spectra. Normalized integral values of the AcP peak following baseline correction and 2 Hz Lorentzian apodization were plotted against time, and the rate constant for AcP hydrolysis was derived from linear least-squares fitting of the data. A control experiment involving 50 mM AcP alone in standard kinetic buffer established that hydrolysis of AcP was insignificant over the same time frame. Throughout all the kinetic measurements, the pH of the reactions was found to be invariant as assessed in situ by the ^{31}P resonance of inorganic phosphate and the ^{1}H resonances of 200 mM HEPES buffer.

^{1}H^{13}C HSQC and 2D CCH-TOCSY Spectra of Glucose 1,6-Bisphosphate Species. Natural-abundance ^{1}H^{13}C HSQC spectra of αG16BP and βG16BP (in 100% $^{2}H_2O$ and 1 mM TSP) were recorded on a Bruker 500 MHz Avance DRX spectrometer equipped with a TXI probe and z-axis gradients (MBB).[30] To assign the bound βG16BP resonances in the reconstituted βPGM$_{D10N}$:βG16BP complexes, ^{1}H^{13}C HSQC and 2D CCH-TOCSY spectra were acquired with 0.5−1 mM ^{15}N-labeled substrate-free βPGM$_{D10N}$ in standard NMR buffer

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

containing 20 mM AcP and 10 mM uniformly ^{13}C-labeled G6P using a Bruker Avance III 800 MHz spectrometer equipped with a TCI cryoprobe and *z*-axis gradients (MIB).

^{1}H^{15}N BEST-TROSY Experiments. Rapid acquisition ^{1}H^{15}N BEST-TROSY spectra[32,33] to follow βPGM$_{D10N}$-catalyzed reactions were acquired using 1 mM substrate-free βPGM$_{D10N}$ in standard kinetic buffer containing either 20 mM AcP or 20 mM AcP and 10 mM βG1P. ^{1}H^{15}N BEST-TROSY spectra were recorded using a Bruker 600 MHz Avance DRX spectrometer equipped with a TXI cryoprobe and *z*-axis gradients (MBB) as 6 min experiments (4 transients, 200 increments and a recycle delay of 0.3 s) with selective ^{1}H pulses centered on the amide region (8.7 ppm). Excitation pulses (90°) were 2 ms at 600 MHz (pulse shape Pc9_4) and 1.7 ms at 600 MHz (pulse shape Eburp2), whereas refocusing pulses (180°) were 1.6 ms at 600 MHz (pulse shape Reburp). The experimental dead-time was approximately 6 min.

Backbone Resonance Assignment of the βPGM$_{D10N}$:βG16BP Complexes. For the ^{1}H, ^{13}C, and ^{15}N backbone resonance assignment of the reconstituted βPGM$_{D10N}$:βG16BP complexes, multidimensional heteronuclear NMR spectra were acquired with 0.5−1 mM ^{2}H^{15}N^{13}C-labeled substrate-free βPGM$_{D10N}$ in standard NMR buffer containing 20 mM AcP and 10 mM G6P using a Bruker 800 MHz Avance III spectrometer equipped with a TCI cryoprobe and *z*-axis gradients (MIB). The standard suite of ^{1}H^{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 nonuniform sampling (NUS) with a multidimensional Poisson Gap scheduling strategy with exponential weighting.[34] NUS data were reconstructed using TopSpin3 and multidimensional decomposition.[35] Backbone resonance assignments of the Mg^{II}-bound βPGM$_{D10N}$:P1G6P and Mg^{II}-free βPGM$_{D10N}$:P1G6P complexes present simultaneously in the spectra were obtained using a simulated annealing algorithm employed by the *asstools* assignment program.[29] Assignments for the two complexes were confirmed by using ^{1}H^{15}N TROSY spectra of separate Mg^{II}-bound and Mg^{II}-free ^{15}N-βPGM$_{D10N}$:P1G6P complexes, together with sequential backbone amide to amide correlations obtained from TROSY-based (H)N(COCA)NNH and H(NCOCA)NNH experiments.[36] The Mg^{II}-free ^{15}N-βPGM$_{D10N}$:P1G6P complex was prepared by dilution of Mg^{II} by over 20 000 fold using buffer exchange into standard NMR buffer in the absence of $MgCl_2$, while the Mg^{II}-bound ^{15}N-βPGM$_{D10N}$:P1G6P complex was prepared in standard NMR buffer containing 50 mM $MgCl_2$.

Determination of the Mg^{II} Dissociation Constant. A Mg^{II}-free ^{15}N-βPGM$_{D10N}$:βG16BP complex was prepared from a reconstituted Mg^{II}-bound ^{15}N-βPGM$_{D10N}$:βG16BP complex by buffer exchange (3000-fold dilution) and overnight equilibration into standard NMR buffer (containing no $MgCl_2$). A discontinuous titration of 0−47.6 mM $MgCl_2$ into separate Mg^{II}-free ^{15}N-βPGM$_{D10N}$:βG16BP samples with overnight equilibration was monitored by ^{1}H^{15}N TROSY spectra recorded using a Bruker 800 MHz Avance I spectrometer equipped with a TXI probe and *z*-axis gradients (MBB). Peak intensities for well-resolved resonances of the Mg^{II}-bound βPGM$_{D10N}$:βG16BP complex (residues N10, G11, A115, K117, and I150) were averaged and normalized against the intensity of the side chain HNε1 resonance of W216, which remains unchanged throughout the titration. The dissociation constant (K_d) was obtained by fitting changes in normalized peak intensity as a function of Mg^{II} concentration to a single-site binding isotherm[37] using a nonlinear least-squares fitting algorithm. The solution concentration of Mg^{II} present at the beginning of the titration was derived from the fitting procedure.

X-ray Crystallography. *Crystallization and Data Collection.* Frozen aliquots of substrate-free βPGM$_{D10N}$ or copurified βPGM$_{D10N}$:βG16BP complex 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. Specific ligands were added to a solution of substrate-free βPGM$_{D10N}$ to generate crystals of the following complexes: βPGM$_{D10N}$:BeF_3 complex (5 mM $BeCl_2$ and 15 mM NaF), βPGM$_{D10N}$:P1G6P and βPGM$_{D10N}$:P6G1P complexes (15 mM βG1P, 5 mM $BeCl_2$ and 15 mM NaF), and βPGM$_{D10N}$:AlF_4:G6P complex (10 mM G6P, 5 mM $AlCl_3$ and 20 mM NaF). Crystals of the βPGM$_{D10N}$:AlF_4:H_2O:βG1P complex were obtained from a solution of the copurified βPGM$_{D10N}$:βG16BP complexes containing 5 mM βG1P, 2 mM $AlCl_3$, and 10 mM NH_4F. Crystals of the copurified βPGM$_{D10N}$:P1G6P complex were obtained from a solution of the copurified βPGM$_{D10N}$:βG16BP complexes. The solutions were adjusted to a protein concentration of 0.6 mM, were incubated for 1 h, and mixed 1:1 with precipitant (24−34% (w/v) PEG 4000 or 19−21% (w/v) PEG 3350, 50−200 mM sodium acetate and 0−100 mM Tris (pH 7.5)). Crystals were grown at 290 K by hanging-drop vapor diffusion using a 2 μL drop suspended on a siliconized glass coverslip 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 either cryo-protected in their mother liquor containing an additional 25% (v/v) ethylene glycol or excess mother liquor was removed[38] 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.

Data Processing, Structural Determination, and Refinement. At the DLS, data were processed using the xia2 pipeline,[39] whereas at the ESRF, data were processed with iMOSFLM.[40] Resolution cut-offs were applied using either CC-half or by consideration of the $\langle I/\sigma(I)\rangle$ and R_{merge} values. All the crystals belonged to the spacegroup $P2_12_12_1$, with cell dimensions that varied depending on the degree of enzyme closure. Structures were determined by molecular replacement with MolRep[41] 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[42] with ligands not included until the final rounds of refinement using REFMAC5[43] so that they could be built into unbiased difference Fourier maps. When structures were refined with down-weighted B-factor restraints, the B-factors of the ligands in the resulting structures were equivalent to those of the surrounding protein, suggesting that the degree of accuracy in the placement of the ligand atoms was equivalent to those of the protein atoms. Structures with a resolution better than 1.4 Å were refined with anisotropic B-factors. Structure validation was carried out in COOT and MolProbity.[44] Superpositions were carried out using PyMOL,[45] maps were generated using FFT,[46] and domain movements were calculated using DynDom.[47] Additional details for X-ray crystallography data

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

collection, data processing and refinement are provided in Table S1 in the Supporting Information.

Crystallization of the βPGM$_{D10N}$:P1G6P and the βPGM$_{D10N}$:P6G1P Complexes. Rod-shaped crystals harvested after 1 week contained predominantly βG16BP in the βPGM$_{D10N}$ active site, with the 6-phosphate group located in the *proximal* site and the 1-phosphate group bound in the *distal* site (βPGM$_{D10N}$:P6G1P complex). After refinement, the ratio of 2Fo − Fc contour thresholds between the 1- and 6-phosphate groups (ca. 6σ and 5σ, respectively) did not correlate with a full βG16BP ligand occupancy in the βPGM$_{D10N}$:P6G1P complex. When modeled at a ligand occupancy of 0.8, B-factor convergence was attained between the βG16BP ligand and neighboring residues in the active site, confirming βG16BP as the dominant ligand. Remaining difference map peaks were consistent with the presence of a minor population of βG1P (with the 1-phosphate in the *distal* site); however, because of poor connectivity at this resolution, βG1P was not modeled into the structure. Crystals from the same drop with the same morphology harvested after 12 weeks contained only βG16BP bound in the alternate orientation with the 1-phosphate group located in the *proximal* site and the 6-phosphate group bound in the *distal* site (βPGM$_{D10N}$:P1G6P complex).

Steady-State Kinetic Assays. Steady-state kinetic assays for βPGM$_{WT}$ and substrate-free βPGM$_{D10N}$ were conducted at 294 K using a FLUOstar OMEGA microplate reader (BMG Labtech) in 200 mM K^+ HEPES buffer (pH 7.2) containing 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 oxidized 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 substrate-free βPGM$_{D10N}$ stock concentrations were determined using a NanoDrop One^C 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, the reaction was initiated by the addition of 10 mM AcP to solutions of 0.5 mM NAD^+ and 5 units mL^{-1} G6PDH containing either 5 nM βPGM$_{WT}$ or 500 nM substrate-free βPGM$_{D10N}$ and variable concentrations of βG1P (5, 15, 35, 50, 70, 100, 160, 230, 330 μM). 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. For the fluoride inhibition experiments monitored using the G6PDH coupled assay, the reaction was initiated by the addition of 10 mM AcP to solutions of 230 μM βG1P, 0.5 mM NAD^+ and 5 units mL^{-1} G6PDH containing either 5 nM βPGM$_{WT}$ or 500 nM substrate-free βPGM$_{D10N}$ and variable concentrations of NaF (0, 1, 2, 3, 5, 7, 10 mM). The linear steady-state portion of G6P production was not used for the analysis of fluoride inhibition as βG16BP formation during the reaction outcompetes fluoride inhibition.[21] The presence of increasing levels of fluoride in the reaction buffer extends the lag phase prior to achieving steady-state kinetics, the duration of which was estimated using a first derivative approach. The time point at which the maximum value was reached in the first derivative vs time plot for each reaction containing fluoride was normalized against the time point for the reaction in the absence of fluoride. A line of best fit for the normalized values vs fluoride concentration was determined using a polynomial function.

RESULTS

Recombinant βPGM$_{D10N}$ Copurifies in Complex with βG16BP. βPGM$_{D10N}$ was produced and purified as for βPGM$_{WT}$ with slight modifications to published procedures.[48−50] A ^{31}P NMR spectrum demonstrated that, unlike βPGM$_{WT}$, βPGM$_{D10N}$ copurifies with tightly bound phosphorylated glucose ligands (Figure 2A). Four ^{31}P resonances are observed, two with chemical shifts corresponding to a 1-phosphate group and two to a 6-phosphate group of glucose. The ratio of intensities of the resonances suggests that the phosphate groups are paired, consistent with the population of two complexes. Ligand extraction was achieved by the removal of heat-denatured βPGM$_{D10N}$ (2 min at 80 °C) using centrifugation followed by membrane filtration of the supernatant. ^{31}P and $^1H^{13}C$ HSQC NMR spectra indicated that a single ligand had been isolated, which revealed that both complexes contained the same phosphorylated glucose species (Figure S1A,C,D and Figure S2B in the Supporting Information). The ligand was identified as βG16BP (the reaction intermediate, Figure 1A) by comparison with synthetic α- and β-glucose 1,6-bisphosphate species (Figure S1E and Figure S2A). The high affinity of βPGM$_{D10N}$ for the βG16BP intermediate is predictable because kinetic data for βPGM$_{WT}$ has identified that βG16BP is the tightest binding species of the native substrates, with K_m = 0.63 μM[8] and K_m = 0.72 μM.[30] Substitution of aspartate with asparagine at residue 10 is likely to increase the binding affinity of βPGM$_{D10N}$ for βG16BP since the deprotonated D10 side chain in βPGM$_{WT}$ does not satisfy charge balance[24] within the complex. Substrate-free βPGM$_{D10N}$ was prepared from the copurified βPGM$_{D10N}$:βG16BP complexes by unfolding the recombinant protein in 4 M guanidine hydrochloride together with a 200-fold dilution of the ligand using buffer exchange and subsequent refolding of βPGM$_{D10N}$ (Figure S1B). A comparison of the $^1H^{15}N$ TROSY spectra of substrate-free βPGM$_{D10N}$ and βPGM$_{WT}$ indicated that βPGM$_{D10N}$ adopts a native conformation following refolding (Figure S3A).

Substrate-Free βPGM$_{D10N}$ Readily Forms a Transient Phospho-Enzyme. βPGM$_{WT}$ can be phosphorylated to generate βPGM$_{WT}^P$ by a number of priming agents, including not only βG16BP (Figure 1A) but also αG16BP, G6P, and acetyl phosphate (AcP).[17,30] In order to establish whether βPGM$_{D10N}$ could be similarly phosphorylated, incubation of 1 mM substrate-free βPGM$_{D10N}$ with 20 mM AcP was followed using a time course of $^1H^{15}N$ BEST-TROSY spectra[32,33] with 6 min time resolution (Figure 2B). The initial spectra overlaid closely with a $^1H^{15}N$ TROSY spectrum of the βPGM$_{D10N}$:BeF_3 complex, which is an analogue of βPGM$_{D10N}^P$ prepared using conditions described previously for the βPGM$_{WT}$:BeF_3 complex (Figure S3B).[23] This established that βPGM$_{D10N}^P$ is generated during the 6 min dead-time of the time course. After 98 min, the $^1H^{15}N$ BEST-TROSY spectrum had reverted entirely to that of substrate-free βPGM$_{D10N}$. Monitoring the same reaction using ^{31}P NMR spectra, the hydrolysis rate constant for βPGM$_{D10N}^P$ was determined to be 0.020 ± 0.002 s^{-1} (Figure S3C). The equivalent rate constant for βPGM$_{WT}^P$ under the same conditions is only 3 fold greater (0.060 ±

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

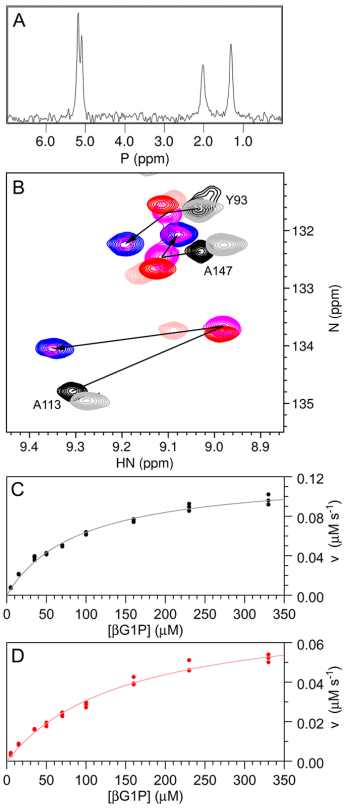


Figure 2. NMR spectra and reaction kinetics of βPGM$_{D10N}$. (A) ^{31}P spectrum of βPGM$_{D10N}$ immediately following purification showing four ^{31}P peaks (5.17, 5.08, 2.01, and 1.30 ppm) consistent with the population of two noncovalently bound βPGM$_{D10N}$:βG16BP complexes (ratio 6:5). Resonances at ~5 ppm and 1−2 ppm correspond to 6-phosphate and 1-phosphate groups of βG16BP, respectively. (B) Overlay of a section of ^{1}H^{15}N TROSY spectra for a range of βPGM$_{D10N}$ complexes: (black) substrate-free βPGM$_{D10N}$; (pink) βPGM$_{D10N}$:BeF$_3$ complex; (red) βPGM$_{D10N}^{P}$ – ^{1}H^{15}N BEST-TROSY spectrum started 6 min after addition of 20 mM AcP to substrate-free βPGM$_{D10N}$; (gray) substrate-free βPGM$_{D10N}$ – ^{1}H^{15}N BEST-TROSY spectrum started after a further 92 min by which time AcP has been depleted and βPGM$_{D10N}^{P}$ has reverted to substrate-free βPGM$_{D10N}$ (the small shift in peak positions is caused by an increase in inorganic phosphate concentration); (magenta) βPGM$_{D10N}^{P}$ as major species – ^{1}H^{15}N BEST-TROSY spectrum started 6 min after addition of 10 mM G6P and 20 mM AcP to substrate-free βPGM$_{D10N}$; (blue) βPGM$_{D10N}$:βG16BP complexes – ^{1}H^{15}N BEST-TROSY spectrum started after a further 145 min by which time AcP has been depleted and the βPGM$_{D10N}$:βG16BP complexes dominate in solution. The arrows indicate progression for the assigned residues from (black) substrate-free βPGM$_{D10N}$ to (magenta) βPGM$_{D10N}^{P}$ to (blue) the βPGM$_{D10N}$:βG16BP complexes. (C and D) Michaelis−Menten plots showing the dependence of the reaction velocity (v) for 5 nM βPGM$_{WT}$ (black circles; n = 3) and 500 nM substrate-free βPGM$_{D10N}$ (red circles; n = 3) on the initial βG1P concentration, monitored using a glucose 6-phosphate dehydrogenase coupled assay. Data were fitted to the standard Michaelis−Menten equation to derive k_{cat} and K_m values and the line of best fit is shown for βPGM$_{WT}$ (gray) and substrate-free βPGM$_{D10N}$ (pink).

0.006 s^{-1}), indicating that the proposed general acid−base (D10) has little involvement in the attack of βPGM^{P} by water. Attempts to crystallize the metastable species βPGM$_{D10N}^{P}$ were unsuccessful. However, the βPGM$_{D10N}$:BeF$_3$ complex was crystallized and the structure was determined to 1.3 Å resolution (PDB 5OJZ; Figure 3A,G, Figure S4A, and Table S1). The cap and core domains were in a predominantly open conformation, as in the βPGM$_{WT}$:BeF$_3$ complex (PDB 2WFA;[23] non-H atom RMSD = 1.06 Å), and the side chain of residue N10 was in the *out* position (Figure 1B), thereby not positioned to contribute to the nucleophilic attack of βPGM$_{D10N}^{P}$ by water. The close similarity of ^{1}H^{15}N TROSY spectra between βPGM$_{WT}$:BeF$_3$, βPGM$_{D10N}$:BeF$_3$, and βPGM$_{D10N}^{P}$ indicates that these structural features are common to all three species in solution.

Substrate-Free βPGM$_{D10N}$ Preparation Has Mutase Activity. In addition to substrate-free βPGM$_{D10N}$ having similar levels of phosphatase activity to βPGM$_{WT}$, the substrate-free βPGM$_{D10N}$ preparation was also found to have mutase activity. The standard glucose 6-phosphate dehydrogenase coupled assay[8,17,18] was used to monitor conversion of βG1P to G6P using AcP as the priming agent. The kinetic profile displayed the characteristic lag phase for βPGM (Figure S3I,J),[30] and a simple steady-state Michaelis−Menten analysis of the linear portion (Figure 2D), yielded values for k_{cat} of 0.15 ± 0.01 s^{-1} and K_m of 150 ± 12 μM. Measurements under the same conditions for βPGM$_{WT}$ (Figure 2C), yielded values of 24.5 ± 0.7 s^{-1} and 92 ± 6 μM, respectively; minor levels of inhibition by the priming agent[17,30] is a likely source of the slightly different values determined here compared with some reported previously for βPGM$_{WT}$.[8,30] Contaminating *E. coli* βPGM$_{WT}$ is unlikely to be the source of mutase activity in the substrate-free βPGM$_{D10N}$ preparation as there is no equilibration of βG1P with G6P over a similar time frame by βPGM$_{D8N}$ (Figure S3D), which has identical chromatography retention characteristics to βPGM$_{D10N}$. To investigate whether the activity of the substrate-free βPGM$_{D10N}$ preparation was the result of recovery by acetate (derived from AcP hydrolysis) substituting for the general acid−base, the equilibration of βG1P with G6P was primed with βG16BP rather than AcP (Figure S3E). Mutase activity was again observed (with a slightly larger rate constant, k_{cat} = 0.6 s^{-1}, as there is no inhibition when βG16BP is used as the priming agent) and thus acetate was not playing a significant role in recovery of activity. In contrast, it has not been possible to eliminate low levels (~0.6%) of contaminating *L. lactis* βPGM$_{WT}$ as the source of mutase activity because the measured K_m values, and degree of inhibition by inorganic phosphate (Figure S3F) and by fluoride (Figure S3G) are not sufficiently different between the substrate-free βPGM$_{D10N}$ preparation and βPGM$_{WT}$. Low levels of βPGM$_{WT}$ can potentially be formed by translational mis-incorporation or by deamidation of βPGM$_{D10N}$ during refolding, where the N10-G11 sequence will have elevated susceptibility.[51] However, it is difficult to rationalize the dominant effect arising either from translational mis-incorporation, when an increase in mutase activity is observed following βG16BP removal (k_{cat} = 0.002 s^{-1} for copurified βPGM$_{D10N}$ vs k_{cat} = 0.2 s^{-1} for the substrate-free βPGM$_{D10N}$ preparation), or from deamidation, when only a 2-fold increase in activity is observed following 2 h vs 48 h incubation with 4 M guanidine hydrochloride prior to refolding (Figure S3K,L).

Substrate-Free βPGM$_{D10N}$ Slowly Reforms Stable βG16BP Complexes. In order to establish that the

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

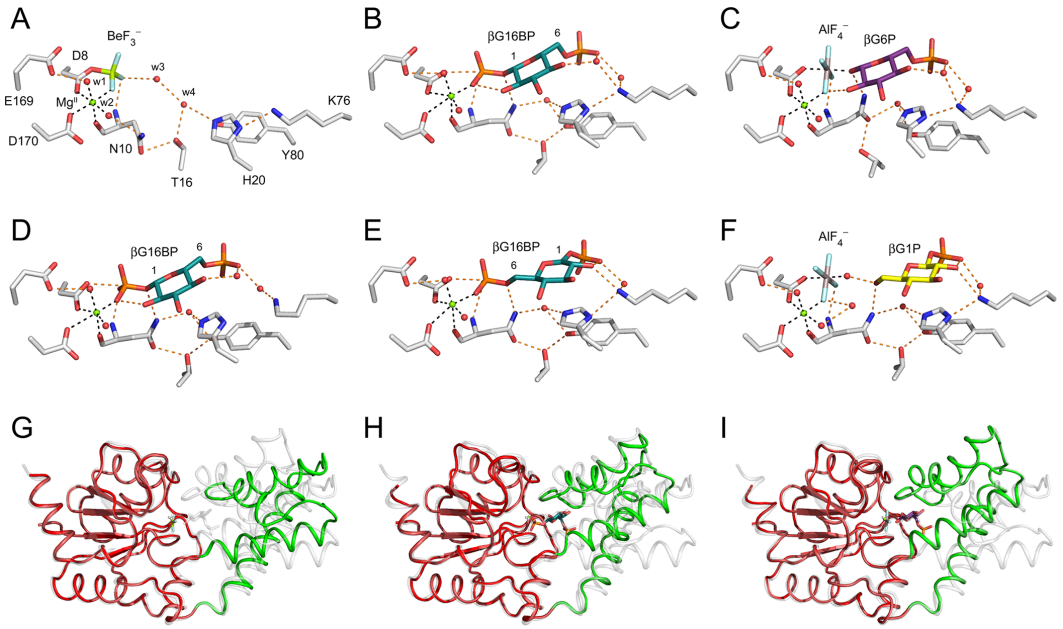


Figure 3. Overviews of the active sites and the extent of domain closure in the βPGM$_{D10N}$ complexes. The active sites of (A) βPGM$_{D10N}$:BeF$_3$ complex (PDB 5OJZ), (B) βPGM$_{D10N}$:P1G6P complex (PDB 5OK1), (C) βPGM$_{D10N}$:AlF$_4$:G6P complex (PDB 5OK2), (D) copurified βPGM$_{D10N}$:P1G6P complex (PDB 5O6P), (E) βPGM$_{D10N}$:P6G1P complex (PDB 5OK0) and (F) βPGM$_{D10N}$:AlF$_4$:H$_2$O:βG1P complex (PDB 5O6R). Selected active site residues and ligands are shown as sticks in standard CPK colors, with beryllium (light green), fluorine (light blue), aluminum (dark gray), βG16BP (teal carbon atoms; with C1 and C6 labeled for clarity), G6P (purple carbon atoms) and βG1P (gold carbon atoms). Structural waters (red) and the catalytic Mg^{II} ion (green) are drawn as spheres. Orange dashes indicate hydrogen bonds and black dashes show metal ion coordination. The extent of domain closure is shown in (G) βPGM$_{D10N}$:BeF$_3$ complex (PDB 5OJZ), (H) βPGM$_{D10N}$:P1G6P complex (PDB 5OK1) and (I) βPGM$_{D10N}$:AlF$_4$:G6P complex (PDB 5OK2). The protein backbone of βPGM$_{D10N}$ is depicted as a ribbon, with the core (red) and the cap (green) domains indicated and the ligands shown as sticks and spheres (colored as above). The pale gray ribbons indicate the open βPGM$_{WT}$ structure (PDB 2WHE[20]) and the fully closed βPGM$_{WT}$:MgF$_3$:G6P TSA complex (PDB 2WF5[20]) superposed on the core domains to show the extent of domain closure in the βPGM$_{D10N}$ complexes.

substrate-free βPGM$_{D10N}$ preparation was capable of reconstituting the βPGM$_{D10N}$:βG16BP complexes in situ, the equilibration of 10 mM βG1P with G6P (and vice versa) by 1 mM substrate-free βPGM$_{D10N}$ in the presence of 20 mM AcP was monitored using a time course of ^{1}H^{15}N BEST-TROSY spectra with 6 min time resolution (Figure 2B). At this concentration of substrate-free βPGM$_{D10N}$, βG1P and G6P were fully equilibrated (via βG16BP, Figure 1A) in the 6 min dead-time of the time course, and the initial enzyme species observed was βPGM$_{D10N}^{P}$. βPGM$_{D10N}^{P}$ was slowly replaced (k_{obs} = 5 × 10^{-4} s^{-1}) by two conformationally distinct species (Figure S5), that reproduce the ^{31}P NMR spectrum of the copurified βPGM$_{D10N}$:βG16BP complexes (Figure 2A). When 20 mM AcP and 10 mM βG1P were added to the reconstituted βPGM$_{D10N}$:βG16BP complex preparation, the rate constant of equilibration was within error of that of the original substrate-free βPGM$_{D10N}$ preparation (Figure S3H).

Nucleophile in the βPGM$_{D10N}$:P1G6P Complex Is Aligned for Attack. The βPGM$_{D10N}$:βG16BP complexes were explored using X-ray crystallography to compare their structures with those of metal fluoride analogue complexes.[19,20,23] A reconstituted βPGM$_{D10N}$:βG16BP complex was crystallized and the structure was determined to 1.9 Å resolution (PDB 5OK1; Figure 3 B,H, Figure S4B and Table S1). In this structure, βG16BP is bound in a single orientation, with the 1-phosphate in the *proximal* site and the 6-phosphate in the *distal* site, and is hence termed the βPGM$_{D10N}$:P1G6P complex. This structure mimics the active site conformation immediately preceding phosphoryl transfer from βG16BP to βPGM in Step 2 (Figure 1A). This conformation requires a protonated general acid−base and its surrogate, N10, forms a hydrogen bond through its side chain amide group to the bridging oxygen of the 1-phosphate of βG16BP. The 1-phosphorus atom is positioned in-line for attack by D8 atom Oδ1 (O−P−O angle = 170°) with a donor−acceptor oxygen atom separation of 4.6 Å and a nucleophile-phosphorus distance of 3.0 Å, which is inside the sum of the van der Waals radii for these two atoms (3.3 Å) (Figure 3B).[52] The donor−acceptor oxygen atom separation is larger than is observed in TSA complexes containing AlF$_4^-$ (3.9 Å; PDB 2WF6) and MgF$_3^-$ (4.3 Å; PDB 2WF5[20]) and in some DFT models of the TS for this chemical step in βPGM$_{WT}$, (4.2 Å[11]; 4.4 Å[12]). A copurified βPGM$_{D10N}$:βG16BP complex was also crystallized and the structure was determined to 2.2 Å resolution (PDB 5O6P; Figure 3D, Figure S4C and Table S1). In this structure, βG16BP is bound in the same orientation as that present in the reconstituted βPGM$_{D10N}$:P1G6P complex and the two complexes overlay closely with a non-H atom RMSD = 0.43

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

Å (Figure S6 and Table S2). The active site arrangement present in both βPGM$_{D10N}$:P1G6P complexes conforms to the definition of an aligned NAC,[23,26] where atomic distances and geometries lie close to those of TS models.[25] Given the close similarity between the complexes, the structure of the reconstituted βPGM$_{D10N}$:βG16BP complex will be used in the comparisons described below.

βPGM$_{D10N}$:P1G6P Complex Is Not Fully Closed. In contrast to all deposited metal fluoride analogue βPGM structures, the alignment of the nucleophile in the βPGM$_{D10N}$:P1G6P complex is satisfied without full closure of the enzyme (Figure 3 B,H and Table S2). Compared to the βPGM$_{WT}$:MgF$_3$:G6P TSA complex (PDB 2WF5[20]), the relative orientation of the cap and core domains undergoes a rotation of 13°, and there are significant changes in the hydrogen bonding network within the vicinity of the general acid−base residue. N10 donates a hydrogen bond to βG16BP (through atom Nδ2), while simultaneously accepting a hydrogen bond (through atom Oδ1) from the backbone amide NH and the side chain OH groups of T16. Crucially, residue T16 dictates the relative degree of closure of the cap and core domains,[8,23] and in the βPGM$_{D10N}$:P1G6P complex the conformation of T16 is near the midpoint of the transition between the substrate-free βPGM$_{WT}$ and the βPGM$_{WT}$:MgF$_3$:G6P TSA structures. The inference is that van der Waals contact between the attacking nucleophile and the 1-phosphorus atom of βG16BP in the βPGM$_{D10N}$:P1G6P complex, resists a donor−acceptor oxygen atom separation of less than 4.6 Å, the effect of which propagates through the structure to prevent the TS hydrogen bonding organization and full domain closure from being established.[11−13,20] Moreover, asymmetrical electron density for the catalytic Mg^{II} ion in the βPGM$_{D10N}$:P1G6P complex shows clear evidence of a deviation from optimal octahedral coordination geometry (Figure S7A), with elongation of distances and distortion of angles, that is not observed in metal fluoride-based ground and transition-state-analogue complexes of βPGM. This result implies that a competition exists in Mg^{II} ion coordination between the oxygen atom of the 1-phosphate group of βG16BP (O − Mg^{II} = 2.0 Å) and the carboxylate oxyanion of residue D170 (O − Mg^{II} = 2.6 Å). The equilibrium position of the Mg^{II} ion lies toward coordination by the phosphate oxygen atom, which is expected to have a higher anionic charge density, with subsequent compromising of coordination by enzymatic oxygen and oxyanion ligands. Together, these observations illustrate the interdependency between donor and acceptor atom separation, optimal hydrogen bond organization, optimal catalytic Mg^{II} ion coordination, and full domain closure to achieve TS architecture.

βPGM$_{D10N}$:AlF$_4$:G6P TSA Complex Is Fully Closed. In order to establish that the antagonism of full closure in the βPGM$_{D10N}$:P1G6P complex was not simply an artifact of the aspartate to asparagine substitution, the βPGM$_{D10N}$:AlF$_4$:G6P TSA complex was crystallized and the structure was determined to 1.1 Å resolution (PDB 5OK2; Figure 3 C,I, Figure S4D and Table S1). This complex superimposes very closely with the βPGM$_{WT}$:AlF$_4$:G6P TSA complex (PDB 2WF6; non-H atom RMSD = 0.13 Å) and it binds G6P with the 6-phosphate in the *distal* site and the square planar AlF_4^- moiety mimicking the transferring phosphoryl group in the *proximal* site between D8 (atom Oδ1) and the 1-OH group of G6P.[53] The donor−acceptor distance and angle of alignment are identical to those in the βPGM$_{WT}$:AlF$_4$:G6P TSA complex (3.8 Å and 173°, respectively). However, a comparison of the hydrogen bonding arrangements between D10/N10 and the 1-oxygen of G6P in the βPGM$_{WT}$:AlF$_4$:G6P TSA and the βPGM$_{D10N}$:AlF$_4$:G6P TSA complexes reveals a difference in the identity of the proton donor and proton acceptor. Whereas in the βPGM$_{WT}$:AlF$_4$:G6P TSA complex, the transferring proton is bonded to the 1-OH group of G6P and is coordinated by the anionic carboxylate group of the general acid−base, the analogous hydrogen bond in the βPGM$_{D10N}$:AlF$_4$:G6P TSA complex has the side chain NH_2 group of N10 coordinated by what is likely to be the deprotonated 1-oxygen of G6P. Owing to the ability of the active site to accommodate the D10 to N10 substitution, the βPGM$_{D10N}$ variant is capable of full domain closure with concomitant formation of TS geometry.

βPGM$_{D10N}$:P6G1P Complex Closely Resembles the βPGM$_{D10N}$:P1G6P Complex. While crystals harvested after 12 weeks consisted exclusively of the βPGM$_{D10N}$:P1G6P complex, a crystal with the same morphology harvested from the same drop after only 1 week yielded a 2.2 Å resolution structure of a different complex. While the resolution of the structure was limited, the electron density clearly showed that the structure contained βG16BP bound in the alternate orientation, with the 6-phosphate in the *proximal* site and the 1-phosphate in the *distal* site, and is hence termed the βPGM$_{D10N}$:P6G1P complex (PDB 5OK0; Figure 3E, Figure S4E and Table S1). Overall, the orientation of βG16BP does not have a strong influence on the degree of domain closure in the βPGM$_{D10N}$:βG16BP complexes (non-H atom RMSD = 0.34 Å). The relative orientation of the cap and core domains compared to the βPGM$_{D10N}$:AlF$_4$:G6P TSA complex have rotations of 13° (βPGM$_{D10N}$:P1G6P) and 14° (βPGM$_{D10N}$:P6G1P) (Table S2). The βPGM$_{D10N}$:P6G1P complex can again be defined as an aligned NAC (O−P−O angle = 176°, a donor−acceptor oxygen atom separation of 4.7 Å and a nucleophile-phosphorus distance of 3.1 Å) and the hydrogen bonding of residue N10 is analogous to that present in the βPGM$_{D10N}$:P1G6P complex. There is also a direct hydrogen bond present between the side chain OH group of S52 and the 3-OH group of βG16BP in the βPGM$_{D10N}$:P6G1P complex, whereas in the βPGM$_{D10N}$:P1G6P complex, hydrogen bonding between βG16BP and the protein is mediated by two water molecules (Figure S8), as observed previously in TSA complexes involving G6P and β-glucose 1-phosphonates.[19] Hence, alignment of the βG16BP intermediate is achieved in both βPGM$_{D10N}$:βG16BP complexes without full closure of the enzyme.

βPGM$_{D10N}$:AlF$_4$:H$_2$O:βG1P Complex Is Partially Open. The structure of the βPGM$_{D10N}$:AlF$_4$:βG1P complex was investigated to ascertain if it behaved analogously to the βPGM$_{D10N}$:AlF$_4$:G6P TSA complex, thus providing a direct comparator for the βPGM$_{D10N}$:P6G1P complex. The crystal structure of the βPGM$_{D10N}$:AlF$_4$:βG1P complex was determined to 1.4 Å resolution (PDB 5O6R; Figure 3F, Figure S4F and Table S1). Surprisingly, the structure did not resemble that of the fully closed βPGM$_{WT}$:AlF$_4$:G6P TSA complex (PDB 2WF6), but instead the protein atoms superimposed almost exactly with the partially open βPGM$_{D10N}$:P6G1P complex (non-H atom RMSD = 0.33 Å). Uniquely in βPGM structures, electron density consistent with a water molecule occupying an axial ligand position of the AlF_4^- moiety (instead of the 6-oxygen of βG1P) was present, with D8 still occupying the

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

other axial position, and this structure is hence termed a βPGM$_{D10N}$:AlF$_4$:H$_2$O:βG1P complex. The water molecule satisfies the demands of the AlF$_4^-$ moiety for octahedral coordination while allowing the cap domain and hydrogen bonding pattern between N10, T16 and D15 to adopt that of the βPGM$_{D10N}$:P6G1P complex. The side chain NH$_2$ group of N10 remains hydrogen bonded to the 6-OH group of βG1P rather than switching to the water molecule, despite the 6-OH group of βG1P being located further from D8 (6-OH − Oδ1 = 5.7 Å), compared with the 6-oxygen of βG16BP in the βPGM$_{D10N}$:P6G1P structure (6-O − Oδ1 = 4.6 Å). This structure implies that there is greater resistance to the formation of the fully closed βPGM$_{D10N}$:AlF$_4^-$ TSA complex with βG1P than with G6P. In contrast to the apparent deprotonation of the 1-oxygen of G6P in the βPGM$_{D10N}$:AlF$_4$:G6P TSA complex, deprotonation of the 6-OH group of βG1P appears not to be the preferred arrangement in the βPGM$_{D10N}$:AlF$_4$:βG1P complex, correlating with the ∼3 unit difference in solution pK_a values for the two hydroxyl groups.[54]

βPGM$_{D10N}$:P1G6P Complex Dominates in Solution. The crystal structures of the βPGM$_{D10N}$:βG16BP complexes with the intermediate bound in the two orientations presents a rationale for the nonequivalent complexes observed in solution using ^{31}P and ^{1}H^{15}N TROSY NMR approaches (Figure 2A and Figure S5). In the βPGM$_{D10N}$:P1G6P complex (Figure 3B), there is close proximity between H4 of βG16BP and the imidazole group of residue H20, which should result in a marked upfield chemical shift change of the H4 resonance through aromatic ring current effects. In the βPGM$_{D10N}$:P6G1P complex (Figure 3E), this chemical shift change should instead be experienced by the H3 resonance because of the change in orientation of the βG16BP ligand. To investigate the two βPGM$_{D10N}$:βG16BP complexes in solution, ^{1}H^{13}C HSQC and CCH-TOCSY spectra were acquired using 1:1 βPGM$_{D10N}$ and 100% U−^{13}C-βG16BP (Figure S2C). In both βPGM$_{D10N}$:βG16BP complexes, only the H4 resonance of βG16BP is shifted markedly upfield on binding ($\Delta\delta$ = 1.05 and 1.18 ppm), while the H3 resonance of βG16BP is shifted slightly downfield ($\Delta\delta$ = 0.08 and 0.14 ppm). Together, these results indicate that the bound orientation of βG16BP is the same in the two solution forms, thus identifying both as βPGM$_{D10N}$:P1G6P complexes. The dominance of βPGM$_{D10N}$:P1G6P over βPGM$_{D10N}$:P6G1P complexes in solution mirrors the relative dissociation constants for G6P (9 μM) and βG1P (46 μM) in the βPGM$_{WT}$:AlF$_4^-$ TSA complexes.[19]

βPGM$_{D10N}$:P1G6P Complex Has Weak Mg^{II} Affinity. The source of the difference between the two solution βPGM$_{D10N}$:P1G6P complexes was investigated using NMR backbone resonance assignment. All 210 of the nonproline residues were assigned, of which 115 showed more than one spin system. No significant structural differences were identified upon calculation of dihedral angles using TALOS-N[55] (Figure S5E,F). Residues with the largest chemical shift differences between the two complexes were principally located within the active site (Figure S5G). For ^{15}N, these comprise L9 (−2.29 ppm), V47 (−2.15 ppm), V141 (−2.78 ppm), and D170 (−2.16 ppm), for ^{13}C′, N10 (−2.69 ppm) and D170 (−1.74 ppm), for ^{13}Cα, D8 (0.81 ppm), N10 (−0.86 ppm) and S144 (−0.90 ppm), and for ^{13}Cβ, K45 (−0.80 ppm) and S171 (−0.93 ppm) (Figure S5A,B). Residues N10 and D170 are involved with the ligation of the catalytic Mg^{II} ion, suggesting that changes in this coordination may be responsible for the chemical shift differences observed. To investigate, an Mg^{II}-free form of the βPGM$_{D10N}$:P1G6P complex was prepared and the ^{1}H^{15}N-TROSY spectrum corresponded to one of the assigned βPGM$_{D10N}$:P1G6P complexes, while addition of Mg^{II} resulted in the other. Overall, the backbone chemical shift differences between the Mg^{II}-bound and Mg^{II}-free βPGM$_{D10N}$:P1G6P complexes are reminiscent of those between the βPGM$_{WT}$:MgF$_3$:G6P TSA complex (BMRB 7234[20]) and the Mg^{II}-bound βPGM$_{D10N}$:P1G6P complex in terms of the residues involved, but are smaller in magnitude (Figure S5C,D). Using changes in ^{1}H^{15}N-TROSY peak intensities on addition of Mg^{II} to the Mg^{II}-free βPGM$_{D10N}$:P1G6P complex, the dissociation constant for Mg^{II} binding was determined to be 7.1 ± 0.6 mM (Figure S7B,C), consistent with the initial purification of the βPGM$_{D10N}$:P1G6P complexes being a mixture of Mg^{II}-bound and Mg^{II}-free forms in the presence of 5 mM MgCl$_2$. In contrast, all metal fluoride analogue complexes of βPGM exist in solution as Mg^{II}-bound species at this concentration of MgCl$_2$. The changes in the ^{31}P NMR chemical shifts between the Mg^{II}-bound and Mg^{II}-free βPGM$_{D10N}$:P1G6P complexes (1-phosphate = +0.71 ppm, 6-phosphate = −0.09 ppm) are small compared with those associated with protonation of βG1P (−3.4 ppm) or G6P (−3.6 ppm) (Figure 2A and Figure S1 F−K), indicating that Mg^{II} binding is not influenced significantly by protonation of either phosphate group. Rather, the surprisingly low affinity for Mg^{II} at this point on the reaction coordinate correlates with its suboptimal coordination geometry in the structure of the βPGM$_{D10N}$:P1G6P complex (Figure S7A), in contrast to the regular Mg^{II} coordination geometry observed in the βPGM$_{D10N}$:BeF$_3$ and βPGM$_{D10N}$:AlF$_4$:G6P TSA complex structures.

DISCUSSION

A unique behavior of the βPGM$_{D10N}$ variant is that, unlike all other forms of βPGM examined to date, it copurifies as tight, noncovalently bound βPGM$_{D10N}$:βG16BP complexes. Effective removal of the bound βG16BP reaction intermediate required an unfolding-dilution-refolding approach. When challenged with substrate in the presence of excess AcP, the substrate-free βPGM$_{D10N_P}$ preparation equilibrates βG1P and G6P, with βPGM$_{D10N}^P$ maintained as the primary enzyme species. On depletion of AcP, the enzyme population shifts slowly to the βPGM$_{D10N}$:P1G6P complex becoming the dominant species. In this complex, the 1-phosphate group of βG16BP is aligned with the carboxylate oxygen atom of D8, and the side chain of N10 is shifted to the *in* position, where it forms a hydrogen bond with the bridging 1-oxygen atom of βG16BP. The enzyme is now caught in the act of phosphoryl transfer, geometrically close to the TS, but unable to complete the reaction (or at least overwhelmingly favoring the 1-phosphate group being bonded to G6P), as N10 will not release the proton hydrogen bonded to the bridging oxygen atom.

The DFT calculations of βPGM$_{WT}$ reflect enzymatic phosphoryl transfer reactions in general[25] in that the point at which proton transfer occurs is controversial. Two βPGM$_{WT}$ models predict that, when D8 attacks βG16BP in Step 2, proton transfer to βG16BP occurs prior to TS formation, and in the TS there is a donor to acceptor atom separation of 4.2 Å[11] or 4.4 Å.[12] In a third model, proton transfer is synchronous with TS formation involving a donor to acceptor atom separation of 4.0 Å,[56] while in a fourth model, proton transfer

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

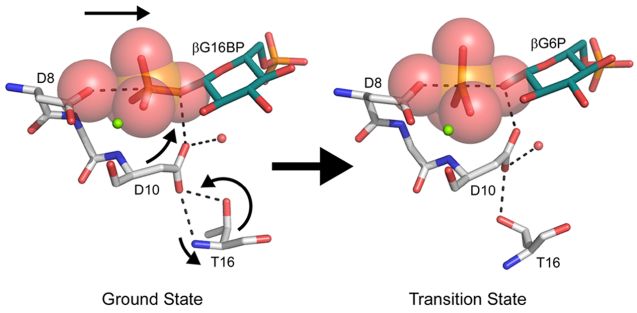


Figure 4. Schematic showing the conformational changes required for ground state to transition-state progression in *β*PGM. Despite van der Waals contact between the attacking nucleophilic carboxylate oxygen atom of D8 and the 1-phosphorus atom of *β*G16BP in the ground state βPGM_{D10N}:P1G6P complex (PDB 5OK1), the hydrogen bonding organization of the transition state is not attained. A shift in hydrogen bonding partners between T16 and D10 is required to allow positional changes in both side chains, which delivers the protonated general acid−base to the bridging oxygen atom of *β*G16BP. Following proton transfer, further compression along the donor−acceptor oxygen atom trajectory occurs, establishing the conformation of the transition state (model derived from the βPGM_{WT}:MgF_3^-:G6P TSA complex; PDB 2WF5[20]). Selected active site residues and ligands are shown as sticks in standard CPK colors, with a structural water (red) and the catalytic Mg^{II} ion (green) drawn as spheres. Large translucent spheres represent van der Waals radii for the oxygen and phosphorus atoms of the transferring phosphoryl group.

to *β*G16BP occurs after TS formation, and in the TS there is a donor to acceptor atom separation of 5.0 Å.[13] The experiment supports the predictions of the first two models, as the βPGM_{D10N}:P1G6P complex rather than the $\beta PGM_{D10N}{}^{P}$:G6P complex is trapped and, without proton transfer, the donor to acceptor atom separation is held at 4.6 Å. Intriguingly, in the 4.4 Å TS model,[12] a compression of the donor to acceptor atom separation to less than 4.6 Å is associated with the start of proton transfer from D10 to *β*G16BP. Moreover, with the donor to acceptor atom separation being held at 0.2−0.4 Å greater than that in the TS, the two domains of *β*PGM do not complete their closure. Full closure, including the hydrogen bonding of T16 and N10/D10 found in the TS, is only stable when there is compression of the reaction coordinate to below the van der Waals contact distance, as mimicked by the TSA complexes (AlF_4^- = 3.9 Å, PDB 2WF6; MgF_3^- = 4.3 Å, PDB 2WF5[20]) (Figure 4). Corroboration of the partial closure of the βPGM_{D10N} complexes is also present in the solution ensembles, where residues of the hinge in the βPGM_{D10N}:P1G6P complex lie in an intermediate position between the open and the TSA conformations, and residues D15 and T16 fail to achieve the hydrogen bond arrangement in the TS model (Figure S9). Together, these observations illustrate the complementarity between the TS and the optimal hydrogen bonding of the fully closed enzyme in the TSA conformation, as opposed to the partially open ground state *β*G16BP complex, and thus a means by which the enzyme discriminates between the TS (binding it tightly enough to have a sufficiently fast chemical step) and product (binding it weakly enough that it does not dissociate too slowly).

The rate constant for hydrolysis of the phospho-enzyme is almost unaffected by the D10N mutation. This result is readily rationalized if hydrolysis occurs with residue 10 in the *out* position, as observed for N10 in the βPGM_{D10N}:BeF_3^- complex (PDB 5OJZ) and D10 in the βPGM_{WT}:BeF_3^- complex (PDB 2WFA[23]). Previously, it had been proposed that D10 was engaged in the hydrolysis reaction of $\beta PGM_{WT}{}^{P}$ on the basis of a rate acceleration by the mutated hinge variant βPGM_{T16P}.[8] However, this mechanism is not dominant in βPGM_{WT}; the water molecule that attacks the phosphate group during hydrolysis must at least as readily transfer a proton to an ancillary base as to residue 10. The identity of the ancillary base remains to be established but the oxygen atoms of the transferring phosphoryl group (via one or more water molecules) are strong local candidates. However, the base may be another residue in βPGM_{WT} (except for residue H20[8]) or the buffer, via extended hydrogen bonded networks involving multiple water molecules.

While the $\beta PGM_{D10N}{}^{P}$ hydrolysis rate constant cannot be rationalized by a contaminant within the substrate-free βPGM_{D10N} preparation, βPGM_{D10N} is not unequivocally the source of the observed mutase activity. However, similarly to the phospho-enzyme hydrolysis reaction, it is plausible that proton transfer to the incipient hydroxyl group of G6P or *β*G1P (as the 1- or 6-phosphoryl group of *β*G16BP transfers to residue D8) is delivered from an ancillary acid by a water molecule. In a model of the βPGM_{D10N}:P1G6P complex with N10 moved to the *out* position (Figure S10), the two water molecules that occupy the space vacated by the side chain of N10 comprise part of an extended hydrogen bonded network, involving active site residues H20, K76, Y80 and the phosphate group in the *distal* site, and reaching to bulk solvent. Any one of these groups or the buffer (or even potentially the phosphate group in the *proximal* site) could act as the ancillary acid via one or more water molecules, allowing low level mutase activity to occur in βPGM_{D10N}.

Regardless of the source of the mutase activity, the replacement of D10 with N10 leads to at least a ∼ 350 fold (Figure S3I,J) reduction in activity. Consequently, the primary effect of introducing the general acid−base into βPGM_{WT} is to elevate the rate of substrate turnover to $\sim 10^3$ fold (Figure S3C,I) greater than the rate of phospho-enzyme hydrolysis, enabling the enzyme to discriminate reaction with substrate over reaction with water. This ensures that *β*PGM is primarily a mutase rather than a phosphatase.

The copurified βPGM_{D10N}:*β*G16BP complexes are present as a near-equimolar mixture of Mg^{II}-bound and Mg^{II}-free βPGM_{D10N}:P1G6P complexes in standard NMR buffer (5 mM Mg^{II}). This reflects the surprisingly low affinity of these complexes for Mg^{II} (K_d = 7.1 mM) compared with the apparent K_m = 270 *μ*M for Mg^{II} in the reaction involving βPGM_{WT},[17] and is similar to the physiological concentration of

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

Mg^{II} for *L. lactis* (~7 mM[57]). The conclusion is that βG16BP binding leads to a suboptimally coordinated catalytic Mg^{II} ion until full closure is achieved. More optimal coordination of the catalytic Mg^{II} ion is found in structures that include the 0.2–0.4 Å reduction in donor to acceptor atom separation associated with the formation of experimental TSA complexes and in DFT models of the TS. In a different class of phosphoryl transfer enzymes, the catalytic Mg^{II} ion has been identified to play a role in the rate of lid opening during the reaction cycle of adenylate kinase,[58] as well as reducing nonproductive active site fluctuations, stabilizing TS architecture, and serving as an anchor to stabilize the nucleophilic phosphate group. In βPGM, rather than acting as a pivot for opening, it appears that the catalytic Mg^{II} ion is favoring TS binding and disfavoring substrate binding by forming a looser association with its ligands as the TS relaxes to ground state complexes.

CONCLUSIONS

The employment of an aspartate to asparagine substitution of the assigned general acid–base of βPGM allowed the examination of stable enzyme:substrate complexes through the ability of βPGM_{D10N} to trap the βG16BP reaction intermediate in situ. Unlike previous structures determined for substrate, transition state, and product analogue complexes involving βG1P and G6P, the βG16BP complex achieves both alignment and contact of the attacking nucleophile with its target but without full closure of the enzyme. This reveals the interplay between compression of the reaction coordinate to below the van der Waals contact distance and the protein conformation that supports the transition state for the chemical step. The coordination of the catalytic Mg^{II} ion is an important element of this interplay on the one hand by complementing the transition state and on the other by facilitating the release of the reaction intermediate on an appropriate time scale.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.8b01612.

^{31}P, $^{1}H^{13}C$ HSQC and $^{1}H^{15}N$ TROSY NMR spectra; βPGM reaction kinetics, electron density difference and omit maps for the βPGM_{D10N} complexes; diagrams of chemical shift differences and backbone dihedral angles; superposition of the βPGM_{D10N}:P1G6P complexes; coordination and binding affinity of the catalytic Mg^{II} ion in the βPGM_{D10N}:P1G6P complex; active site coordination in the βPGM_{D10N}:βG16BP complexes; comparison of backbone amide chemical shifts in the βPGM_{D10N} complexes; model of the potentially catalytically competent form of the βPGM_{D10N}:βP1G6P complex; tables of X-ray data collection and refinement statistics; and pairwise domain rotations between the βPGM complexes (PDF)

Accession Codes

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.rcsb.org) with the following codes: βPGM_{D10N}:BeF_3 complex (PDB 5OJZ), βPGM_{D10N}:P1G6P complex (PDB 5OK1), copurified βPGM_{D10N}:P1G6P complex (PDB 5O6P), βPGM_{D10N}:P6G1P complex (PDB 5OK0), βPGM_{D10N}:AlF_4:G6P complex (PDB 5OK2) and βPGM_{D10N}:AlF_4:H_2O:βG1P complex (PDB 5O6R). The NMR chemical shifts have been deposited in the BioMagResBank (www.bmrb.wisc.edu) with the following accession numbers: Mg^{II}-bound βPGM_{D10N}:P1G6P complex (BMRB 27174) and Mg^{II}-free βPGM_{D10N}:P1G6P complex (BMRB 27175).

AUTHOR INFORMATION

Corresponding Author

*E-mail for J.P.W.: j.waltho@sheffield.ac.uk.

ORCID

Jonathan P. Waltho: 0000-0002-7402-5492

Present Addresses

[¶]L.A.J. and Y.J.: School of Chemistry, Cardiff University, Cardiff, CF10 3AT, United Kingdom

[§]C.B.: Institute of Structural and Molecular Biology, Department of Biological Sciences, Birkbeck, University of London, London, WC1E 7HX, United Kingdom

Author Contributions

[#]L.A.J. and A.J.R.: These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank Dr Tooba Alizadeh for the preparation of the βPGM_{D10N} plasmid construct and for the acquisition and interpretation of preliminary NMR experiments. We would also like to thank the beamline scientists at the Diamond Light Source (DLS) and the European Synchrotron Radiation Facility (ESRF) for the provision of synchrotron radiation facilities and assistance with data collection. This research was supported by the Biotechnology and Biological Sciences Research Council (N.J.B. – grant number: BB/M021637/1; C.T. – grant number: BB/K016245/1) and the Engineering and Physical Sciences Research Council (NMR spectrometer core capability–grant number: EP/K039547/1).

REFERENCES

(1) Hunter, T. Why Nature Chose Phosphate to Modify Proteins. *Philos. Trans. R. Soc., B* **2012**, *367*, 2513–2516.

(2) Lad, C.; Williams, N. H.; Wolfenden, R. The Rate of Hydrolysis of Phosphomonoester Dianions and the Exceptional Catalytic Proficiencies of Protein and Inositol Phosphatases. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 5607–5610.

(3) Buechler, J. A.; Taylor, S. S. Identification of Aspartate-184 as an Essential Residue in the Catalytic Subunit of cAMP-Dependent Protein Kinase. *Biochemistry* **1988**, *27*, 7356–7361.

(4) Green, P. C.; Tripathi, R. L.; Kemp, R. G. Identification of Active Site Residues in Pyrophosphate-Dependent Phosphofructo-1-Kinase by Site-Directed Mutagenesis. *J. Biol. Chem.* **1993**, *268*, 5085–5088.

(5) Denu, J. M.; Lohse, D. L.; Vijayalakshmi, J.; Saper, M. A.; Dixon, J. E. Visualization of Intermediate and Transition-State Structures in Protein-Tyrosine Phosphatase Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 2493–2498.

(6) Wu, L.; Zhang, Z.-Y. Probing the Function of Asp128 in the Lower Molecular Weight Protein-Tyrosine Phosphatase-Catalyzed Reaction. A Pre-Steady-State and Steady-State Kinetic Investigation. *Biochemistry* **1996**, *35*, 5426–5434.

(7) Skamnaki, V. T.; Owen, D. J.; Noble, M. E. M.; Lowe, E. D.; Lowe, G.; Oikonomakos, N. G.; Johnson, L. N. Catalytic Mechanism of Phosphorylase Kinase Probed by Mutational Studies. *Biochemistry* **1999**, *38*, 14718–14730.

(8) Dai, J.; Finci, L.; Zhang, C.; Lahiri, S.; Zhang, G.; Peisach, E.; Allen, K. N.; Dunaway-Mariano, D. Analysis of the Structural

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140–8153

Determinants Underlying Discrimination Between Substrate and Solvent in β-Phosphoglucomutase Catalysis. *Biochemistry* **2009**, *48*, 1984−1995.
(9) Valiev, M.; Kawai, R.; Adams, J. A.; Weare, J. H. The Role of the Putative Catalytic Base in the Phosphoryl Transfer Reaction in a Protein Kinase: First-Principles Calculations. *J. Am. Chem. Soc.* **2003**, *125*, 9926−9927.
(10) Asthagiri, D.; Liu, T.; Noodleman, L.; Van Etten, R. L.; Bashford, D. On the Role of the Conserved Aspartate in the Hydrolysis of the Phosphocysteine Intermediate of the Low Molecular Weight Tyrosine Phosphatase. *J. Am. Chem. Soc.* **2004**, *126*, 12677−12684.
(11) Webster, C. E. High-Energy Intermediate or Stable Transition State Analogue: Theoretical Perspective of the Active Site and Mechanism of β-Phosphoglucomutase. *J. Am. Chem. Soc.* **2004**, *126*, 6840−6841.
(12) Marcos, E.; Field, M. J.; Crehuet, R. Pentacoordinated Phosphorus Revisited by High-Level QM/MM Calculations. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 2405−2411.
(13) Elsässer, B.; Dohmeier-Fischer, S.; Fels, G. Theoretical Investigation of the Enzymatic Phosphoryl Transfer of β-Phosphoglucomutase: Revisiting Both Steps of the Catalytic Cycle. *J. Mol. Model.* **2012**, *18*, 3169−3179.
(14) Kim, K.; Cole, P. A. Measurement of a Brønsted Nucleophile Coefficient and Insights into the Transition State for a Protein Tyrosine Kinase. *J. Am. Chem. Soc.* **1997**, *119*, 11096−11097.
(15) Zhou, J.; Adams, J. A. Is There a Catalytic Base in the Active Site of cAMP-Dependent Protein Kinase? *Biochemistry* **1997**, *36*, 2977−2984.
(16) Kim, K.; Cole, P. A. Kinetic Analysis of a Protein Tyrosine Kinase Reaction Transition State in the Forward and Reverse Directions. *J. Am. Chem. Soc.* **1998**, *120*, 6851−6858.
(17) Zhang, G.; Dai, J.; Wang, L.; Dunaway-Mariano, D.; Tremblay, L. W.; Allen, K. N. Catalytic Cycling in β-Phosphoglucomutase: A Kinetic and Structural Analysis. *Biochemistry* **2005**, *44*, 9404−9416.
(18) Dai, J.; Wang, L.; Allen, K. N.; Radstrom, P.; Dunaway-Mariano, D. Conformational Cycling in β-Phosphoglucomutase Catalysis: Reorientation of the β-D-Glucose 1,6-(Bis)phosphate Intermediate. *Biochemistry* **2006**, *45*, 7818−7824.
(19) Jin, Y.; Bhattasali, D.; Pellegrini, E.; Forget, S. M.; Baxter, N. J.; Cliff, M. J.; Bowler, M. W.; Jakeman, D. L.; Blackburn, G. M.; Waltho, J. P. α-Fluorophosphonates Reveal How a Phosphomutase Conserves Transition State Conformation Over Hexose Recognition in Its Two-Step Reaction. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 12384−12389.
(20) 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.; Waltho, J. P. Atomic Details of Near-Transition State Conformers for Enzyme Phosphoryl Transfer Revealed by MgF_3^- Rather Than by Phosphoranes. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 4555−4560.
(21) Baxter, N. J.; Olguin, L. F.; Goličnik, M.; Feng, G.; Hounslow, A. M.; Bermel, W.; Blackburn, G. M.; Hollfelder, F.; Waltho, J. P.; Williams, N. H. A Trojan Horse Transition State Analogue Generated by MgF_3^- Formation in an Enzyme Active Site. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 14732−14737.
(22) Baxter, N. J.; Hounslow, A. M.; Bowler, M. W.; Williams, N. H.; Blackburn, G. M.; Waltho, J. P. MgF_3^- and α-Galactose 1-Phosphate in the Active Site of β-Phosphoglucomutase Form a Transition State Analogue of Phosphoryl Transfer. *J. Am. Chem. Soc.* **2009**, *131*, 16334−16335.
(23) 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.; Waltho, J. P. Near Attack Conformers Dominate β-Phosphoglucomutase Complexes Where Geometry and Charge Distribution Reflect Those of Substrate. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6910−6915.
(24) Jin, Y.; Richards, N. G. J.; Waltho, J. P.; Blackburn, G. M. Metal Fluorides as Analogues for Studies on Phosphoryl Transfer Enzymes. *Angew. Chem., Int. Ed.* **2017**, *56*, 4110−4128.
(25) Jin, Y.; Molt, R. W.; Blackburn, G. M. Metal Fluorides: Tools for Structural and Computational Analysis of Phosphoryl Transfer Enzymes. *Top. Curr. Chem.* **2017**, *375*, 36−59.
(26) Bruice, T. C. Some Pertinent Aspects of Mechanism as Determined With Small Molecules. *Annu. Rev. Biochem.* **1976**, *45*, 331−374.
(27) Jin, Y.; Molt, R. W.; Waltho, J. P.; Richards, N. G. J.; Blackburn, G. M. ^{19}F NMR and DFT Analysis Reveal Structural and Electronic Transition State Features for RhoA-Catalyzed GTP Hydrolysis. *Angew. Chem., Int. Ed.* **2016**, *55*, 3318−3322.
(28) Jin, Y.; Molt, R. W.; Pellegrini, E.; Cliff, M. J.; Bowler, M. W.; Richards, N. G. J.; Blackburn, G. M.; Waltho, J. P. Assessing the Influence of Mutation on GTPase Transition States Using X-Ray Crystallography, ^{19}F NMR and DFT Approaches. *Angew. Chem., Int. Ed.* **2017**, *56*, 9732−9735.
(29) Reed, M. A. C.; Hounslow, A. M.; Sze, K. H.; Barsukov, I. G.; Hosszu, L. L. P.; Clarke, A. R.; Craven, C. J.; Waltho, J. P. Effects of Domain Dissection on the Folding and Stability of the 43 kDa Protein PGK Probed by NMR. *J. Mol. Biol.* **2003**, *330*, 1189−1201.
(30) Goličnik, M.; Olguin, L. F.; Feng, G.; Baxter, N. J.; Waltho, J. P.; Williams, N. H.; Hollfelder, F. Kinetic Analysis of β-Phosphoglucomutase and Its Inhibition by Magnesium Fluoride. *J. Am. Chem. Soc.* **2009**, *131*, 1575−1588.
(31) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins: Struct., Funct., Genet.* **2005**, *59*, 687−696.
(32) Schulte-Herbrüggen, T.; Sørensen, O. W. Clean TROSY: Compensation for Relaxation-Induced Artefacts. *J. Magn. Reson.* **2000**, *144*, 123−128.
(33) Lescop, E.; Schanda, P.; Brutscher, B. A Set of BEST Triple-Resonance Experiments for Time-Optimized Protein Resonance Assignment. *J. Magn. Reson.* **2007**, *187*, 163−169.
(34) Hyberts, S. G.; Robson, S. A.; Wagner, G. Exploring Signal-to-Noise Ratio and Sensitivity in Non-Uniformly Sampled Multi-Dimensional NMR Spectra. *J. Biomol. NMR* **2013**, *55*, 167−178.
(35) Hyberts, S. G.; Milbradt, A. G.; Wagner, A. B.; Arthanari, H.; Wagner, G. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-Uniformly Sampled With Multidimensional Poisson Gap Scheduling. *J. Biomol. NMR* **2012**, *52*, 315−327.
(36) Sun, Z.-Y. J.; Frueh, D. P.; Selenko, P.; Hoch, J. C.; Wagner, G. Fast Assignment of ^{15}N-HSQC Peaks Using High-Resolution 3D HNcocaNH Experiments With Non-Uniform Sampling. *J. Biomol. NMR* **2005**, *33*, 43−50.
(37) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1−16.
(38) Pellegrini, E.; Piano, D.; Bowler, M. W. Direct Cryocooling of Naked Crystals: Are Cryoprotection Agents Always Necessary? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *D67*, 902−906.
(39) Winter, G. *xia2*: An Expert System for Macromolecular Crystallography Data Reduction. *J. Appl. Crystallogr.* **2010**, *43*, 186−190.
(40) Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. iMOSFLM: A New Graphical Interface for Diffraction-Image Processing With MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *D67*, 271−281.
(41) Vagin, A.; Teplyakov, A. MOLREP: An Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **1997**, *30*, 1022−1025.
(42) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of COOT. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 486−501.
(43) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *D53*, 240−255.
(44) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

Macromolecular Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 12−21.
(45) *PyMOL Molecular Graphics System*, Version 1.8.; Schrödinger, LLC, 2015.
(46) Read, R. J.; Schierbeek, A. J. A Phased Translation Function. *J. Appl. Crystallogr.* **1988**, *21*, 490−495.
(47) Hayward, S.; Berendsen, H. J. C. Systematic Analysis of Domain Motions in Proteins From Conformational Change: New Results on Citrate Synthase and T4 Lysozyme. *Proteins: Struct., Funct., Genet.* **1998**, *30*, 144−154.
(48) Lahiri, S. D.; Zhang, G.; Radstrom, P.; Dunaway-Mariano, D.; Allen, K. N. Crystallization and Preliminary X-Ray Diffraction Studies of β-Phosphoglucomutase From *Lactococcus lactis*. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *D58*, 324−326.
(49) Lahiri, S. D.; Zhang, G.; Dunaway-Mariano, D.; Allen, K. N. Caught in the Act: The Structure of Phosphorylated β-Phosphoglucomutase From *Lactococcus lactis*. *Biochemistry* **2002**, *41*, 8351−8359.
(50) Lahiri, S. D.; Zhang, G.; Dunaway-Mariano, D.; Allen, K. N. The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction. *Science* **2003**, *299*, 2067−2071.
(51) Tyler-Cross, R.; Schirch, V. Effects of Amino Acid Sequence, Buffers, and Ionic Strength on the Rate and Mechanism of Deamidation of Asparagine Residues in Small Peptides. *J. Biol. Chem.* **1991**, *266*, 22549−22556.
(52) Rowland, R. S.; Taylor, R. Intermolecular Nonbonded Contact Distances in Organic Crystal Structures: Comparison With Distances Expected From van der Waals Radii. *J. Phys. Chem.* **1996**, *100*, 7384−7391.
(53) 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.; Waltho, J. P. Anionic Charge Is Prioritized Over Geometry in Aluminum and Magnesium Fluoride Transition State Analogs of Phosphoryl Transfer Enzymes. *J. Am. Chem. Soc.* **2008**, *130*, 3952−3958.
(54) Woolley, E. M.; Tomkins, J.; Hepler, L. G. Ionization Constants for Very Weak Organic Acids in Aqueous Solution and Apparent Ionization Constants for Water in Aqueous Organic Mixtures. *J. Solution Chem.* **1972**, *1*, 341−351.
(55) Shen, Y.; Bax, A. Protein Backbone and Sidechain Torsion Angles Predicted From NMR Chemical Shifts Using Artificial Neural Networks. *J. Biomol. NMR* **2013**, *56*, 227−241.
(56) Barrozo, A.; Liao, Q.; Esguerra, M.; Marloie, G.; Florian, J.; Williams, N. H.; Kamerlin, S. C. L. Computer Simulations of the Catalytic Mechanism of Wild-Type and Mutant β-Phosphoglucomutase. *Org. Biomol. Chem.* **2018**, *16*, 2060−2073.
(57) Zamberlin, Š.; Antunac, N.; Havranek, J.; Samaržija, D. Mineral Elements in Milk and Dairy Products. *Mljekarstvo* **2012**, *62*, 111−125.
(58) Kerns, S. J.; Agafonov, R. V.; Cho, Y.-J.; Pontiggia, F.; Otten, R.; Pachov, D. V.; Kutter, S.; Phung, L. A.; Murphy, P. N.; Thai, V.; Alber, T.; Hagan, M. F.; Kern, D. The Energy Landscape of Adenylate Kinase During Catalysis. *Nat. Struct. Mol. Biol.* **2015**, *22*, 124−131.

DOI: 10.1021/acscatal.8b01612
ACS Catal. 2018, 8, 8140−8153

SUPPORTING INFORMATION

van der Waals Contact between Nucleophile and Transferring Phosphorus Is Insufficient To Achieve Enzyme Transition-State Architecture

Luke A. Johnson,[†,¶,#] Angus J. Robertson,[†,#] Nicola J. Baxter,[†,‡] Clare R. Trevitt,[†] Claudine Bisson,[†,§] Yi Jin,[†,¶] Henry P. Wood,[†] Andrea M. Hounslow,[†] Matthew J. Cliff,[‡] G. Michael Blackburn,[†] Matthew W. Bowler,[‖] and Jonathan P. Waltho,[*,†,‡]

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

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

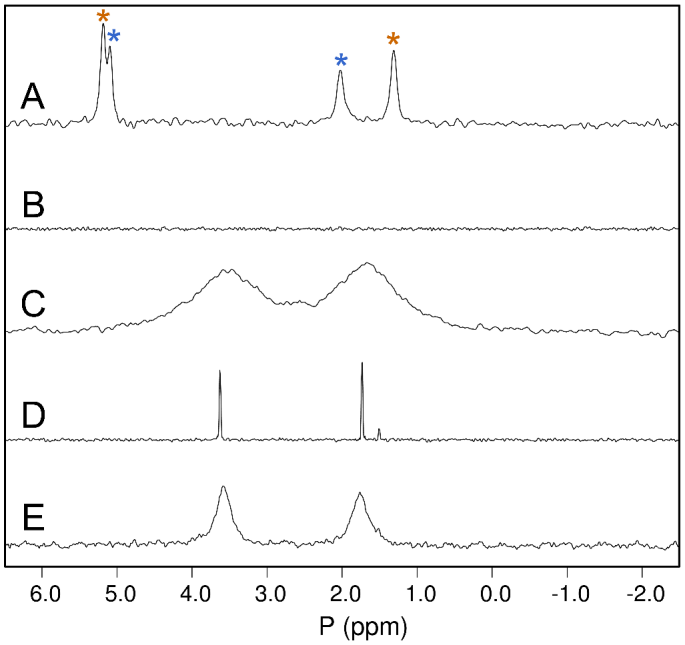
[‖] European Molecular Biology Laboratory, Grenoble Outstation, 71 avenue des Martyrs, CS 90181 F-38042 Grenoble, France

[#] (L.A.J. and A.J.R.) These authors contributed equally

[¶] (L.A.J. and Y.J.) School of Chemistry, Cardiff University, Cardiff, CF10 3AT, United Kingdom

[§] (C.B.) Institute of Structural and Molecular Biology, Department of Biological Sciences, Birkbeck, University of London, London, WC1E 7HX, United Kingdom

* E-mail for J.P.W.: j.waltho@sheffield.ac.uk


A
B
C
D
E
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
P (ppm)

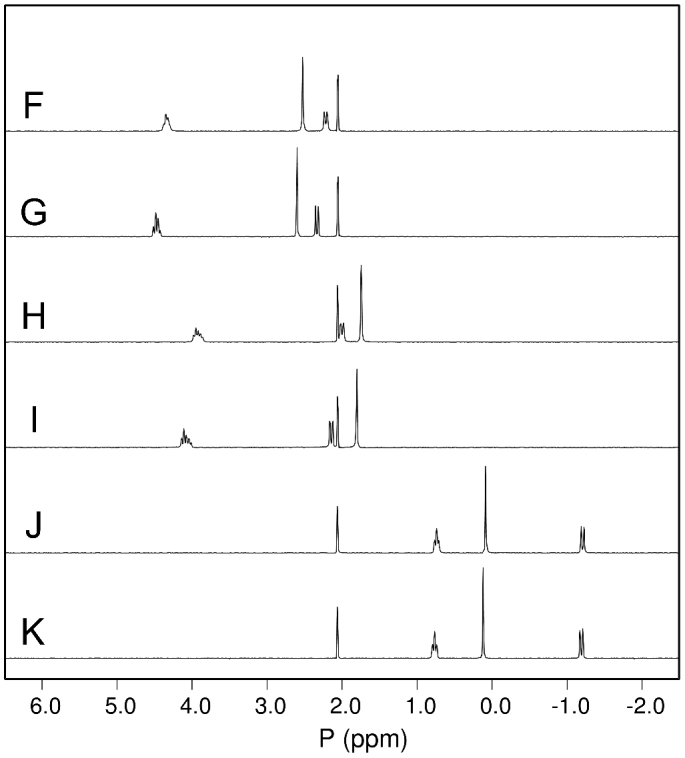

F
G
H
I
J
K
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
P (ppm)

Figure S1. ^{31}P NMR spectra illustrating steps in the purification of βPGM_{D10N} and βG16BP, together with the dependence of G6P and βG1P chemical shifts on pH and Mg^{II} concentration. (A – E) Samples were prepared in standard NMR buffer and spectra were acquired typically with 10000 transients over 50 ppm using proton-phosphorus decoupling. (A) βPGM_{D10N} immediately following purification showing four ^{31}P resonances consistent with the population of two tight noncovalently bound βPGM_{D10N}:βG16BP complexes. The orange asterisks indicate ^{31}P peaks from the 6-phosphate (5.17 ppm) and the 1-phosphate (1.30 ppm) groups of βG16BP in the Mg^{II}-bound βPGM_{D10N}:P1G6P complex and the blue asterisks indicate ^{31}P peaks from the 6-phosphate (5.08 ppm) and the 1-phosphate (2.01 ppm) groups of βG16BP in the Mg^{II}-free βPGM_{D10N}:P1G6P complex. (B) Substrate-free βPGM_{D10N} generated by unfolding the βPGM_{D10N}:βG16BP complexes in 4 M guanidine hydrochloride, with 200-fold dilution of the ligand using buffer exchange, and subsequent refolding of βPGM_{D10N} to a native conformation. The absence of ^{31}P resonances indicates that βG16BP no longer occupies the active site. (C) βG16BP extracted by membrane filtration from heat denatured (2 min at 80 °C) βPGM_{D10N}:βG16BP complexes. The two ^{31}P resonances are broadened significantly due to exchange of coordination between the phosphate groups of βG16BP and Mg^{II} ions present in the sample. (D) Addition of 6 mM EDTA to the sample in (C) chelates the Mg^{II} ions resulting in a significant narrowing of linewidths for the two ^{31}P peaks (3.63 and 1.74 ppm). This sample was used to record the $^{1}H^{13}C$ HSQC spectrum shown in Figure S2B. (E) Chemically synthesized βG16BP (Prof. Nicholas Williams, Department of Chemistry, The University of Sheffield) in standard NMR buffer. Correspondence in chemical shift values between the two ^{31}P resonances (C – E) is consistent with βG16BP being isolated from the βPGM_{D10N}:βG16BP complexes. (F – K) Samples contained 10 mM G6P, 10 mM βG1P and 20 mM sodium phosphate in 10 mM Tris and 10 mM sodium acetate buffer at (F and G) pH 9.0, (H and I) pH 7.0 and (J and K) pH 4.0, containing either (F, H, J) 10 mM $MgCl_2$ or (G, I, K) no Mg^{II}. A capillary containing 200 mM sodium phosphate at pH* 7.2 in 100% $^{2}H_2O$ was included in the sample as a chemical shift reference (2.06 ppm) and for the deuterium lock. Other resonances are assigned as follows: G6P (left hand multiplet), inorganic phosphate (singlet) and βG1P (right hand doublet). Spectra were acquired with 256 transients over 50 ppm and without proton-phosphorus decoupling to differentiate the G6P and βG1P resonances.

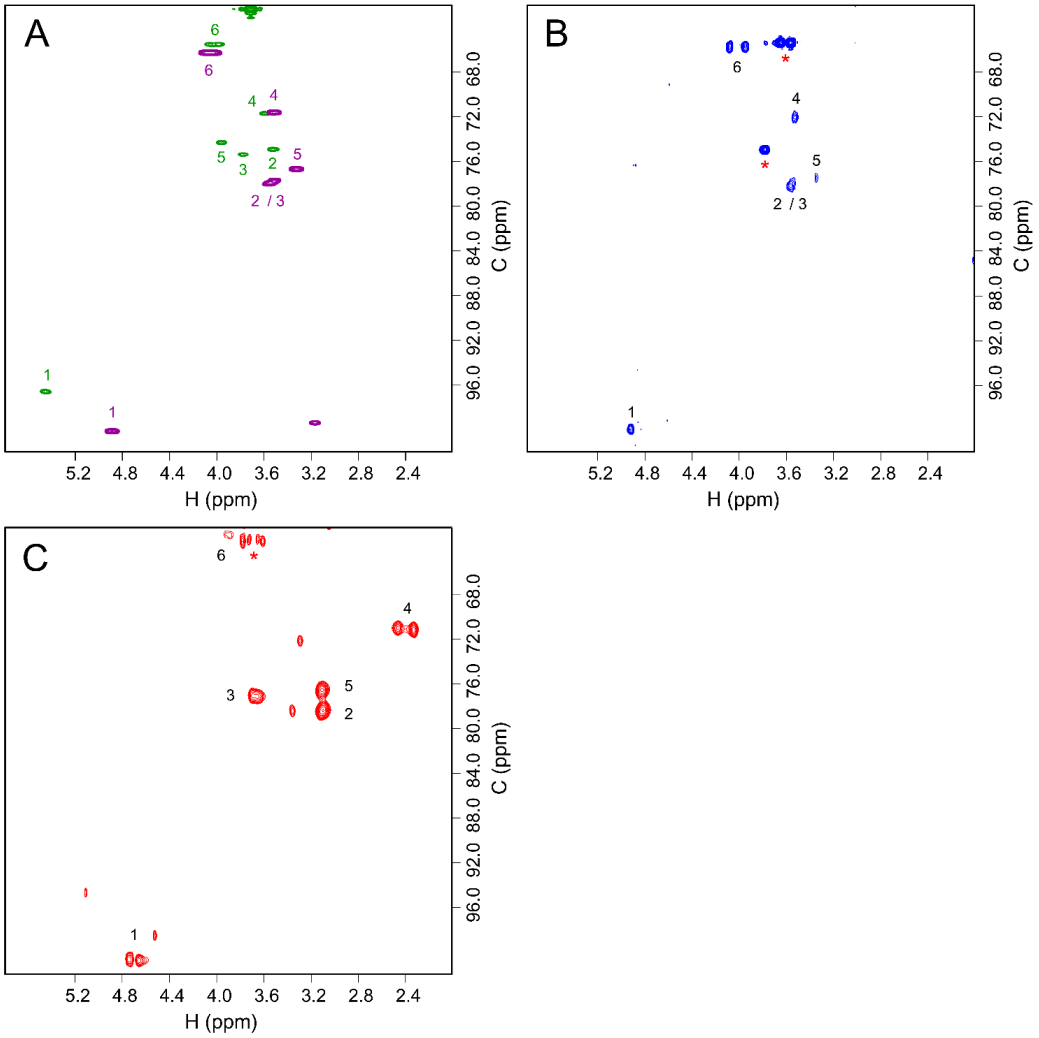
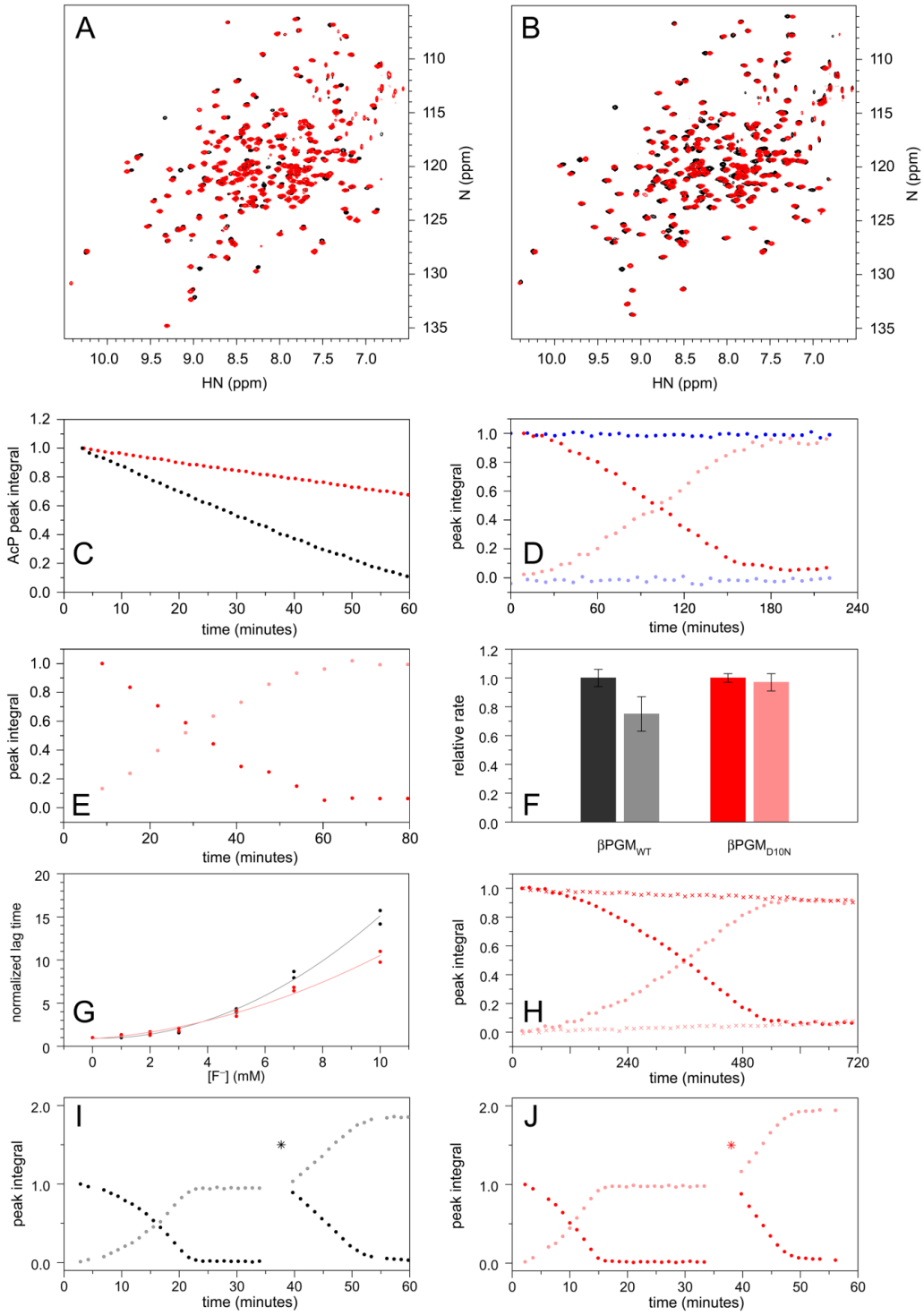


Figure S2. Assigned $^{1}H^{13}C$ HSQC spectra of glucose 1,6-bisphosphate species. (A) Commercially produced αG16BP (Sigma) in 100% $^{2}H_2O$ (green) and chemically synthesized βG16BP in 100% $^{2}H_2O$ (magenta). (B) βG16BP extracted by membrane filtration from heat denatured (2 min at 80 °C) $βPGM_{D10N}$:βG16BP complexes in standard NMR buffer containing 6 mM EDTA. The red asterisks denote peaks arising from the buffer. (C) Uniformly ^{13}C-labeled βG16BP in the Mg^{II}-bound $βPGM_{D10N}$:P1G6P and Mg^{II}-free $βPGM_{D10N}$:P1G6P complexes in standard NMR buffer.


A
B
N (ppm)
HN (ppm)
C
AcP peak integral
time (minutes)
D
peak integral
E
F
relative rate
βPGM$_{WT}$
βPGM$_{D10N}$
G
normalized lag time
[F$^-$] (mM)
H
I
J

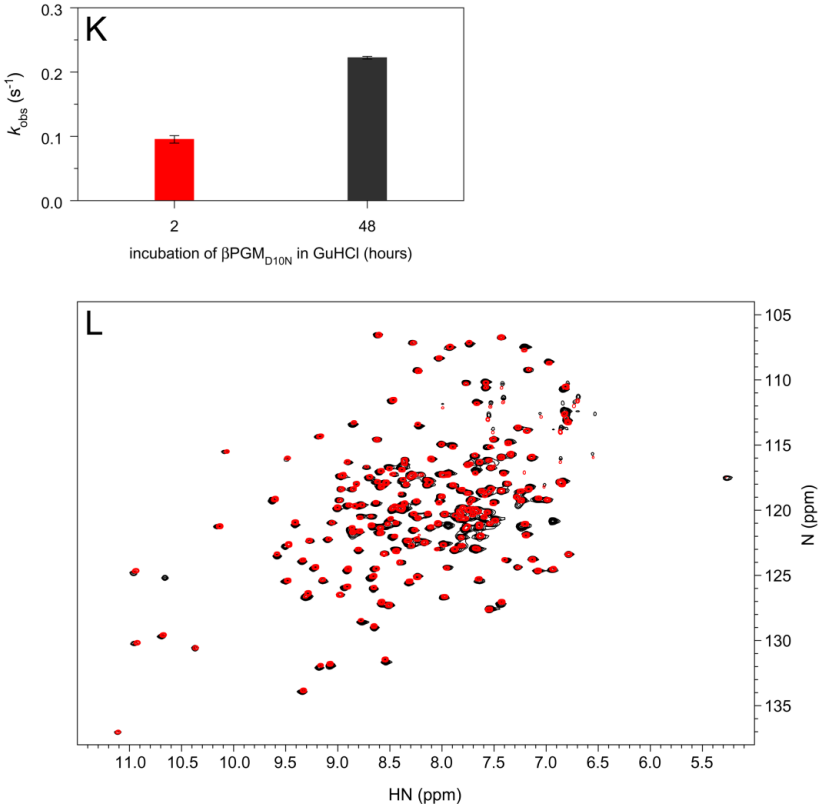


Figure S3. Comparison of $^{1}H^{15}N$ TROSY spectra, acetyl phosphate (AcP) hydrolysis kinetics, inhibition of βPGM$_{WT}$ and βPGM$_{D10N}$ by inorganic phosphate and fluoride, and βG1P equilibration by βPGM$_{D10N}$, βPGM$_{D8N}$ and the reconstituted βPGM$_{D10N}$:βG16BP complexes. (A) Superposed $^{1}H^{15}N$ TROSY spectra of (black) βPGM$_{WT}$ and (red) substrate-free βPGM$_{D10N}$, generated by unfolding the copurified βPGM$_{D10N}$:βG16BP complexes in 4 M guanidine hydrochloride, with 200-fold dilution of the ligand using buffer exchange, and subsequent refolding to a native conformation. Samples typically contained either 0.5 mM βPGM$_{WT}$ or 0.5 mM substrate-free βPGM$_{D10N}$ in standard NMR buffer with 50 mM $MgCl_2$. (B) Superposed $^{1}H^{15}N$ TROSY spectra of (black) βPGM$_{WT}$:BeF_3 complex and (red) βPGM$_{D10N}$:BeF_3 complex. The complexes containing the BeF_3^- moiety coordinated at D8 are structural mimics of βPGM$_{WT}^{P}$ and βPGM$_{D10N}^{P}$ and were generated from samples typically containing either 0.5 mM βPGM$_{WT}$ or 0.5 mM substrate-free βPGM$_{D10N}$ in standard NMR buffer with 5 mM $BeCl_2$ and 10 mM NH_4F. (C) Hydrolysis kinetics of AcP to inorganic phosphate and acetate

by (black) βPGM_{WT} and (red) substrate-free βPGM_{D10N}, monitored by ^{31}P NMR spectra with integration of the AcP peak as a function of time. The samples contained either 250 μM βPGM_{WT} or 250 μM substrate-free βPGM_{D10N}, together with 50 mM AcP in standard kinetic buffer containing 50 mM $MgCl_2$ and 1 mM EDTA. Hydrolysis rate constants were obtained from linear fitting of the data (βPGM_{WT} = 0.06 ± 0.006 s^{-1}; βPGM_{D10N} = 0.02 ± 0.002 s^{-1}). A control experiment involving 50 mM AcP alone in standard kinetic buffer established that hydrolysis of AcP was insignificant over the same time frame. (D) Reaction kinetics of (red / pink) substrate-free βPGM_{D10N} (5 μM) and (blue / light blue) βPGM_{D8N} (10 μM) for the equilibration of 10 mM βG1P with G6P in the presence of 20 mM AcP monitored by ^{31}P NMR spectra using normalized integral values of (red / blue) the βG1P peak and (pink / light blue) the G6P peak as a function of time. (E) Reaction kinetics of substrate-free βPGM_{D10N} (5 μM) for the equilibration of 10 mM βG1P with G6P in the presence of βG16BP (extracted from the copurified βPGM_{D10N}:βG16BP complexes) monitored by ^{31}P NMR spectra using normalized integral values of (red) the βG1P peak and (pink) the G6P peak as a function of time. (F) Relative reaction rates monitored by ^{31}P NMR spectra of βPGM_{WT} (0.1 – 0.25 μM; n = 3) and substrate-free βPGM_{D10N} (45 μM; n = 3) for the equilibration of 10 mM βG1P with G6P in the presence of 20 mM AcP in (dark gray / red) the standard kinetic buffer and (light gray / pink) with the addition of 20 mM sodium phosphate. (G) Reaction kinetics monitored by the glucose 6-phosphate dehydrogenase coupled assay of βPGM_{WT} (5 nM) and substrate-free βPGM_{D10N} (500 nM) for the equilibration of 10 mM βG1P with G6P in the presence of 20 mM AcP in the standard kinetic buffer with increasing concentrations of fluoride (0, 1, 2, 3, 5, 7 and 10 mM). Time points corresponding to the end of the lag phase (as measured by first derivative analysis) for each of the fluoride concentrations were normalized against data recorded in the absence of fluoride. (H) Reaction kinetics of the reconstituted βPGM_{D10N}:βG16BP complexes (2.5 μM) for the equilibration of 10 mM βG1P with G6P in the (crosses) absence and (circles) presence of 20 mM AcP monitored by ^{31}P NMR spectra using normalized integral values of (red) the βG1P peak and (pink) the G6P peak as a function of time. (I and J) Reaction kinetics of (black / gray) βPGM_{WT} (0.25 μM) and (red / pink) substrate-free βPGM_{D10N} (45 μM) for the equilibration of 10 mM βG1P with G6P in the presence of 20 mM AcP monitored by ^{31}P NMR spectra using normalized integral values of the βG1P peak (black / red) and G6P peak (gray / pink) as a function of time. Asterisks denote the time points at which samples were recharged with additional 10 mM βG1P. Missing ^{31}P data at ca. 5 and 55 minutes in the time courses is to allow for the acquisition of ^{1}H NMR spectra. A k_{cat} of 0.2 ± 0.08 s^{-1} (n = 8) was derived for βPGM_{D10N} from the linear segment of the first kinetic profile, compared with 70 ± 30 s^{-1} (n = 7)

for βPGM_{WT}. Note that for the reaction kinetics monitored by ^{31}P NMR spectra, the enzyme concentration was adjusted to allow for similar signal-to-noise ratios to be obtained on the different spectrometers used. (K) Observed catalytic rate constants (k_{obs}) monitored by the glucose 6-phosphate dehydrogenase coupled assay of βPGM_{D10N} (500 nM) for the equilibration of 230 μM βG1P with G6P in the presence of 10 mM AcP for (red) substrate-free βPGM_{D10N} following ca. 2 h incubation with 4 M guanidine hydrochloride and (black) substrate-free βPGM_{D10N} following ca. 48 h incubation with 4 M guanidine hydrochloride in the unfolding-dilution-refolding procedure. (L) Superposed $^{1}H^{15}N$ TROSY spectra of the reconstituted Mg^{II}-bound βPGM_{D10N}:P1G6P complex in standard NMR buffer containing 50 mM $MgCl_2$, 20 mM AcP and 10 mM G6P with (red) substrate-free βPGM_{D10N} following ca. 2 h incubation with 4 M guanidine hydrochloride and (black) substrate-free βPGM_{D10N} following ca. 48 h incubation with 4 M guanidine hydrochloride in the unfolding-dilution-refolding procedure.

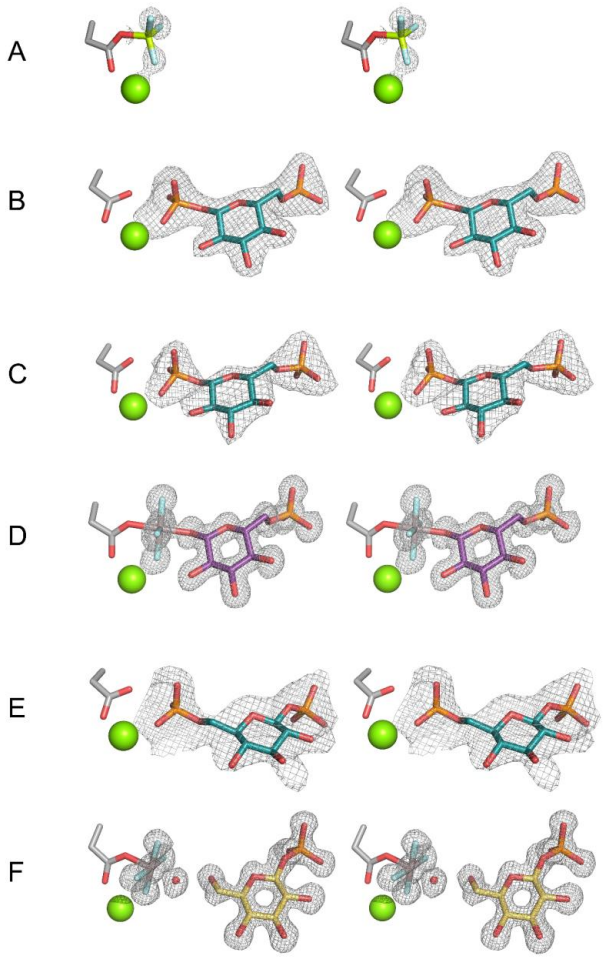
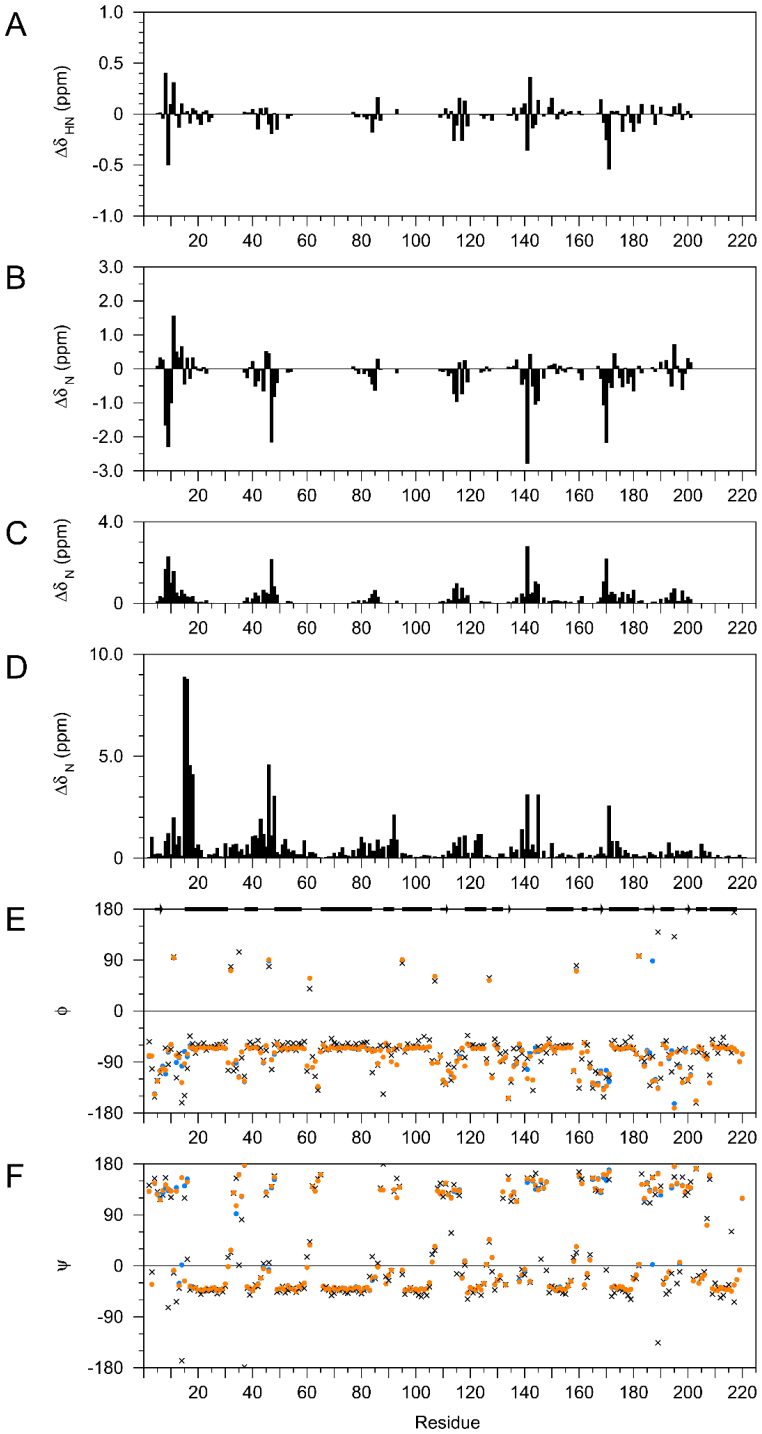


Figure S4. Stereoviews of difference density (Fo – Fc) for the βPGM_{D10N} complexes. The active sites of (A) βPGM_{D10N}:BeF_3 complex (PDB 5OJZ), (B) βPGM_{D10N}:P1G6P complex (PDB 5OK1), (C) copurified βPGM_{D10N}:P1G6P complex (PDB 5O6P), (D) βPGM_{D10N}:AlF_4:G6P complex (PDB 5OK2), (E) βPGM_{D10N}:P6G1P complex (PDB 5OK0) and (F) βPGM_{D10N}:AlF_4:H_2O:βG1P complex (PDB 5O6R). The side chain of D8 and active site ligands are shown as sticks in standard CPK colors, with beryllium (light green), fluorine (light blue), aluminum (dark gray), βG16BP (teal carbon atoms), G6P (purple carbon atoms) and βG1P (gold carbon atoms). An axially coordinated water (red) and the catalytic Mg^{II} ion (green) are drawn as spheres. Difference density (Fo – Fc; gray mesh) was generated following ligand omission from the final structures, and is contoured selectively at 2.5σ (E) and 3σ (A–D, F) for the BeF_3^-, βG16BP, AlF_4^-, G6P, βG1P and water ligands.


A
B
C
D
E
F
$\Delta\delta_{HN}$ (ppm)
$\Delta\delta_N$ (ppm)
ϕ
ψ
Residue

G

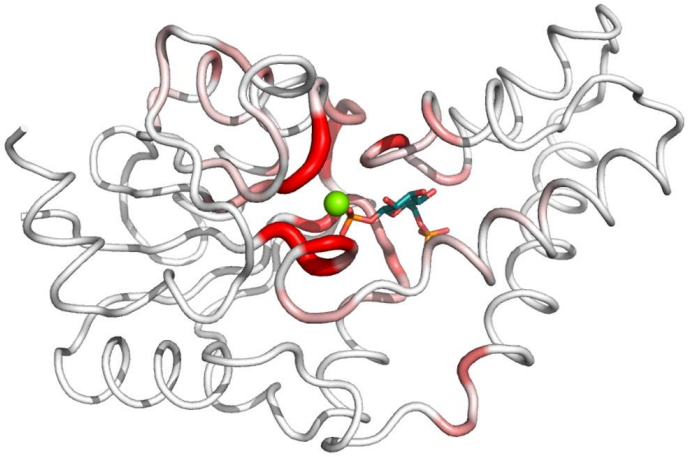


Figure S5. Chemical shift analysis of the Mg^{II}-bound βPGM_{D10N}:P1G6P and the Mg^{II}-free βPGM_{D10N}:P1G6P complexes. Histograms of residue specific chemical shift changes for the βPGM_{D10N}:P1G6P complexes calculated as $\Delta\delta = \delta_{Mg\text{-}bound} - \delta_{Mg\text{-}free}$ for (A) backbone H_N atoms, (B) backbone N atoms and (C) as $\Delta\delta = [(\delta_{Mg\text{-}bound} - \delta_{Mg\text{-}free})^2]^{1/2}$ for the backbone N atoms. (D) Residue specific chemical shift changes between the Mg^{II}-bound βPGM_{D10N}:P1G6P complex and the βPGM_{WT}:MgF_3:G6P TSA complex (BMRB 7234)[20] calculated as $\Delta\delta = [(\delta_{\beta PGM\text{-}D10N\text{-}P1G6P} - \delta_{\beta PGM\text{-}WT\text{-}TSA})^2]^{1/2}$ for the backbone N atoms. The data have been plotted with the same vertical scaling as (C) so that the size of $\Delta\delta_N$ can be compared. (E and F) Backbone dihedral angle prediction of βPGM_{D10N} in the Mg^{II}-bound βPGM_{D10N}:P1G6P complex (orange circles) and the Mg^{II}-free βPGM_{D10N}:P1G6P complex (blue circles) obtained with TALOS-N[55] using the backbone ^1H_N, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$ and $^{13}C'$ chemical shifts. For comparison, backbone dihedral angles were extracted from the βPGM_{D10N}:P1G6P crystal structure (PDB 5OK1) and are shown as black crosses. Secondary structure elements from βPGM_{WT} (PDB 2WHE)[20] are indicated by bars (α-helices) and arrows (β-strands) at the top of the panel. (G) Structure of the βPGM_{D10N}:P1G6P complex (PDB 5OK1) with residues colored according chemical shift changes calculated as $\Delta\delta = [\Delta\delta_{HN}^2 + (0.12 \times \Delta\delta_N)^2]^{1/2}$, between the Mg^{II}-bound βPGM_{D10N}:P1G6P complex and the Mg^{II}-free βPGM_{D10N}:P1G6P complex, with the intensity of color and thickness of the backbone corresponding to larger $\Delta\delta$ values. The βG16BP ligand is shown as CPK-colored sticks and the catalytic Mg^{II} ion is indicated as a green sphere.

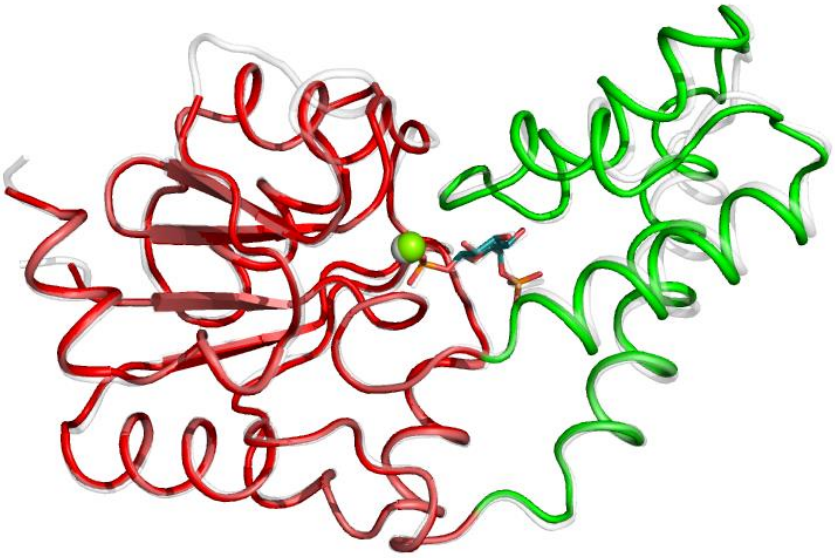


Figure S6. Comparison of the two crystal structures of the βPGM_{D10N}:P1G6P complex. The reconstituted βPGM_{D10N}:P1G6P complex (PDB 5OK1; pale gray ribbon) and the copurified βPGM_{D10N}:P1G6P complex (PDB 5O6P; red and green ribbon) have been superposed on the core domains (left). The βG16BP ligands are drawn as sticks (in CPK colors for PDB 5O6P) and the catalytic Mg^{II} ions are shown as spheres (green sphere for PDB 5O6P).

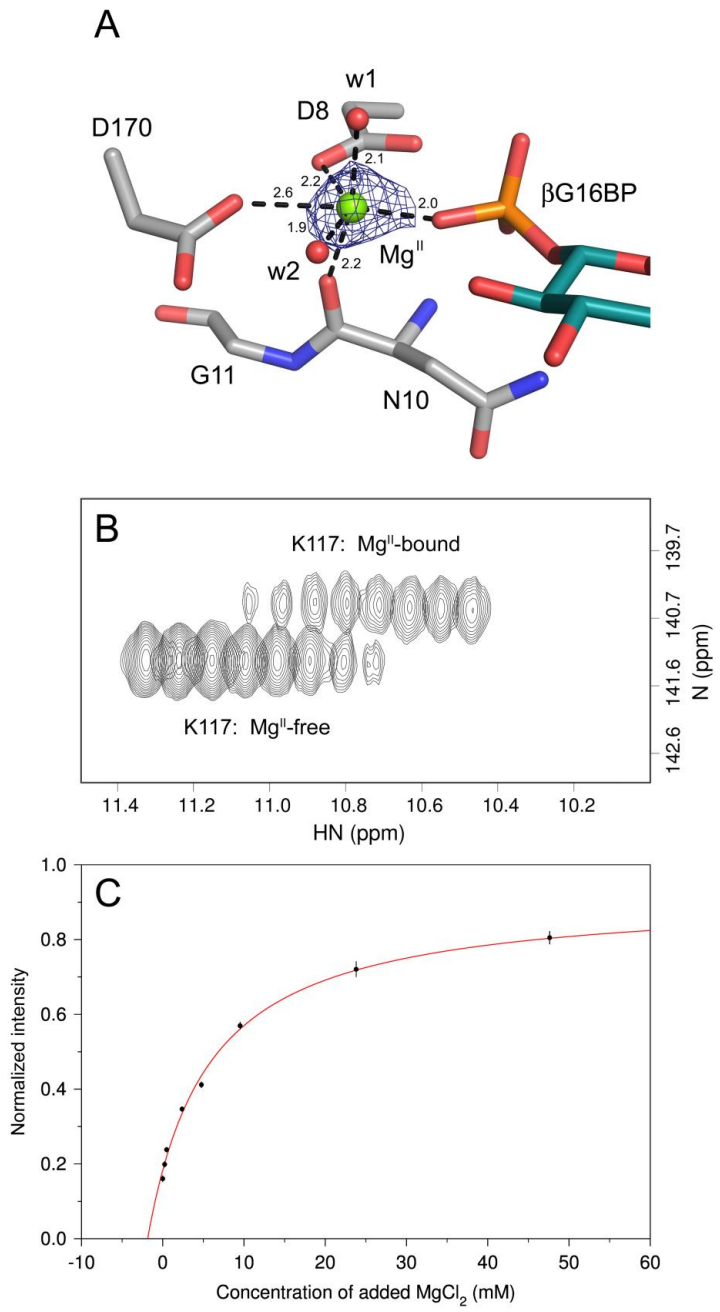

A
w1
D8
D170
2.1
2.2
2.6
2.0
1.9
βG16BP
Mg^{II}
w2
2.2
G11
N10
B
K117: Mg^{II}-bound
K117: Mg^{II}-free
139.7
140.7
141.6
142.6
N (ppm)
11.4
11.2
11.0
10.8
10.6
10.4
10.2
HN (ppm)
C
Normalized intensity
1.0
0.8
0.6
0.4
0.2
0.0
-10
0
10
20
30
40
50
60
Concentration of added MgCl$_2$ (mM)

Figure S7. Coordination and binding affinity of the catalytic Mg^{II} ion in the βPGM_{D10N}:P1G6P complex. (A) Active site of the βPGM_{D10N}:P1G6P complex (PDB 5OK1) with βG16BP and selected residues shown as CPK-colored sticks, structural waters shown as red spheres and the catalytic Mg^{II} ion indicated as a green sphere. The asymmetrical 2Fo – Fc electron density for the catalytic Mg^{II} ion is contoured at 2.5σ (blue mesh), with Mg^{II} coordination (black dashes) and atomic distances (Å) indicated. Restrained refinement of the βPGM_{D10N}:P1G6P complex (1.9 Å) results in a suboptimal coordination geometry for the catalytic Mg^{II} ion, as the cumulative atomic distance is ~0.2 Å too long between the oxygen atom of the 1-phosphate group of βG16BP (O – Mg^{II} = 2.0 Å) and the side chain carboxylate Oδ1 atom of D170 (O – Mg^{II} = 2.6 Å). The locations of the side chain carboxylate group of D8, the backbone carbonyl group of N10 and the water molecules present suggest that a more optimal binding geometry is accessible for the Mg^{II} ion when centered ~0.2 Å further towards the side chain carboxylate Oδ1 atom of D170. (B) Changes in peak intensity for residue K117 in a superposed series of $^{1}H^{15}N$ TROSY spectra (offset in ^{1}H frequency for clarity) as $MgCl_2$ is titrated into the Mg^{II}-free βPGM_{D10N}:P1G6P complex. As the concentration of $MgCl_2$ increases (left to right), the population of the Mg^{II}-free βPGM_{D10N}:P1G6P complex decreases with a concomitant increase in the population of the Mg^{II}-bound βPGM_{D10N}:P1G6P complex, consistent with a slow conformational exchange on the NMR time scale. The slow rate of Mg^{II} exchange most likely reflects the exclusion of its binding site by βG16BP. (C) Calculation of the binding affinity of Mg^{II} for the Mg^{II}-free βPGM_{D10N}:P1G6P complex using nonlinear least-squares fitting (red line) of normalized changes in averaged $^{1}H^{15}N$ TROSY peak intensities (black circles) for residues N10, G11, A115, K117 and I150. The dissociation constant (K_d) was determined to be 7.1 ± 0.6 mM. The initial concentration of Mg^{II} in the solution was evaluated as 1.9 ± 0.1 mM. Errors in peak intensity measurements are indicated as vertical black lines on each data point.

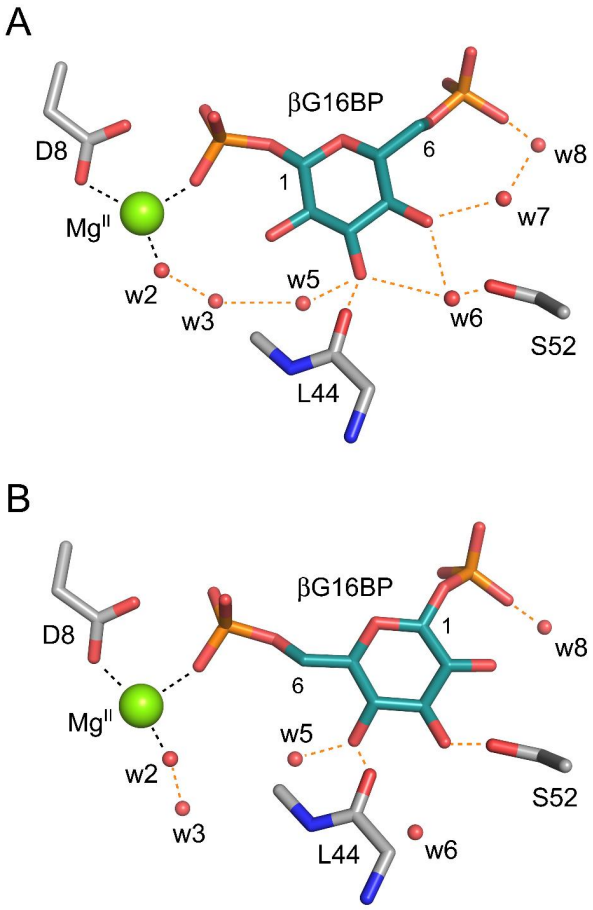


Figure S8. Comparison of βG16BP and structural water coordination in the βPGM$_{D10N}$:βG16BP complexes. The active sites of (A) βPGM$_{D10N}$:P1G6P complex (PDB 5OK1) and (B) βPGM$_{D10N}$:P6G1P complex (PDB 5OK0). Selected residues and the βG16BP ligand are shown as sticks in standard CPK colors, with structural waters (red) and the catalytic Mg^{II} ion (green) drawn as spheres. Orange dashes indicate hydrogen bonds and black dashes show catalytic Mg^{II} ion coordination.

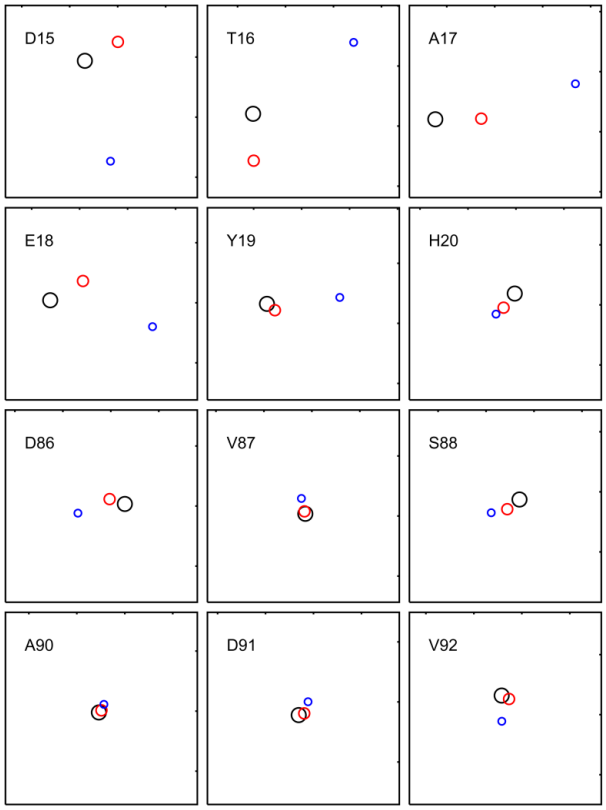


Figure S9. Comparison of backbone amide group peak positions in $^{1}H^{15}N$ TROSY spectra of βPGM_{D10N} complexes. $^{1}H^{15}N$ TROSY peak positions are shown for twelve hinge residues (D15, T16, A17, E18, Y19, H20, D86, V87, S88, A90, D91, V92) of the open βPGM_{D10N}:BeF_3 complex (black circle), the Mg^{II}-bound βPGM_{D10N}:βP1G6P complex (red circle) and the fully closed βPGM_{D10N}:AlF_4:G6P TSA complex (blue circle). The chemical shifts of these hinge residues are sensitive to the degree of closure of the cap and core domains and apart from D15 and T16, the $^{1}H^{15}N$ TROSY peaks of the Mg^{II}-bound βPGM_{D10N}:βP1G6P complex lie in an intermediate position between those of the open βPGM_{D10N}:BeF_3 complex and the fully closed βPGM_{D10N}:AlF_4:G6P TSA complex. These results indicate that the Mg^{II}-bound βPGM_{D10N}:βP1G6P complex is partially open in agreement with the crystal structures. Residues D15 and T16 do not follow this pattern and the crystal structures indicate that these residues play a crucial role in governing optimal hydrogen bonding for substrate coordination by positioning of the general acid–base and closure of the domains through rotation in backbone dihedral angles, which will be different in each of the complexes. For each panel, the x-axis

denotes the backbone amide proton ($^{1}H_{N}$) frequency, with a range of 2 ppm and the y-axis denotes the backbone amide nitrogen (^{15}N) frequency, with a range of 16 ppm. For the βPGM_{D10N}:BeF_3 complex, the $^{1}H_{N}$ and ^{15}N chemical shifts are (in ppm): D15 (8.34, 120.31), T16 (8.84, 113.98), A17 (9.12, 128.97), E18 (9.31, 119.79), Y19 (7.36, 118.38), H20 (8.01, 119.06), D86 (7.35, 114.83), V87 (7.07, 124.81), S88 (9.15, 125.64), A90 (7.77, 120.11), D91 (8.15, 116.09) and V92 (7.21, 123.41). For the Mg^{II}-bound βPGM_{D10N}:βP1G6P complex, the $^{1}H_{N}$ and ^{15}N chemical shifts are (in ppm): D15 (8.00, 118.73), T16 (8.83, 117.88), A17 (8.64, 128.91), E18 (8.97, 118.19), Y19 (7.28, 118.91), H20 (8.13, 120.23), D86 (7.51, 114.42), V87 (7.08, 124.61), S88 (9.28, 126.44), A90 (7.74, 119.94), D91 (8.10, 115.95) and V92 (7.13, 123.70). For the βPGM_{D10N}:AlF_4:G6P TSA complex, the $^{1}H_{N}$ and ^{15}N chemical shifts are (in ppm): D15 (8.07, 128.67), T16 (7.79, 108.03), A17 (7.66, 126.01), E18 (8.24, 121.99), Y19 (6.60, 117.84), H20 (8.21, 120.77), D86 (7.84, 115.59), V87 (7.11, 123.54), S88 (9.45, 126.74), A90 (7.72, 119.44), D91 (8.06, 114.98) and V92 (7.21, 125.55).

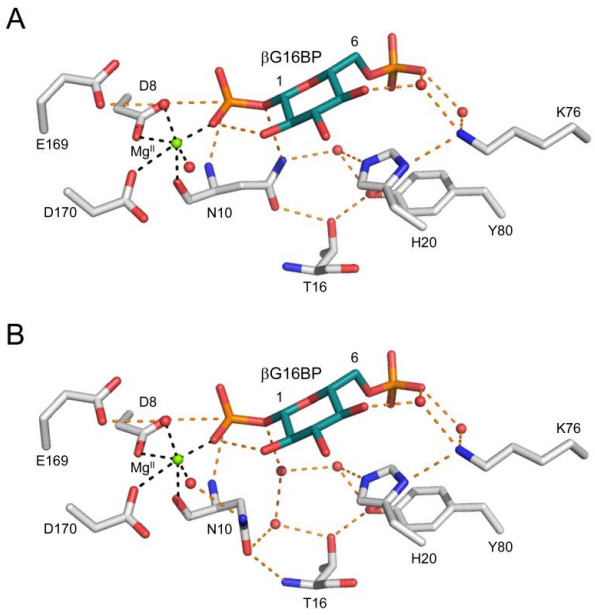


Figure S10. A model showing a potential mechanism for mutase activity in βPGM$_{D10N}$. Selected active site residues and ligands are shown as sticks in standard CPK colors, with structural waters (red) and the catalytic Mg^{II} ion (green) drawn as spheres. Orange dashes indicate hydrogen bonds and black dashes show metal ion coordination. (A) The βPGM$_{D10N}$:βP1G6P complex (PDB 5OK1; Figure 3B) with residue N10, the mimic of the protonated form of the general acid–base in the *in* position. The active site arrangement is analogous to that present in the copurified βPGM$_{D10N}$:βP1G6P complex (PDB 5O6P; Figure 3D). (B) A model of the βPGM$_{D10N}$:βP1G6P complex with N10 in the *out* position. In this model, the carbonyl oxygen atom of the carboxamide group of N10 forms a hydrogen bond to the amide group of T16, as observed in the βPGM$_{D10N}$:BeF_3 complex (PDB 5OJZ). Two water molecules which occupy the position of the general acid–base side chain when in the *in* position, comprise part of an extended hydrogen bonded network in the active site involving residues H20, K76, Y80, the phosphate group in the *distal* site, as well as structural and bulk water molecules. Any one of these groups could facilitate proton transfer to the bridging oxygen atom of the transferring phosphoryl group, allowing catalysis to occur in βPGM$_{D10N}$. The model was prepared by rotation of the N10 side chain and the addition of two water molecules in the active site of the PDB 5OK1 structure. Geometry was optimized against the existing electron density in COOT[42].

Table S1.

Data collection and data processing statistics for the βPGM$_{D10N}$ complexes

Complex	**βPGM$_{D10N}$:BeF$_3$**	**βPGM$_{D10N}$:P1G6P**	**Copurified βPGM$_{D10N}$:P1G6P**	**βPGM$_{D10N}$:P6G1P**	**βPGM$_{D10N}$:AlF$_4$:G6P**	**βPGM$_{D10N}$:AlF$_4$:H$_2$O:βG1P**
PDB code	PDB 5OJZ	PDB 5OK1	PDB 5O6P	PDB 5OK0	PDB 5OK2	PDB 5O6R
Crystallization conditions	0.6 mM substrate-free βPGM$_{D10N}$ 5 mM $BeCl_2$ 15 mM NaF	0.6 mM substrate-free βPGM$_{D10N}$ 15 mM βG1P, 5 mM $BeCl_2$ 15 mM NaF	0.6 mM copurified βPGM$_{D10N}$:P1G6P	0.6 mM substrate-free βPGM$_{D10N}$ 15 mM βG1P, 5 mM $BeCl_2$ 15 mM NaF	0.6 mM substrate-free βPGM$_{D10N}$ 10 mM G6P, 5 mM $AlCl_3$ 20 mM NaF	0.6 mM copurified βPGM$_{D10N}$:P1G6P 5 mM βG1P, 2 mM $AlCl_3$ 10 mM NH_4F
Crystal morphology	Thin plate crystals	Rod shaped crystals	Small needle crystals	Rod shaped crystals	Thin plate crystals	Large plate crystals
Wavelength (Å) **Beamline, Facility**	0.97625 Beamline i03, DLS	0.97950 Beamline i04, DLS	0.933 Beamline ID14-2, ESRF	0.97950 Beamline i04, DLS	0.97625 Beamline i03, DLS	0.933 Beamline ID14-2, ESRF
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions: **a, b, c (Å)** **α, β, γ (°)**	 52.8, 53.8, 81.6 90.0, 90.0, 90.0	 36.7, 74.5, 78.6 90.0, 90.0, 90.0	 31.8, 68.3, 83.2 90.0, 90.0, 90.0	 36.8, 54.9, 103.3 90.0, 90.0, 90.0	 37.5, 54.3, 104.7 90.0, 90.0, 90.0	 36.3, 54.9, 107.6 90.0, 90.0, 90.0
Resolution (Å) [1]	44.3 – 1.3 (1.33 – 1.30)	39.3 – 1.86 (1.91 – 1.86)	20.0 – 2.2 (2.26 – 2.20)	48.5 – 2.15 (2.21 – 2.15)	54.3 – 1.1 (1.12 – 1.10)	20.0 – 1.36 (1.40 – 1.36)
R_{merge} [1,2]	0.05 (0.93)	0.06 (1.90)	0.08 (0.27)	0.31 (1.50)	0.07 (1.03)	0.08 (0.27)
R_{pim} [1]	0.023 (0.542)	0.018 (0.558)	0.052 (0.167)	0.148 (0.721)	0.029 (0.572)	0.018 (0.077)
CC-half [1]	0.999 (0.489)	1.000 (0.599)	–	0.980 (0.391)	0.999 (0.484)	–
$\langle I/\sigma(I) \rangle$ [1]	16.1 (1.3)	25.8 (1.2)	10.0 (4.0)	5.3 (1.2)	12.5 (1.2)	24.5 (7.4)
Completeness (%) [1]	99.0 (88.3)	100.0 (99.9)	98.2 (99.5)	100.0 (100.0)	99.5 (91.6)	93.6 (91.3)
Multiplicity [1]	6.8 (4.4)	12.9 (13.3)	3.3 (3.3)	6.2 (6.4)	6.8 (4.7)	3.0 (2.8)
Total reflections	388282	242890	–	74044	591468	–
Unique reflections	57228	18807	31189	11990	87229	43148
Molecular replacement model	PDB 2WFA	PDB 2WF5	PDB 1O08	PDB 2WF5	PDB 2WF6	PDB 1O08

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

[2] $R_{merge} = \sum_h \sum_i |I(h) - I(h)_i| / \sum_h \sum_i I(h)_i$, where I(h) is the mean weighted intensity after rejection of outliers

Table S1 continued.

Refinement statistics for the βPGM$_{D10N}$ complexes

Complex	βPGM$_{D10N}$:BeF_3	βPGM$_{D10N}$:P1G6P	Copurified βPGM$_{D10N}$:P1G6P	βPGM$_{D10N}$:P6G1P	βPGM$_{D10N}$:AlF_4:G6P	βPGM$_{D10N}$:AlF_4:H_2O:βG1P
PDB code	PDB 5OJZ	PDB 5OK1	PDB 5O6P	PDB 5OK0	PDB 5OK2	PDB 5O6R
***R* (%) [3] / R_{free} (%) [4]**	13.6 / 17.0	19.8 / 24.6	19.5 / 24.8	22.0 / 29.2	14.8 / 17.0	10.7 / 14.6
Number of atoms:						
Protein	1724	1688	1611	1680	1812	1690
Ligands	4	20	20	20	21	21
Metal ions	2	1	1	1	3	2
Water	246	67	70	84	244	379
Protein residues	218	218	209	218	218	218
RMS deviations:						
Bonds (Å)	0.014	0.011	0.024	0.012	0.010	0.024
Angles (°)	1.48	1.49	2.39	1.50	1.47	2.15
Average B factors (Å^2):	20.44	42.68	39.63	31.86	15.41	15.01
Main chain	16.30	41.23	39.02	30.54	12.63	10.64
Side chains	20.83	44.75	41.39	33.62	15.40	14.15
Ligands	13.74	35.43	27.16	32.29	8.67	7.62
Metal ions	24.35	37.50	23.43	46.62	15.18	16.48
Water	32.26	38.77	31.28	28.34	25.91	27.33
Ramachandran analysis:						
Favored/allowed (%)	98.6	95.8	94.2	97.7	97.7	97.2
Disallowed (%)	0.0	0.5	0.5	0.0	0.0	0.0
MolProbity score	0.69 (100th percentile)	1.11 (100th percentile)	2.62 (41st percentile)	1.12 (100th percentile)	1.11 (96th percentile)	1.17 (97th percentile)

[3] $R = \sum_{hkl} \left| |F_{obs}| - k|F_{calc}| \right| \Big/ \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes

[4] $R_{free} = \sum_{hkl \subset T} \left| |F_{obs}| - k|F_{calc}| \right| \Big/ \sum_{hkl \subset T} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes and T is the test set of data omitted from refinement (5% in this case)

Table S2.
Pairwise cap domain rotations calculated using DynDom for selected βPGM complexes

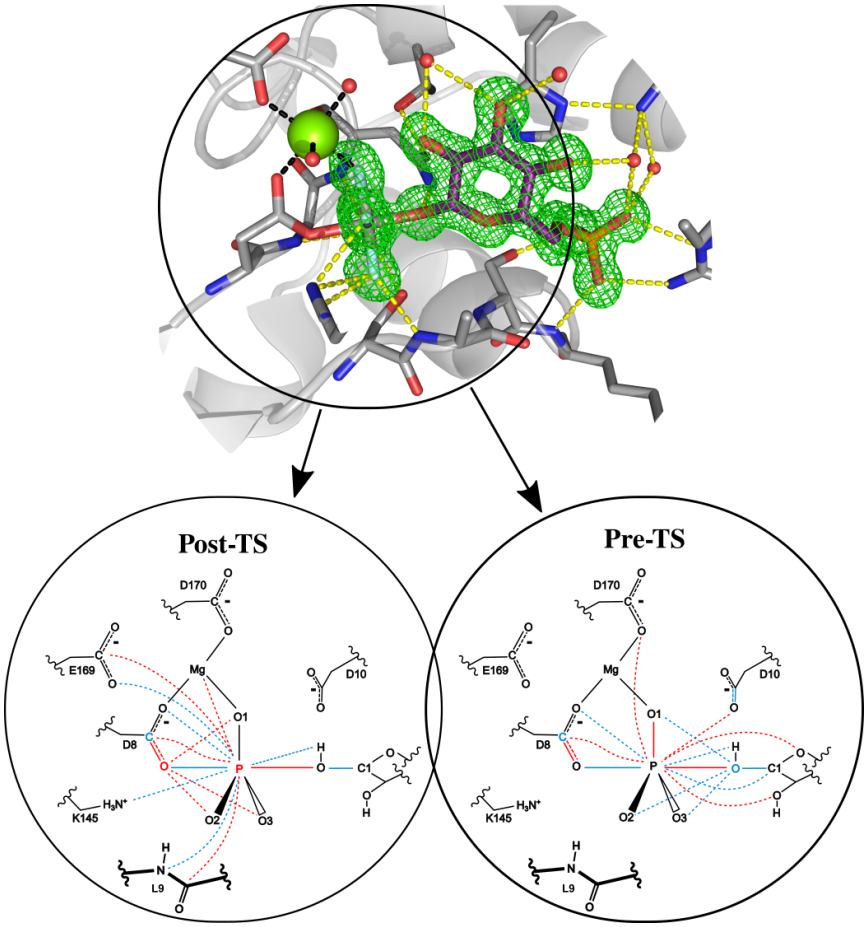
βPGM complex	**βPGM complex**	**Cap domain rotation (°)** [1]	**Cap domain translation (Å)** [2]
$βPGM_{WT}$ (PDB 2WHE)	$βPGM_{WT}$:MgF_3:G6P (PDB 2WF5)	35.0	1.4
$βPGM_{WT}$:BeF_3 (PDB 2WFA)	$βPGM_{WT}$:MgF_3:G6P (PDB 2WF5)	35.5	1.5
$βPGM_{D10N}$:BeF_3 (PDB 5OJZ)	$βPGM_{WT}$:BeF_3 (PDB 2WFA)	5.9	0.0
$βPGM_{D10N}$:P1G6P (PDB 5OK1)	$βPGM_{WT}$:MgF_3:G6P (PDB 2WF5)	13.2	–0.8
Copurified $βPGM_{D10N}$:P1G6P (PDB 5O6P)	$βPGM_{WT}$:MgF_3:G6P (PDB 2WF5)	13.5	–0.6
$βPGM_{D10N}$:P6G1P (PDB 5OK0)	$βPGM_{WT}$:MgF_3:G6P (PDB 2WF5)	14.0	–1.0
$βPGM_{D10N}$:AlF_4:G6P (PDB 5OK2)	$βPGM_{WT}$:MgF_3:G6P (PDB 2WF5)	No dynamic domains found	No dynamic domains found
$βPGM_{D10N}$:AlF_4:H_2O:βG1P (PDB 5O6R)	$βPGM_{WT}$:MgF_3:G6P (PDB 2WF5)	14.1	–1.0
$βPGM_{D10N}$:AlF_4:H_2O:βG1P (PDB 5O6R)	$βPGM_{WT}$:MgF_3:β$G1CF_5P$ (PDB 4C4S) [3]	14.2	–1.1
$βPGM_{D10N}$:AlF_4:G6P (PDB 5OK2)	$βPGM_{D10N}$:AlF_4:H_2O:βG1P (PDB 5O6R)	13.8	–0.9
$βPGM_{D10N}$:P1G6P (PDB 5OK1)	$βPGM_{D10N}$:AlF_4:G6P (PDB 5OK2)	13.0	–0.8
Copurified $βPGM_{D10N}$:P1G6P (PDB 5O6P)	$βPGM_{D10N}$:AlF_4:G6P (PDB 5OK2)	13.4	–0.5
$βPGM_{D10N}$:P6G1P (PDB 5OK0)	$βPGM_{D10N}$:AlF_4:G6P (PDB 5OK2)	13.7	–0.9
$βPGM_{D10N}$:P1G6P (PDB 5OK1)	Copurified $βPGM_{D10N}$:P1G6P (PDB 5O6P)	5.2	–0.3
$βPGM_{D10N}$:P6G1P (PDB 5OK0)	Copurified $βPGM_{D10N}$:P1G6P (PDB 5O6P)	8.3	–0.4
$βPGM_{D10N}$:P1G6P (PDB 5OK1)	$βPGM_{D10N}$:P6G1P (PDB 5OK0)	3.4	–0.1

[1] Hinge located at T16 for all pairwise comparisons

[2] DynDom translation term was less than ± 1.5 Å for all pairwise comparisons and so was not included in the text

[3] Where β$G1CF_5P$ corresponds to the α-fluorophosphonate analogue of β-glucose 1-phosphate[19]

A.2 Paper II: X-ray, NMR and QM approaches reveal the relationship between protein conformational change, proton transfer, and phosphoryl transfer in an archetypal enzyme



Contribution: I expressed and purified protein with different isotope enrichment schemes and enzymatically synthesized the U[^{13}C]C1-G6P for the NMR experiments. I performed the crystallography, NMR spectroscopy (including backbone assignment) and backbone model free analysis. I analysed and interpreted the data, I wrote programs to analyse and interpret chemical shift perturbations and display the output. I wrote the manuscript together with AW and JPW.

X-ray, NMR and QM approaches reveal the relationship between protein conformational change, proton transfer, and phosphoryl transfer in an archetypal enzyme.

Angus J. Robertson,[1] Alex L. Wilson,[2] Matthew J. Cliff,[2] Paul L.A. Popelier,[2] 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

* E-mail for J.P.W.: j.waltho@sheffield.ac.uk

Abstract

Molecular details for the timing and role of proton transfer in phosphoryl transfer reactions are poorly understood. Using a combination of NMR, X-ray crystallography and DFT approaches, we characterize pre- and post- proton transfer models of a phosphoryl transfer reaction in the archetypal phosphoryl transfer enzyme βPGM. We observe that the ionic nature of the AlF_4^- TSA may be highly useful in the investigation of proton transfer in phorphoryl transfer enzymes as out-of-plane distortion of the central Al^{3+} ion closely correlates with proton timing across the reaction coordinate. Backbone order parameters (ps-ns rigidity measurement) were used to guide QM model generation and residue truncation in pre- and post- proton transfer TSA models. The TS model displays a key contribution of this proton transfer to/from the GAB on the charge distribution within the transferring group, and consequently, the electrostatic interactions with surrounding residues in the active site. Given the free energy profile of the reaction, the GS of the reaction indicates that a high degree of proton transfer has already occurred to substrate which is closely reflected in ^{19}F and X-ray crystallographic observations which may further empower the use of ^{19}F NMR in the investigation of phosphoryl transfer reactions.

Introduction

I1) Background and the GAB controversy

Phosphate monoesters are labile in the active sites of phosphoryl transfer enzymes but extremely inert in aqueous solution (Lad, Williams, and Wolfenden 2003), in part due to the negative charge providing a strong repulsion to potential attacking nucleophiles. Some phosphoryl transfer enzymes alleviate this repulsion by populating near attack complexes (NACs) in which the attacking hydroxyl group hydrogen bonds with the transferring phosphate group in a nonproductive orientation (Griffin et al. 2012; Jin, Richards, et al. 2017). The residue that provides general acid-base (GAB) catalysis (Lassila, Zalatan, and Herschlag 2011; Kamerlin et al. 2013) is thereby utilized not only to activate the hydroxyl group for nucleophilic attack of the target phosphate group, but also to stimulate the alignment of the nucleophilic oxygen atom with the phosphorus atom, in a conformation that is geometrically close to the transition state (TS). Structural investigations of near TS species have made use of both MgF_3^- and AlF_4^- as transition state analogues (TSAs) that mimic the transferring phosphate group, as they are planar and have a net single negative charge when complexed with substrate in the enzyme active site (Baxter et al. 2008; Cliff et al. 2010; Jin, Richards, et al. 2017; Jin, Molt, and Blackburn 2017). The TSA structures have indicated that the engagement of GAB residues is concurrent with phosphoryl group transfer. However, controversy remains as to the timing of proton transfer associated with GAB catalysis meaning that any interpretation of the mechanism and the energy barrier of the chemical step is open to question. β-phosphoglucomutase (βPGM) [EC 5.4.2.6] is an archetypal phosphoryl transfer enzyme that utilizes GAB catalysis and has been well characterized enzymatically and structurally (Lahiri et al. 2004; Zhang et al. 2005; Dai et al. 2006, 2009, Baxter et al. 2006, 2008; Golicnik et al. 2009; Baxter et al. 2010; Griffin et al. 2012; Jin et al. 2014; Johnson et al. 2018). βPGM catalyzes the reversible isomerization of β-glucose 1-phosphate (βG1P) and glucose 6-phosphate (G6P) via a β-glucose 1,6-bisphosphate (βG16BP) intermediate. Previous computational studies on the phosphoryl transfer between βG16BP and residue D8 of βPGM (generating G6P), have presented conflicting timings for the proton transfer associated with the GAB residue (D10). Analyzing the trajectories in the direction of phosphoryl group transfer from βG16BP to D8, these studies range in prediction from “early” (Webster 2004; Marcos, Field, and Crehuet 2010), through “concerted” (Barrozo et al. 2018), to “late” (Elsässer, Dohmeier-Fischer, and Fels 2012) proton transfer events, with predicted barrier heights ranging from 14 to 64 kJ mol^{-1}.

I2) D10N literature and what we establish here

The D10N variant of βPGM (βPGM_{D10N}), which serves as a model of wild-type βPGM (βPGM_{WT}) with the GAB residue in its protonated form, was found to trap a ground state (GS) analog complex in which the phosphorus atom of the 1-phosphate group of βG16BP is at van der Waals contact distance from the nucleophilic carboxylate oxygen of D8 (Johnson et al. 2018). This observation demonstrated that without proton transfer from the GAB to the bridging oxygen of βG16BP, the phosphate group prefers to remain associated with βG16BP. This is consistent with an "early" proton transfer step during phosphorylation of D8 by βG16BP, and suggests that the N10 variant provides a good model of the pre-proton-transfer state. However, in order to determine whether the GAB proton is likely to be transferred to the nascent hydroxyl group before the peak of the energy barrier in the native reaction, a post-proton-transfer model is also required. In this report, we establish that the AlF_4^- complex of wild-type βPGM with G6P provides a suitable post-proton-transfer model while the equivalent complex of the D10N variant remains a suitable pre-proton-transfer model, which allows a direct comparison of states on either side of the proton transfer step. While aluminium fluorides act as surrogates for transferring phosphoryl groups in the transition state, they have reduced atomic charges and little covalency in their bonding (Griffin et al. 2012). Correspondingly, they are shown to report on the electronic distribution within the active site pre- and post- proton-transfer, since they distort their geometry in line with the preferred positions of the phosphorus atom in each scenario. In parallel, solution NMR measurements are used to calibrate DFT calculations to generate reliable models of the reaction trajectory for phosphoryl transfer. Collectively, these studies establish the timing of proton transfer in this reaction using a novel approach that is applicable to other phosphoryl transfer enzymes that rely on GAB catalysis.

Results

R1) NMR investigation of the complexes.

Both βPGM_{WT}:AlF_4:G6P and βPGM_{D10N}:AlF_4:G6P complexes have been crystallized previously (Baxter et al. 2010; Johnson et al. 2018). In the WT complex, it was assumed that the GAB proton (the proton that transfers between atom Oδ1 of residue D10 ($O\delta1_{D10}$) and the 1-oxygen group of G6P ($O1_{G6P}$)) is associated with the sugar (i.e. a post-proton transfer model), whereas in the D10N complex it was assumed that $O1_{G6P}$ was deprotonated (i.e. a pre-proton transfer model) (Johnson et al. 2018). However, the position of the GAB proton was not established independently in either complex, and therefore was investigated here using solution NMR methods. The βPGM_{D10N}:AlF_4:G6P complex was prepared as described previously (Baxter et al. 2010; Johnson et al. 2018) and 97% backbone assignment of non-proline residues was determined (BMRB ID: 27697, Supporting information (SI) Section 3). The chemical shifts were compared to the previously assigned βPGM_{WT}:AlF_4:G6P complex (Baxter et al. 2010). Only subtle perturbations were present and these occurred in four distinct regions, each of which is in direct contact with the substrate (SI Fig. S1). This indicates that the protein conformation and the accommodation of substrate in the active site is very similar in the two complexes. NMR relaxation measurements of fast (ps-ns) dynamics corroborate this interpretation, with few significant differences in the observed order parameters. The average order parameter in the βPGM_{D10N}:AlF_4:G6P complex is only 0.02 lower than that in the WT complex (SI Section 6, Fig. S2). Almost all of the discernible changes are distant from the active site and are juxtaposed to changes of the opposite sign, indicative of local compensatory mechanisms (SI Fig. S2, S3, S4).

In the βPGM_{WT}:AlF_4:G6P complex, the assumption that the GAB proton is associated solely with the sugar is challenged by ^{13}C chemical shift measurements (Table S2). For example, the Cβ ^{13}C chemical shift of D10 is the most upfield of the Asp residues in this complex implying protonation to some extent (SI Table S1). It is *ca*. 2 ppm upfield compared to in the more open βPGM_{WT}:BeF_3^- complex, where the D10 residue has rotated out of the active site (Griffin et al. 2012) (SI Table S1). However, it is only 0.44 ppm upfield of the equivalent resonance of D180 in the βPGM_{WT}:AlF_4:G6P complex, which is surface exposed and likely to be deprotonated at the experimental pH. Together, this indicates that some sharing of the GAB proton between atoms $O1_{G6P}$ and $O\delta1_{D10}$ in the βPGM_{WT}:AlF_4:G6P complex is likely. To shed further light on the behavior of the GAB proton, 1D ^{19}F NMR spectra of the βPGM_{WT}:AlF_4:G6P and βPGM_{D10N}:AlF_4:G6P complexes were compared, and an average downfield chemical shift change of 4 ppm is observed for AlF_4^- peaks in the D10N complex (Fig. 2). The hydrogen bonding to the fluorides in the two complexes was assessed using solvent induced isotope shifts (SIIS), which are sensitive to the distance between hydrogen bonding partners and the fluoride ions (SI Section 13). The SIIS values for the βPGM_{D10N}:AlF_4:G6P complex reflect those of the WT complex. While there is a small overall reduction in SIIS values of the AlF_4^- moiety (*ca.* 0.1 ppm), this is consistent with only a minor change in hydrogen bonding between the enzyme and the AlF_4^- group (SI Fig S7, SI Table S4-6). Thus, while changes in hydrogen bonding geometry can be eliminated as the primary source of the average downfield chemical shift change between the two complexes, the value is too small for any differences in the GAB proton position to be confidently predicted without further corroborating evidence.

R2) X-ray investigation of the D10N complex.

The assumption that the $O1_{G6P}$ atom was not protonated in the βPGM_{D10N}:AlF_4:G6P complex (PDB: 5OK2) was based on the assumed orientation of the carboxamide of N10, but this was not explicitly validated. Re-refinement of the N10 carboxamide in the opposite orientation yielded a difference map peak of >3 σ for the alternately modeled atoms (SI Section 8, Fig S5), which indicates that the carboxamide indeed adopts the previously assumed orientation. A crystal of the βPGM_{D10N}:AlF_4:G6P complex at a higher resolution (1.02 Å) corroborates this interpretation, and further supports a model where the $O1_{G6P}$ atom is deprotonated (PDB: 6L03; SI Section 8,10,11; SI Fig. S6). Unexpectedly, the higher resolution structure also revealed a clear distortion of the AlF_4^- group. Since AlF_4^- is predominantly an ionic moiety, it is not strictly tied to a particular geometry in the active site of phosphoryl transfer enzymes, and hence has the potential to report (through its distortion) on which axial ligand has the higher charge density. If the $O1_{G6P}$ atom is deprotonated in the βPGM_{D10N}:AlF_4:G6P complex, it is expected to have a much higher charge density than atom Oδ1 of residue D8 ($O\delta1_{D8}$) on the basis of their solution pKa values. In this scenario, the position of the Al^{3+} atom of the AlF_4^- might be expected to be biased towards the $O1_{G6P}$ atom. In the high resolution crystal structure of the βPGM_{D10N}:AlF_4:G6P complex an out-of-plane distortion of the Al^{3+} atom towards the $O1_{G6P}$ atom of 3° is observed (SI Section 12). A re-examination of 5OK2 indicates that this distortion is also present in the lower resolution structure but the angle of distortion cannot be defined accurately. The observed distortion corroborates that in the environment of the enzyme in the D10N complex the $O1_{G6P}$ atom has a higher charge density compared with the $O\delta1_{D8}$ atom, which is consistent with hydrogen bonding between the NH_2 group of the carboxamide mimic of the GAB residue and a deprotonated $O1_{G6P}$ atom. In the WT complex, no distortion from planarity of the AlF_4^- is observed within error (SI Section 12), which is consistent with a substantial degree of protonation of the $O1_{G6P}$ atom. Indeed, the behavior of the βPGM_{WT}:AlF_4:G6P complex implies that the attraction and repulsion provided by the two axial ligands of the Al^{3+} atom is balanced by the extent of proton transfer towards the sugar compared with in the D10N complex. The corollary of these observations is that significant proton transfer from the GAB residue to the $O1_{G6P}$ atom is required for the AlF_4^- mimic of the transferring phosphoryl group to be most stable in its planar form, in line with an “early” proton transfer event stabilizing a planar phosphoryl group during phosphoryl transfer from βG16BP to D8 in the native reaction.

R3) QM model generation and validation.

In order to investigate more quantitatively the interactions and charge distribution that give rise to the geometries of the AlF_4 moieties observed in the crystal structures, large Quantum Mechanical (QM) models of the βPGM_{WT}:AlF_4:G6P and βPGM_{D10N}:AlF_4:G6P complexes were constructed using previously established methodology (Jin et al. 2016; Jin, Molt, et al. 2017; Himo et al. 2005; Noodleman et al. 2004). A QM model consisting of 386 atoms was initially computed for the βPGM_{WT}:AlF_4:G6P complex (starting from PDB: 2WF6 (Baxter et al. 2010)) using Gaussian09 (Frisch et al., n.d.), before the D10N mutation was introduced and the model reoptimized (SI Section 14). An increase in the average F_x-Al-$O1_{G6P}$ angle by 4° was observed in the βPGM_{D10N}:AlF_4:G6P complex model when the $O1_{G6P}$ atom is formally deprotonated, which is in good agreement with the geometrical distortion in the experimental crystal structure (3°). Computed ^{19}F NMR chemical shift changes between the two complexes reveal an average downfield chemical shift change of 3.1 ppm in the βPGM_{D10N}:AlF_4:G6P complex when the $O1_{G6P}$ atom is formally deprotonated, which also is in good agreement with experiment (4.0 ppm downfield) (Fig. 2; SI Section 15). Hence the QM models provide a firm basis from which to quantify the relationship between proton and phosphoryl group transfer.

In the QM model of the βPGM_{WT}:AlF_4:G6P complex, the GAB proton is primarily associated with the $O1_{G6P}$ atom with a bond order of *ca*. 0.5, compared with a bond order of *ca*. 0.2 to the $O\delta1_{D10}$ atom. This sharing of the GAB proton is consistent with the chemical shift of the Cβ ^{13}C resonance of D10 observed in solution. To defend the assertion that the $O1_{G6P}$ atom is deprotonated in the βPGM_{D10N}:AlF_4:G6P complex, the $O1_{G6P}$ atom of the βPGM_{D10N}:AlF_4:G6P model was protonated and the structure reoptimized. This model predicted a planar distortion of the AlF_4^- moiety (Fx-Al-$O1_{G6P}$ angle) in the opposite direction compared to both the 1.02 Å crystal structure (6°) and the deprotonated QM model (10°). Furthermore, calculated ^{19}F chemical shifts (F1:146.5, F2: 138.1, F3:129.6, F4: 139.4 ppm) are upfield relative to WT, rather than the downfield chemical shift change observed both experimentally, and computationally in the deprotonated βPGM_{D10N}:AlF_4:G6P model. Thus, the protonation states of the $O1_{G6P}$ atom and the GAB residue (or analog) are established in both TSA complexes.

R4) Relationship between proton transfer, charge distribution, and the energy barrier.

To investigate further the effects of proton transfer between the GAB and the $O1_{G6P}$ atom, the electron density of each model was rigorously partitioned into atomic basins according to the Quantum Theory of Atoms in Molecules (QTAIM) (Bader 1990; P L A Popelier 2000) using the software package AIMAll17 (Keith 2017), and relevant atomic charges were evaluated (SI Section 19). An increase in electron density of 100 me is observed on the $O1_{G6P}$ atom when the GAB proton is associated with the GAB compared to with G6P. The Interacting Quantum Atoms (IQA) method (Pendás, Francisco, and Blanco 2005; Blanco, Martin Pendas, and Francisco 2005; Francisco, Pendás, and Blanco 2006) was then chosen as an appropriate energy decomposition scheme for the approximation of various interaction energies, since it makes use of the electron density partitioning within QTAIM. This allows the evaluation of relative changes in pairwise energies between the βPGM_{WT}:AlF_4:G6P and βPGM_{D10N}:AlF_4:G6P complexes. Using this approach, a stronger electrostatic interaction (-2504 kJ mol^{-1} compared with -2179 kJ mol^{-1}) was observed between Al^{3+} and the $O1_{G6P}$ atom in the βPGM_{D10N}:AlF_4:G6P complex as a result of increased electron density on the $O1_{G6P}$ atom (SI Fig. S9), which is in line with the observed reduction in the Al^{3+} - $O1_{G6P}$ bond length.

R5) The implications of proton transfer on phosphoryl group transfer

Translating from metal fluoride analogue complexes to the native reaction, the AlF_4^- groups in both the βPGM_{WT}:AlF_4:G6P and βPGM_{D10N}:AlF_4:G6P QM models were replaced with PO_3^- (SI Section 16, 17). During geometrical relaxation, all atoms were held fixed except for those in the PO_3^- group and the GAB proton. This allowed the assessment of the geometrical and electronic effects of the phosphate group (SI Fig. S9), which has substantially greater polarization than AlF_4^-, in the protein environment defined by the metal fluoride complexes. In the PO_3^- complexes, electron density redistributes towards the equatorial oxygens, leaving the phosphorus atom with a much larger positive charge in comparison to the Al^{3+} ion. This results in a stronger electrostatic interaction between the $O1_{G6P}$ atom and the PO_3^- group, and hence the planar distortion is exaggerated for PO_3^- in these models (*ca.* 4°). This result leads to the hypothesis that proton transfer is a necessary initial step to mediate the ground state P - $O1_{G6P}$ electrostatic interaction via localized electron density redistribution, thereby reducing the overall energetic barrier to reaction.

R6) Consequences for the energy barrier

In order to analyze the energetic barrier associated with phosphoryl group transfer with βPGM in the TSA conformation, the initial $βPGM_{WT}$:PO_3:G6P model was trimmed to 163 atoms to be computationally viable while maintaining all atoms that interacted with the transferring phosphoryl group (SI Section 16 and 17). The resulting structure was optimized to a transition state (TS), with one vibrational mode corresponding to motion of the transferring PO_3^- group along the reaction coordinate (SI Fig. S8). The principal geometrical features of the computed TS are the changing bond lengths of the atoms undergoing bond-making and bond-breaking processes (Fig. 3). However, the relative contributions of each intra- and inter-atomic energy term to the overall energy profile of phosphoryl group transfer were evaluated. This involved a full energy decomposition of the molecular wavefunction using an IQA approach at each snapshot along the reaction coordinate, utilizing the Relative Energy Gradient (REG) method implemented in the software program ANANKE (Thacker and Popelier 2018, 2017; Alkorta, Thacker, and Popelier 2018)(SI section 19).

The reaction coordinate was split into 4 segments, defined according to stationary points on the energy profile. All possible intra- and inter-atomic energy terms were ranked by their relative contributions to the overall energy profile of phosphoryl group transfer. Analyzing the trajectories when transferring phosphate from the $O1_{G6P}$ atom to the $O\delta1_{D8}$ atom, segment 1 describes the contributions of all energy terms to the total energy prior to formation of what constitutes a ground state (GS_{12}) in this protein conformation (Fig. 3). Segments 2 and 3 describe the pre- and post-transition state (TS_{23}) respectively, and segment 4 describes post-formation of the product phospho-enzyme ground state (GS_{34}). Examination of the two ground states associated with the reaction trajectory reveals that in the protein conformation adopted by the transition state analogue complexes, the phosphoryl group is already partially dissociated from the leaving group oxygens. In GS_{12} the P - $O1_{G6P}$ bond order is 0.61, and in GS_{34} the P - $O\delta1_{D8}$ bond order is 0.53. For comparison, at the transition state (TS_{23}) the P - $O1_{G6P}$ bond order is 0.21 and the P - $O\delta1_{D8}$ bond order is 0.36. It is also apparent that in GS_{12} there is already substantial proton transfer from D10 to the sugar phosphate (Bond order: H - $O1_{G6P}$ = 0.41, H - $O\delta1_{D10}$ = 0.34). This illustrates that proton transfer is coupled to phosphoryl group cleavage from βG16BP. However, in this trajectory proton transfer is far from completed by GS_{34} (Bond order: H - $O1_{G6P}$ = 0.50, H - $O\delta1_{D10}$ = 0.24). For comparison, in the TS_{23} the H - $O1_{G6P}$ bond order is 0.54 and the H - $O\delta1_{D10}$ bond order is 0.20. Overall the simulation shows that when the protein is in the conformation associated with the transition state the GAB proton is preferentially associated with sugar throughout. Hence, the proton transfer step can be considered to be “early” (when transferring phosphate from the $O1_{G6P}$ atom to the $O\delta1_{D8}$ atom) but essentially the proton remains shared throughout the phosphoryl group transfer process. The corollary of the observation is that the adoption by the protein of the conformation associated with the TSA complex structures is synergistic with partial proton transfer and partial dissociation of the phosphoryl group from the leaving group oxygen atoms.

The subset of intra- and inter-energetic terms required to reproduce the relative energies of reactant, transition state, and product were elucidated by analysis of the intra- and inter-atomic energy terms in segments 2 and 3 of the trajectory. In both segments, the principal inter- and intra-atomic terms are between atoms directly involved in the phosphoryl group transfer. However, for segment 2 there are inter-atomic terms with significant REG values that provide both stabilizing and destabilizing interactions on both sides of the transferring phosphate, whereas in segment 3, the dominant inter-atomic terms lie predominantly on the D8 side of the transferring phosphate (Fig. 4). This increase in destabilizing interactions on the D8 side of the reaction correlates with the low observed energy barrier in the 'reverse' direction (~2 kJ mol^{-1}) and the higher energy of the product state relative to the ground state (35 kJ mol^{-1}).

Note: The energies were extracted from the figure, could Alex check against data?

Discussion

Together the models illustrate that proton transfer is important in stimulating phosphoryl transfer from $O1_{G6P}$ to $O\delta1_{D8}$ when the protein is in the TSA conformation, but that neither the phosphoryl transfer nor the proton transfer processes are complete in the corresponding GS on either side of the barrier. Completion of phosphoryl and proton transfer therefore requires a change in protein conformation.

In the experimental βPGM_{D10N}:βG16BP ground state complex (Johnson et al. 2018) the $O\delta1_{D8}$ atom is positioned in line with the 1-phosphate group at van der Waals contact distance from the phosphorus atom, and the NH_2 group of N10 is hydrogen bonding with the bridging oxygen of βG16BP. However, for the protein to adopt this conformation the relative orientation of the cap and core domains undergoes a rotation of 13° relative to the TSA conformation. This rotation also disrupts the hydrogen bonding network of part of the catalytic machinery in the vicinity of the GAB residue.

The relevance of the TSA conformation to the phosphoryl transfer reaction is illustrated by the success of EVB approaches in accurately reproducing the height of the transition state free energy barrier, when starting in this conformation (Barrozo et al. 2018) . This study also highlighted the importance of the hydrogen bonding network in the vicinity of the phosphoryl group during its transfer.

The energy barrier calculated in our model (~2 kJ mol^{-1} between GS_{34} and TS_{23}) is almost zero and significantly lower than previously reported in equivalent simulations of QM models with fewer atoms (25-56 kJ mol^{-1}) (Marcos, Field, and Crehuet 2010). In our model the atoms with fixed positions in line with the making/breaking P-O bonds of the transferring PO_3^- group are the backbone nitrogen of D8, and the phosphorus atom of the phosphate group in the *distal* site. In previous models, the unconstrained atoms did not include the entirety of the sugar phosphate and P-O bond making/breaking is able to go to completion. Hence, these simulations appear to have sufficient freedom to mimic some protein conformational change outside of that dictated by the TSA conformation, which leads to more stable apparent ground states and therefore higher apparent barriers than observed in our simulation.

Conclusion

In conclusion, the experimental X-ray structures and NMR measurements, in combination with QM models and their partitioning into atomic basins, all point towards a synergy between proton transfer from the GAB residue and the early stages of phosphoryl transfer. Both of these processes are assisted by the transition of the protein conformation between NAC III and that adopted in the metal fluoride TSA complexes. This final closure of the two domains is associated with a significant degree of transfer of the GAB proton from the $O\delta1_{D10}$ to the $O1_{G6P}$ atom of βG16BP, which couples to partial dissociation of the phosphoryl group. The fixing of the peripheral atoms in the simulation at the positions defined by the TSA conformation thus prevented the enzyme from going through full reaction trajectory of phosphoryl transfer but defines an almost barrierless transition from the phospho-enzyme to βG16BP. The corollary of this is that the transformation from the TSA conformation to the NAC III is associated with the completion of bond formation of both phosphoryl and proton transfer. It is therefore not possible to partition the whole energy barrier into a conformational term and a chemical term as the two are coupled in this case, which strongly implicates protein conformational change between alternatively closed structures to be instrumental in the catalysis of phosphoryl transfer in βPGM.

These data confirm that mutation of the GAB residue from Asp to Asn serves as a good model of the pre-proton transfer state, and that the corresponding WT complexes serve as a good post-proton-transfer model in a GAB catalyzed phosphoryl transfer reaction. The study also shows that the distortion of the predominantly ionic metal fluoride TSA moieties can be used to report on the relative charges of the axial oxygens that constitute the nucleophile and the leaving group for the reaction *in situ* within the enzyme. Finally, ^{19}F NMR measurements can be used in combination with QM models to corroborate the protonation state of the nucleophile and the leaving group oxygen atoms in the TSA complexes, thereby validating the reliability of QM models of the native reaction trajectory.

Experimental Section

Details of the experimental methods for X-ray crystallography, ^{19}F NMR, and DFT computations are given in the Supporting Information. The nomenclature system used here to described oxygen and phosphorus atoms in the structures is as recommended by IUPAC.

Acknowledgments

These studies were supported, in part, by BBSRC (Grants BB/E017541, BB/K016245 and BB/M021637 to J.P.W.). AJR and AW were funded by department studentships and we would also like to thank the Universities of Sheffield and Manchester for support. We would also like to thank Geoff Kelly at the Sir Francis Crick institute, London UK for his expertise in setting up experiments on the 950 MHz sepctrometer.

Data Availability.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.rcsb.org) with the following PDB codes:

βPGM$_{D10N}$:AlF4:G6P complex (1.10 Å;5OK2)

βPGM$_{D10N}$:AlF4:G6P complex (1.02 Å; 6L03)

The NMR chemical shifts have been deposited in the BioMagResBank (www.bmrb.wisc.edu) with the accession number: 27697

Figures

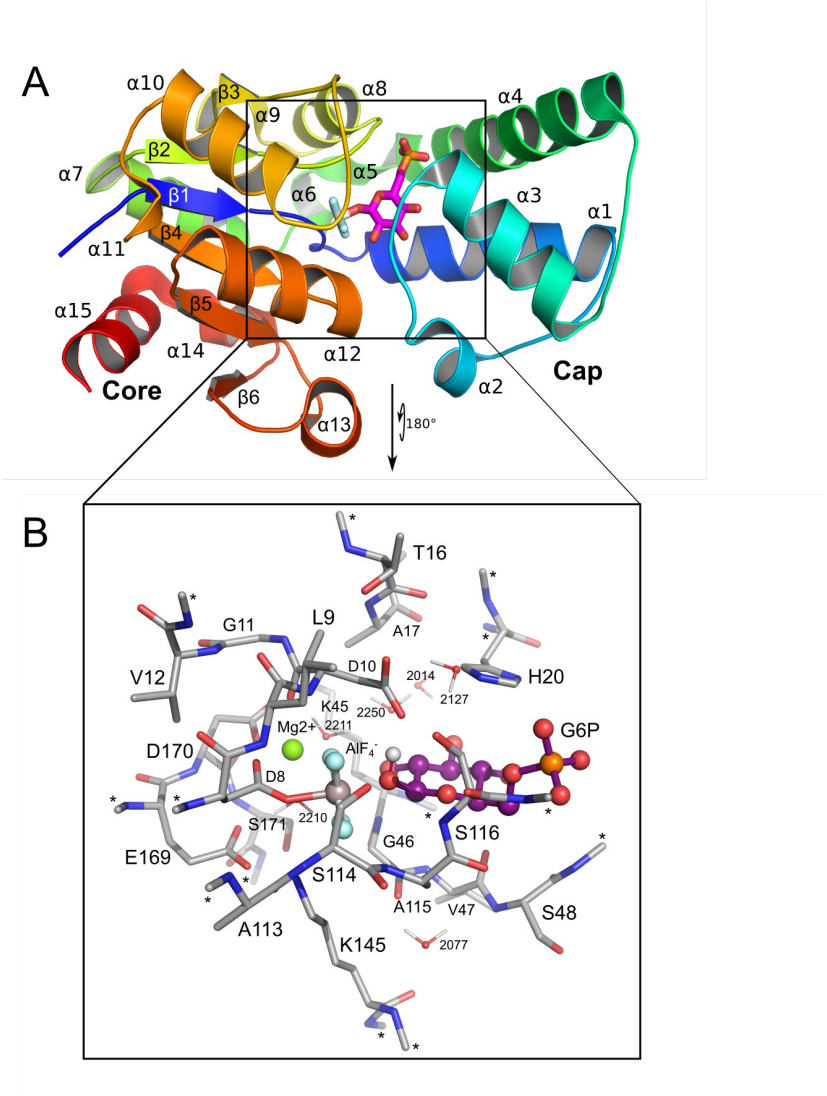
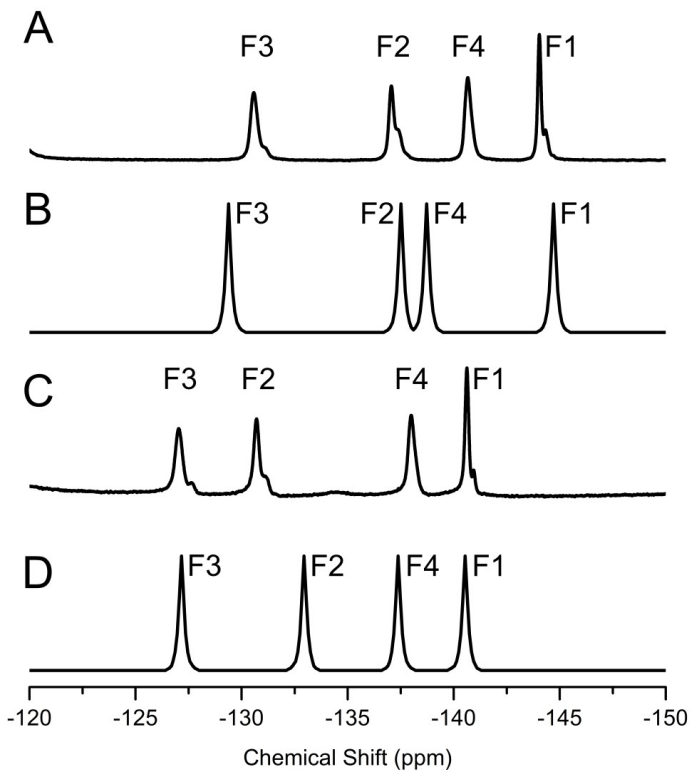
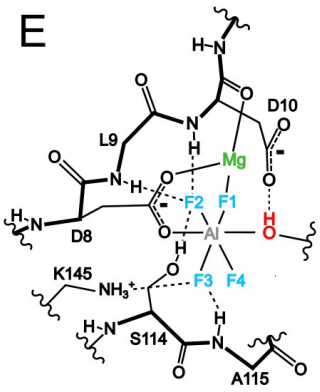


Fig 1.

An overview of the βPGM enzyme and QM model generated. **A)** An annotated cartoon illustration of the 2WF6 βPGM_{WT}:AlF_4:G6P TSA complex with G6P (pink) and AlF_4^- (grey) ligands illustrated. **B)** The active site region used in the QM calculations is shown as sticks in standard CPK colors, but with carbon atoms (dark gray), fluorine (light blue spheres), aluminum (dark gray sphere), βG6P (purple carbon atoms as spheres), and MgII ion (green sphere) . Structural waters (red) are drawn as ball and sticks and annotated with the water number in the model (SI Section 14) . An asterisk denotes truncation points in the model which are consistently adjacent to a backbone amide (SI Section 14).



E



F

	Expt.	Calcd.	SIIS
F1	-144.0	-144.7	0.8
F2	-137.0	-137.5	1.1
F3	-130.6	-129.4	0.9
F4	-140.6	-138.7	0.8

G

	Expt.	Calcd.	SIIS
F1	-140.6	-140.5	0.5
F2	-130.7	-132.9	0.9
F3	-127.1	-127.2	0.9
F4	-138.0	-137.4	0.8

389 **Fig. 2**

390 Comparison of experimental and calculated fluorine shifts for WT and D10N :AlF_4:G6P complexes.
391 **A)** Experimental ^{19}F 1D NMR spectra of βPGM_{WT}:AlF_4:G6P complex in 90% H_2O and 10% D_2O,
392 **B)** calculated ^{19}F 1D NMR spectra of βPGM_{WT}:AlF_4:G6P complex, **C)** experimental ^{19}F 1D NMR
393 spectra of βPGM_{D10N}:AlF_4:G6P complex in 90% H_2O and 10% D_2O, **D)** calculated ^{19}F 1D NMR
394 spectra of βPGM_{D10N}:AlF_4:G6P complex. E) An active site schematic to correlate fluorine label and
395 geometric position. Experimental and calculated chemical shifts for the βPGM_{WT}:AlF_4:G6P
396 complex (F; spectra A and B), and βPGM_{D10N}:AlF_4:G6P complex (G; spectra C and D) are
397 presented alongside solvent induced isotope shift (SIIS) values for each of the resonances (SI
398 section 13).

399

400

401

402

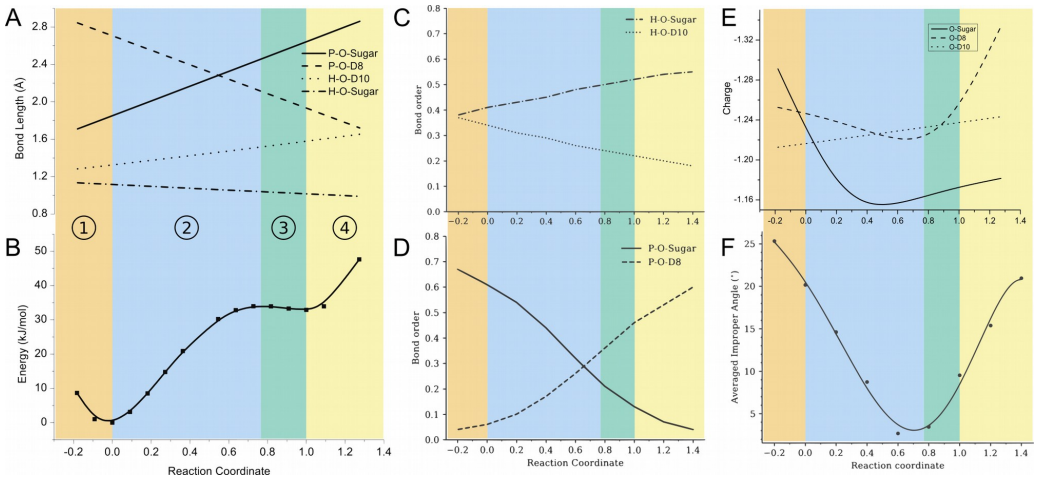


Fig. 3

The TS model of phosphoryl transfer in βPGM_{WT}. **A)** The interatomic distances for key atoms in the phosphoryl group transfer reaction. **B)** The resultant free energy profile with GS at reaction coordinate 0 and product state at reaction coordinate 1. The central numbering (and coloring) scheme corresponds to the four segments defined between stationary points on the energy profile. The central numbering (and coloring) scheme corresponds to the four segments defined between stationary points on the energy profile. **C)** The bond order of the H-O-G6P and H-O-D10(carboxylate) bonds across the reaction trajectory in B using the quantum chemical topology (QCT) method (Outeiral et al. 2018; Vosko, Wilk, and Nusair 1980; Paul L A Popelier 2005) **D)** The bond order of the P-O-G6P and P-O-D8(carboxylate) bonds across the reaction trajectory in B using the QCT method. **E)** The charge partitioned to the 1-oxygen of G6P, and the carboxylate oxygens of residues D8 and D10 across the reaction trajectory in B using the IQA method. **F)** The improper angle that the phosphate atom makes to the plane of its associated 3 oxygen atoms across the reaction trajectory in B.

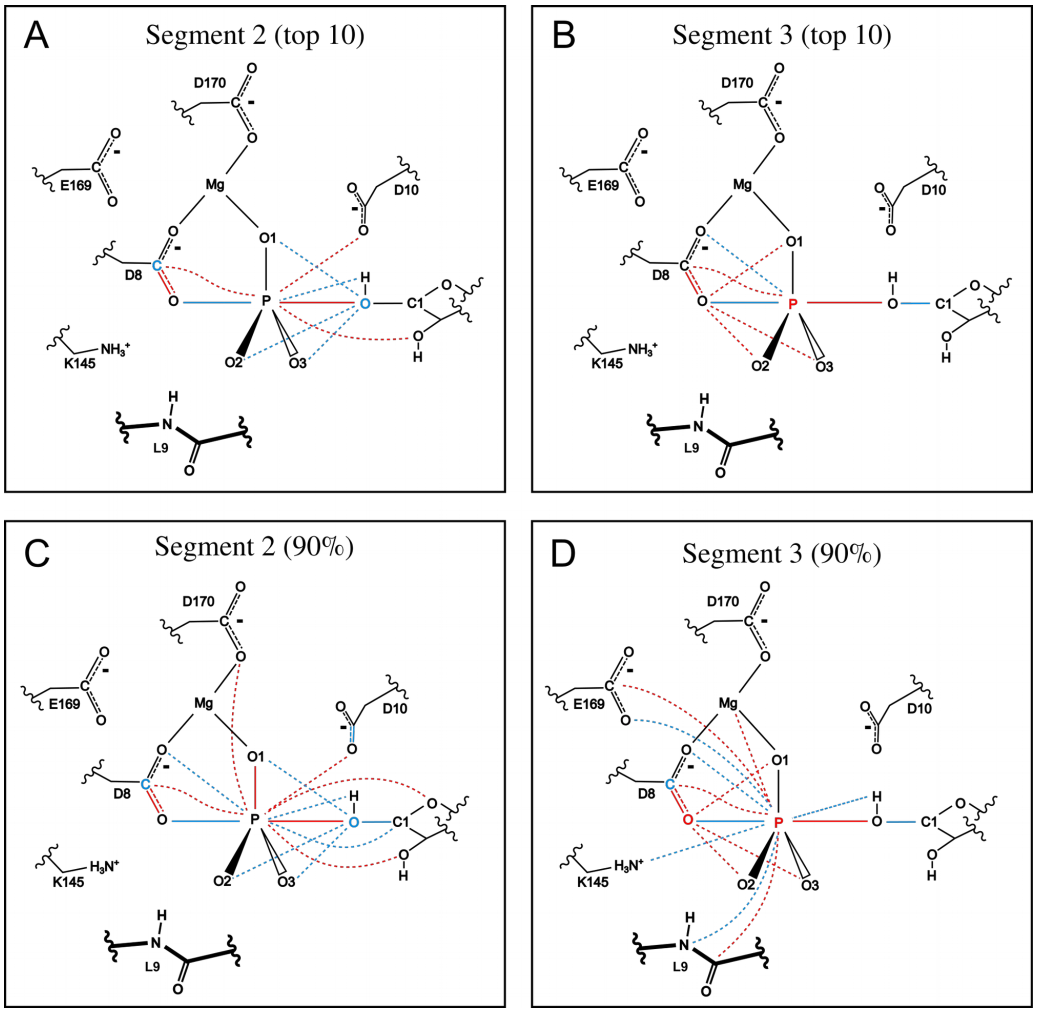


Fig 4.

ANANKE derived REG values which describe the principal inter- and intra-atomic interactions in the phosphoryl group transfer reaction. Solid lines describe interactions with covalent character, whereas dashed lines illustrate interactions, while the coloring of a particular atom denotes a significant self term. The color red (blue) is used to illustrate interactions that increase (decrease) the energy barrier for the phosphoryl transfer reaction.

References

Alkorta, Ibon, Joseph C.R. Thacker, and Paul L.A. Popelier. 2018. “An Interacting Quantum Atom Study of Model SN2 Reactions (X–···CH3X, X = F, Cl, Br, and I).” *Journal of Computational Chemistry* 39 (10): 546–56. https://doi.org/10.1002/jcc.25098.

Bader, R F W. 1990. *Atoms in Molecules: A Quantum Theory*.

Barrozo, Alexandre, Qinghua Liao, Mauricio Esguerra, Gaël Marloie, Jan Florian, Nicholas H. Williams, and Shina C. L. Kamerlin. 2018. “Computer Simulations of the Catalytic Mechanism of Wild–Type and Mutant β-Phosphoglucomutase.” *Organic & Biomolecular Chemistry* 8: 2060–73. https://doi.org/10.1039/C8OB00312B.

Baxter, Nicola J., G. Michael Blackburn, James P. Marston, Andrea M. Hounslow, Matthew J. Cliff, Wolfgang Bermel, Nicholas H. Williams, Florian Hollfelder, David E. Wemmer, and Jonathan P. Waltho. 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): 3952–58. https://doi.org/10.1021/ja078000n.

Baxter, Nicola J, Matthew W Bowler, Tooba Alizadeh, Matthew J Cliff, Andrea M Hounslow, Bin Wu, David B Berkowitz, Nicholas H Williams, G Michael Blackburn, and Jonathan P Waltho. 2010. “Atomic Details of Near-Transition State Conformers for Enzyme Phosphoryl Transfer Revealed by MgF-3 Rather than by Phosphoranes.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (10): 4555–60. https://doi.org/10.1073/pnas.0910333106.

Baxter, Nicola J, Luis F Olguin, Marko Golic, Guoqiang Feng, Andrea M Hounslow, Wolfgang Bermel, G Michael Blackburn, Florian Hollfelder, Jonathan P Waltho, and Nicholas H Williams. 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): 14732–37. https://doi.org/10.1073/pnas.0604448103.

Blanco, M A, A Martin Pendas, and E Francisco. 2005. “Interacting Quantum Atoms: A Correlated Energy Decomposition Scheme Based on the Quantum Theory of Atoms in Molecules.” *Journal of Chemical Theory and Computation* 1 (6): 1096–1109. https://doi.org/10.1021/ct0501093.

Cliff, M J, M W Bowler, A Varga, J P Marston, J Szabó, A M Hounslow, N J Baxter, G M Blackburn, M Vas, and J P Waltho. 2010. “Transition State Analogue Structures of Human Phosphoglycerate Kinase Establish the Importance of Charge Balance in Catalysis.” *J Am Chem Soc* 132: 6507–16. https://doi.org/10.1021/ja100974t.

Dai, Jianying, Lorenzo Finci, Chunchun Zhang, Sushmita Lahiri, Guofeng Zhang, Ezra Peisach, Karen N Allen, and Debra Dunaway-Mariano. 2009. “Analysis of the Structural Determinants Underlying Discrimination between Substrate and Solvent in $β$-Phosphoglucomutase Catalysis.” *Biochemistry* 48 (9): 1984–95. https://doi.org/10.1021/bi801653r.

Dai, Jianying, Liangbing Wang, Karen N. Allen, Peter Radstrom, and Debra Dunaway-Mariano. 2006. “Conformational Cycling in β-Phosphoglucomutase Catalysis: Reorientation of the β-D-

Glucose 1,6-(Bis)Phosphate Intermediate." *Biochemistry* 45 (25): 7818–24. https://doi.org/10.1021/bi060136v.

Elsässer, Brigitta, Silvia Dohmeier-Fischer, and Gregor Fels. 2012. "Theoretical Investigation of the Enzymatic Phosphoryl Transfer of $β$-Phosphoglucomutase: Revisiting Both Steps of the Catalytic Cycle." *Journal of Molecular Modeling* 18 (7): 3169–79. https://doi.org/10.1007/s00894-011-1344-5.

Francisco, E., A. Martín Pendás, and M. A. Blanco. 2006. "A Molecular Energy Decomposition Scheme for Atoms in Molecules." *Journal of Chemical Theory and Computation* 2 (1): 90–102. https://doi.org/10.1021/ct0502209.

Frisch, M J, G W Trucks, H B Schlegel, G E Scuseria, M A Robb, J R Cheeseman, G Scalmani, et al. n.d. "Gaussian∼09 {R}evision {E}.01."

Golicnik, Marko, Luis F Olguin, Guoqiang Feng, Nicola J Baxter, Jonathan P Waltho, Nicholas H Williams, and Florian Hollfelder. 2009. "Kinetic Analysis of Beta-Phosphoglucomutase and Its Inhibition by Magnesium Fluoride." *Journal of the American Chemical Society* 131 (4): 1575–88. https://doi.org/10.1021/ja806421f.

Griffin, Joanna L, Matthew W Bowler, Nicola J Baxter, Katherine N Leigh, Hugh R W Dannatt, Andrea M Hounslow, G Michael Blackburn, Charles Edwin Webster, Matthew J Cliff, and Jonathan P Waltho. 2012. "Near Attack Conformers Dominate β-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): 6910–15. https://doi.org/10.1073/pnas.1116855109.

Himo, Fahmi, Timothy Lovell, Robert Hilgraf, Vsevolod V. Rostovtsev, Louis Noodleman, K. Barry Sharpless, and Valery V. Fokin. 2005. "Copper(I)-Catalyzed Synthesis of Azoles. DFT Study Predicts Unprecedented Reactivity and Intermediates." *Journal of the American Chemical Society* 127 (1): 210–16. https://doi.org/10.1021/ja0471525.

Jin, Yi, Debabrata Bhattasali, Erika Pellegrini, Stephanie M Forget, Nicola J Baxter, Matthew J Cliff, Matthew W Bowler, David L Jakeman, G Michael Blackburn, and Jonathan P Waltho. 2014. "α-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): 12384–89. https://doi.org/10.1073/pnas.1402850111.

Jin, Yi, Robert W. Molt, and G. Michael Blackburn. 2017. "Metal Fluorides: Tools for Structural and Computational Analysis of Phosphoryl Transfer Enzymes." *Topics in Current Chemistry* 375 (2): 1–31. https://doi.org/10.1007/s41061-017-0130-y.

Jin, Yi, Robert W. Molt, Erika Pellegrini, Matthew J. Cliff, Matthew W. Bowler, Nigel G. J. Richards, G. Michael Blackburn, and Jonathan P. Waltho. 2017. "Assessing the Influence of Mutation on GTPase Transition States by Using X-Ray Crystallography, ^{19}F NMR, and DFT Approaches." *Angewandte Chemie International Edition* 56 (33): 9732–35. https://doi.org/10.1002/anie.201703074.

Jin, Yi, Robert W. Molt, Jonathan P. Waltho, Nigel G. J. Richards, and G. Michael Blackburn. 2016. "^{19}F NMR and DFT Analysis Reveal Structural and Electronic Transition State Features for RhoA-Catalyzed GTP Hydrolysis." *Angewandte Chemie International Edition* 55 (10): 3318–22. https://doi.org/10.1002/anie.201509477.

Jin, Yi, Nigel G Richards, Jonathan P Waltho, and G Michael Blackburn. 2017. "Metal Fluorides as Analogues for Studies on Phosphoryl Transfer Enzymes." *Angewandte Chemie - International Edition* 56 (15): 4110–28. https://doi.org/10.1002/anie.201606474.

Johnson, Luke A., Angus Robertson, Nicola Jane Baxter, Clare Trevitt, Claudine Bisson, Yi Jin, Henry Wood, et al. 2018. "Van Der Waals Contact between Nucleophile and Transferring Phosphorus Is Insufficient to Achieve Enzyme Transition State Architecture." *ACS Catalysis*. https://doi.org/10.1021/acscatal.8b01612.

Kamerlin, Shina C L, Pankaz K Sharma, Ram B Prasad, and Arieh Warshel. 2013. "Why Nature Really Chose Phosphate." *Quarterly Reviews of Biophysics* 46 (01): 1–132. https://doi.org/10.1017/S0033583512000157.

Keith, Todd A. 2017. "AIMAll (Version 17.11.14),." TK Gristmill Software, Overland Park KS, USA.

Lad, C, N H Williams, and R Wolfenden. 2003. "The Rate of Hydrolysis of Phosphomonoester Dianions and the Exceptional Catalytic Proficiencies of Protein and Inositol Phosphatases." *Proc Natl Acad Sci U S A* 100: 5607–10.

Lahiri, Sushmita D, Guofeng Zhang, Jianying Dai, Debra Dunaway-Mariano, and Karen N Allen. 2004. "Analysis of the Substrate Specificity Loop of the HAD Superfamily Cap Domain." *Biochemistry* 43 (10): 2812–20. https://doi.org/10.1021/bi0356810.

Lassila, Jonathan K, Jesse G Zalatan, and Daniel Herschlag. 2011. *Biological Phosphoryl-Transfer Reactions: Understanding Mechanism and Catalysis. Annual Review of Biochemistry*. Vol. 80. https://doi.org/10.1146/annurev-biochem-060409-092741.

Marcos, Enrique, Martin J. Field, and Ramon Crehuet. 2010. "Pentacoordinated Phosphorus Revisited by High-Level QM/MM Calculations." *Proteins: Structure, Function and Bioinformatics* 78 (11): 2405–11. https://doi.org/10.1002/prot.22758.

Noodleman, Louis, Timothy Lovell, Wen Ge Han, Jian Li, and Fahmi Himo. 2004. "Quantum Chemical Studies of Intermediates and Reaction Pathways in Selected Enzymes and Catalytic Synthetic Systems." *Chemical Reviews* 104 (2): 459–508. https://doi.org/10.1021/cr020625a.

Outeiral, Carlos, Mark A. Vincent, Ángel Martín Pendás, and Paul L.A. Popelier. 2018. "Revitalizing the Concept of Bond Order through Delocalization Measures in Real Space." *Chemical Science* 9 (25): 5517–29. https://doi.org/10.1039/c8sc01338a.

Pendás, A. Martín, E. Francisco, and M. A. Blanco. 2005. "Two-Electron Integrations in the Quantum Theory of Atoms in Molecules with Correlated Wave Functions." *Journal of Computational Chemistry* 26 (4): 344–51. https://doi.org/10.1002/jcc.20173.

Popelier, P L A. 2000. *Atoms in Molecules. An Introduction*. United Kingdom: Pearson Education.

Popelier, Paul L A. 2005. “Quantum Chemical Topology: On Bonds and Potentials.” In *Intermolecular Forces and Clusters I*, edited by D J Wales, 1–56. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/b135617.

Thacker, Joseph C.R., and Paul L.A. Popelier. 2017. “The ANANKE Relative Energy Gradient (REG) Method to Automate IQA Analysis over Configurational Change.” *Theoretical Chemistry Accounts* 136 (7): 1–13. https://doi.org/10.1007/s00214-017-2113-z.

———. 2018. “Fluorine Gauche Effect Explained by Electrostatic Polarization Instead of Hyperconjugation: An Interacting Quantum Atoms (IQA) and Relative Energy Gradient (REG) Study.” *Journal of Physical Chemistry A* 122 (5): 1439–50. https://doi.org/10.1021/acs.jpca.7b11881.

Vosko, S H, L Wilk, and M Nusair. 1980. “Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis.” *Canadian Journal of Physics* 58 (8): 1200–1211. https://doi.org/10.1139/p80-159.

Webster, Charles Edwin. 2004. “High-Energy Intermediate or Stable Transition State Analogue: Theoretical Perspective of the Active Site and Mechanism of $β$-- Phosphoglucomutase.” *Journal of the American Chemical Society* 126 (22): 6840–41. https://doi.org/10.1021/ja049232e.

Zhang, Guofeng, Jianying Dai, Liangbing Wang, Debra Dunaway-Mariano, Lee W. Tremblay, and Karen N. Allen. 2005. “Catalytic Cycling in??-Phosphoglucomutase: A Kinetic and Structural Analysis.” *Biochemistry* 44 (27): 9404–16. https://doi.org/10.1021/bi050558p.

Supporting information

X-ray, NMR and QM approaches reveal the relationship between protein conformational change, proton transfer, and phosphoryl transfer in an archetypal enzyme.

Angus J. Robertson,[1] Alex L. Wilson,[2] Matthew J. Cliff,[2] Paul L.A. Popelier,[2] 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

* E-mail for J.P.W.: j.waltho@sheffield.ac.uk

Table of contents

1. Protein production and purification.

The *pgmB* gene from Lactococcus lactis together with the *pgmB* gene containing the D10N mutation were cloned in pET22b+ expression vectors and used to express βPGM_{WT} and βPGM_{D10N} proteins in *E. coli* strain BL21(DE3). One liter cell cultures were grown to log phase in M9 media (with ^{15}N isotopic enrichment), induced with 1 mM IPTG and grown for a further 16 h at 25 ◦ C. Perdeuterated protein preparations for enzyme dynamics were grown in 100% D_2O and included >97% ^{2}H isotope labelling of the carbon source, glucose (CortecNet). Cells were harvested by centrifugation at 10,000 rpm for 10 min at 4 °C, decanted and frozen at –80 °C. Cell pellets were resuspended in ice-cold standard native buffer (50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3) supplemented with one tablet of cOmplete TM protease inhibitor cocktail (Roche). The cell suspension was lysed on ice by sonication for 6 cycles of pulsation for 20 s with 60 s cooling intervals. The cell lysate was then separated by ultracentrifugation (Beckman Coulter Avanti centrifuge) at 24,000 rpm for 35 min at 4 °C to remove insoluble matter. The cleared cell lysate was filtered using a 0.2 µm syringe filter and loaded onto a DEAE-Sepharose fast flow ion exchange column connected to an ÄKTA purification system that had been washed previously with 1 column volume of 6 M guanidine hydrochloride (GuHCl), 1 column volume of 1 M NaOH and equilibrated with > 2 column volumes of standard native buffer. Following extensive washing, proteins bound to the DEAE-Sepharose column were eluted with a gradient of 0 to 100% standard native buffer containing 0.5 M NaCl. Fractions containing βPGM were checked for purity using SDS-PAGE, were pooled together and concentrated by Vivaspin (10 kDa MWCO). The protein sample was filtered using a 0.2 µm syringe filter and loaded onto a prepacked Hiload 26/60 Superdex 75 size-exclusion column connected to an ÄKTA purification system that had been pre-equilibrated with filtered and degassed standard native buffer containing 1 M NaCl. βPGM eluted as a single peak and fractions containing βPGM were checked for purity using SDS-PAGE, were pooled together, buffer exchanged into standard native buffer and concentrated to 1 mM by Vivaspin (10 kDa MWCO) for storage as 1 mL aliquots at –20 °C. The overall yield for βPGM was ca. 60 mg protein from 1 L of bacterial culture.

2. NMR spectroscopy general methods

$^{2}H^{15}N$ backbone dynamics and 3D multi-dimensional heteronuclear experiments for dynamic characterisation and resonance assignment of the βPGM_{D10N}:AlF_4:G6P complex were acquired at 298 K using 1 mM $^{2}H^{15}N$ or $^{1}H^{15}N^{13}C$-labeled apo-βPGM_{D10N} in standard NMR buffer (50 mM K+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3 with 10% (v/v) D_2O and 1 mM TSP) containing 5mM $AlCl_3$, 20mM NaF, and 20 mM G6P. Reference ^{19}F spectra for βPGM_{WT} and βPGM_{D10N} AlF_4 TSA complexes with either G1P or G6P ligands were acquired at 298 K using 0.5 – 1 mM ^{15}N-βPGM in standard NMR buffer also containing 5mM $AlCl_3$, 20mM NaF, 20mM βG1P/G6P.

3. Backbone assignment of βPGM_{D10N}:AlF_4:G6P complex

The βPGM_{D10N}:AlF_4:G6P complex was generated using 1 mM $^{1}H^{15}N^{13}C$-labeled βPGM_{D10N} in standard NMR buffer (50 mM K+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3 with 10% (v/v) D_2O and 1 mM TSP) with the addition of 5mM $AlCl_3$, 20mM NaF, and 20 mM G6P. Multi-dimensional heteronuclear NMR spectra for backbone resonance assignment of the $^{2}H,^{15}N,^{13}C$-labeled βPGM_{D10N}:AlF_4:G6P complex were acquired at 298 K on either a Bruker 800 MHz Avance III spectrometer equipped with a TCI cryoprobe and z-axis gradients (MIB) or Bruker 800 MHz Avance spectrometer equipped with a TXI probe and z-axis gradinents (Sheffield). The standard suite of $^{1}H^{15}N$-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 TopSpin3 and multidimensional decomposition (Hyberts et al. 2012). Backbone resonance assignments of the βPGM_{D10N}:AlF_4:G6P complex were obtained using a simulated annealing algorithm employed by the *asstools* assignment program (Reed et al. 2003). The backbone assignment is available from the BMRB with accession code (BMRB: 27697).

4. Chemical shift comparisons between the βPGM_{WT}:AlF_4:G6P and βPGM_{D10N}:AlF_4:G6P TSA complexes.

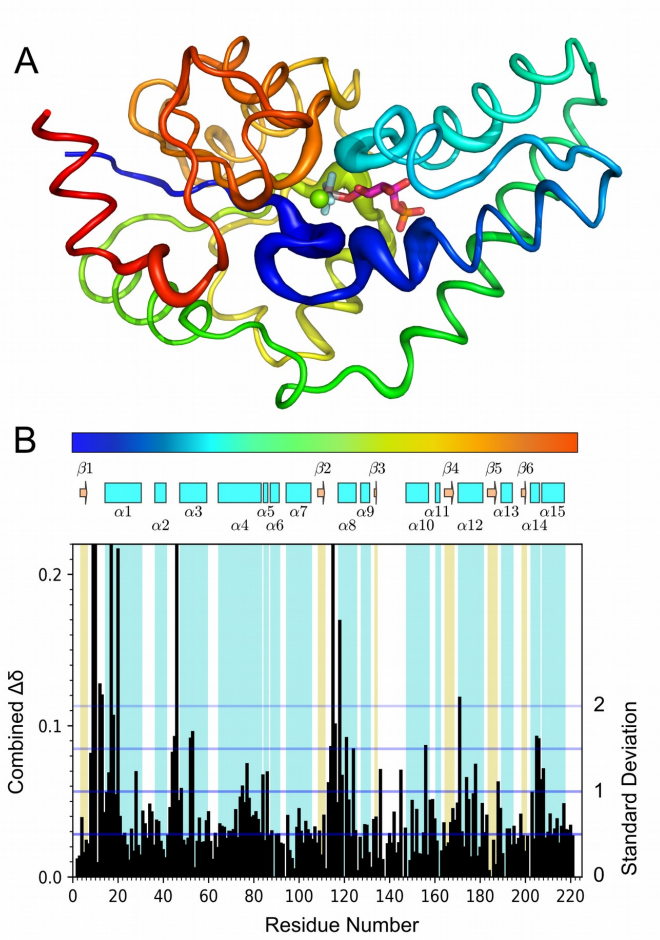


Fig. S1

The combined chemical shift differences between the backbone assignments of WT and D10N variants of βPGM complexed with AlF_4^- and G6P. The combined chemical shift was generated using the equation presented in Williamson 2013 (Williamson 2013) , with a the alpha scaling factor set to 0.133 to account for the difference in spectral widths of the ^1H and ^{15}N dimensions. **A)** The magnitude of the combined chemical shift change as a b-factor putty on a crystal structure of the βPGM_{WT}:AlF_4:G6P complex (**PDB: 2WF6**; (Baxter et al. 2010)) structure, with the polypeptide chain coloured from blue to red (N-term to C-term) and with the G6P ligand represented as pink sticks. **B)** The magnitude of the combined chemical shift difference between the two complexes by residue, with secondary structure indicated and a colour bar that refers to the colouring of the structure in part A.

5. Comparison of Cβ chemical shift of aspartates in βPGM complexes

Complex	WT:BeF_3	WT:MgF_3:G6P	WT:AlF_4:G6P	WT:BeF_3:G6P
BMRB ID	17851	7234	15467	17852
Residue:				
10	40.244	38.331	38.444	-
180	38.648	38.771	38.742	38.736
149	38.971	38.976	38.976	38.951
78	39.130	39.238	39.212	39.177
61	39.412	39.444	39.389	39.319
51	39.489	39.476	39.617	39.459
58	39.560	39.502	39.446	39.488
102	39.593	39.561	39.559	39.602
197	39.558	39.650	39.621	39.603
15	43.166	39.855	39.911	41.010
86	40.819	40.733	40.675	40.706
196	41.111	41.085	41.084	41.060
137	41.396	41.259	–	41.356
91	40.935	41.319	41.261	41.129
203	41.524	41.376	41.379	41.482
170	42.545	41.405	41.358	41.764
193	41.815	41.611	41.609	41.633
37	41.697	41.791	41.786	41.801
133	42.371	42.170	42.136	42.184
8	43.163	43.811	44.011	43.285

Table S1. Cβ carbon chemical shifts for all of the assigned aspartate residues in several βPGM complexes, both open (BeF_3^- complex; (Griffin et al. 2012)) and closed (MgF_3^- and AlF_4^- complexes; (Baxter et al. 2010)). Residues were ordered by chemical shift in the MgF_3:G6P complex.

6. Protein backbone relaxation measurments and modelfree analysis

NMR Relaxation Measurements on WT:AlF4:G6P and D10N:AlF4:G6P complexes.

βPGM$_{WT}$:AlF$_4$:G6P and βPGM$_{D10N}$:AlF$_4$:G6P samples for ^{15}N fast timescale relaxation measurements were performed using ^{2}H^{15}N labelled βPGM enzyme in 5mm Shigemi D_2O matched tubes. Sample conditions for βPGM$_{WT}$:AlF$_4$:G6P complex - 1mM βPGM$_{WT}$, 2mM $AlCl_3$, 10mM NaF, 10mM G6P, 5mM $MgCl_2$, 2mM NaN_3, 1mM TSP. Sample conditions for the βPGM$_{D10N}$:AlF$_4$:G6P complex were 1mM βPGM$_{D10N}$, 5mM $AlCl_3$, 20mM NaF, 20mM G6P, 5mM $MgCl_2$, 2mM NaN_3, 1mM TSP.

Experiments were acquired using a Bruker 600 MHz Avance DRX spectrometer equipped with a 5-mm TXI cryoprobe and z-axis gradients (Sheffield), a Bruker 800 MHz Avance spectrometer equipped with a 5-mm TXI probe and z-axis gradients (Sheffield), and for the βPGM$_{D10N}$:AlF$_4$:G6P complex an additional data series using a 950MHz Avance III spectrometer equipped with a TCI probe and z-axis gradients (Crick Institute, London).

Spin-lattice ^{15}N relaxation rates (R1), rotating frame ^{15}N relaxation rates (R1ρ) and heteronuclear steady-state ^{15}N-{^{1}H} NOE (hNOE) values were obtained using interleaved TROSY-readout pulse sequences (Lakomek, Ying, and Bax 2012). Temperature compensation was applied in the R1 experiment by incorporating a spin-lock pulse placed off resonance in the inter-scan delay, equal to the longest spin-lock time and the RF power of the R1ρ experiment. Relaxation delays of 0, 80, 240, 400, 400, 640, 800, 1200, 1760, and 2400 ms were used to calculate R1, and delays of 1, 20, 20, 30, 40, 60, 90, 110, 150, and 200 ms were used to calculate R1ρ at 600 MHz and 800 MHz for both complexes. Relaxation delays of 20, 40, 80, 240, 400, 640, 800, 1200, 1200, 1760, 2400, 3200, 4800, 6400, ms were used to calculate R1, and delays of 1, 5, 5, 10, 15, 20, 20, 40, 60, 90, 110, 140, 160, 200 ms were used to calculate R1ρ for the βPGM$_{D10N}$:AlF$_4$:G6P at 950MHz. The inter scan delay was 3.5 s and the strength of the RF spin-lock field during R1ρ measurement was 1400 Hz at 600MHz, 1866.7 at 800MHz, and 1500 HZ at 950MHz. For the hNOE measurement, two interleaved experiments were acquired with relaxation delays of 10s.

Experiments were processed in NMRpipe (Delaglio et al. 1995) using a squared sine bell window function, without linear prediction in either dimension. R1 and R2 values were determined in PINT (Niklasson et al. 2017) by fitting the integral of the assigned peak to a decaying exponential function across the relaxation series . R2 values were calculated in PINT from fitted R1 values. hNOE values were also fitted in PINT by calculating the difference in peak integral between saturated and unsaturated spectra.

Model free analysis

Model free analysis (Lipari and Szabo 1982a, 1982b; Halle 2009; Halle and Wennerström 1981; Halle and Carlström 1981; Halle et al. 1981) was performed using *relax* (E J d'Auvergne and Gooley 2008a, 2008b, Edward J. d'Auvergne and Gooley 2007, 2006, 2003; Bieri, d'Auvergne, and Gooley 2011). R1, R1ρ, and hNOE values at 600MHz and 800MHz were used for the βPGM_{WT}:AlF_4:G6P complex with backbone amide coordinate geometry provided by a crystal structure of the βPGM_{WT}:AlF_4:G6P complex (**PDB: 2WF6**; (Baxter et al. 2010)). R1, R1ρ, and hNOE values at 600MHz and 800MHz, and 950MHz were used for the βPGM_{D10N}:AlF_4:G6P complex with backbone amide coordinate geometry provided by a crystal structure of the βPGM_{D10N}:AlF_4:G6P (**PDB: 5OK2**; (Johnson et al. 2018)). Model free analysis was performed in *relax* using models m0-m5 in both complexes (ie. without using the extended MF formula presented by Clore and Co-workers (Clore et al. 1990)), however the three terminal residues were excluded due to their absence in both crystal structures. The resulting ellipsiodal diffusion tensors closely matched the geometry of the input crystal structures.

βPGM$_{WT}$:AlF$_4$:G6P complex

Residue	Model	S2	S2f	te < 100 or tf	te > 100 or ts	Rex (800.343762 MHz)
				ps	ps	s^-1
1						
2						
3	m1	0.826±0.003				
4	m3	0.845±0.002				2.442±0.152
5	m3	0.831±0.002				1.006±0.136
6	m3	0.829±0.002				2.648±0.092
7	m3	0.814±0.002				1.635±0.094
8	m3	0.826±0.003				3.116±0.111
9	m3	0.857±0.004				2.391±0.213
10	m3	0.842±0.008				5.007±0.349
11	m1	0.882±0.002				
12	m3	0.845±0.002				1.886±0.221
13	m4	0.801±0.005		2.02±0.79		1.370±0.248
14	m3	0.862±0.002				1.027±0.144
15	m4	0.788±0.002		24.35±0.65		0.891±0.093
16	m4	0.801±0.002		8.87±0.71		3.094±0.131
17	m4	0.890±0.002		13.97±1.36		2.448±0.165
18	m4	0.852±0.005		9.20±1.15		1.417±0.272
19						
20	m4	0.895±0.029		7.39±4.09		1.247±0.735
21	m3	0.835±0.001				2.838±0.096
22	m3	0.822±0.001				4.407±0.093
23	m3	0.848±0.001				4.255±0.102
24	m3	0.820±0.001				4.210±0.070
25	m3	0.832±0.002				4.760±0.145
26	m4	0.829±0.002		1.71±0.89		0.620±0.152
27						
28	m3	0.732±0.001				11.036±0.082
29						
30	m4	0.734±0.001		3.68±0.57		6.054±0.382
31	m4	0.795±0.002		4.13±0.71		2.770±0.131
32	m5	0.791±0.004	0.841±0.003		1233.13±69.43	
33	m2	0.821±0.002		7.90±0.93		
34	m5	0.769±0.011	0.817±0.010		1113.52±162.00	
35	m5	0.458±0.009	0.736±0.010		1628.99±26.65	
36	m4	0.837±0.002		5.50±0.92		0.851±0.137
37	m4	0.768±0.001		6.16±0.64		0.985±0.060
38	m4	0.845±0.019		2.77±1.30		1.011±0.678
39	m4	0.809±0.002		10.75±0.76		1.028±0.125
40	m3	0.871±0.002				1.563±0.138
41	m4	0.850±0.002		3.59±1.16		2.029±0.120
42	m3	0.785±0.002				5.736±0.140
43	m3	0.834±0.001				1.960±0.107
44	m4	0.818±0.002		6.28±0.93		1.068±0.132
45	m4	0.810±0.002		4.43±0.93		2.004±0.098
46	m2	0.851±0.002		4.08±1.09		
47	m4	0.801±0.001		5.47±0.82		2.715±0.069
48	m5	0.837±0.003	0.859±0.002		610.85±61.02	
49	m4	0.850±0.002		5.09±1.22		3.159±0.124
50	m4	0.814±0.002		4.83±0.89		2.454±0.117
51	m4	0.825±0.001		5.94±0.87		1.917±0.190
52	m4	0.854±0.002		2.55±1.21		1.175±0.169
53	m3	0.825±0.002				4.081±0.090
54	m4	0.842±0.001		1.80±1.05		2.593±0.092

Residue	Model	S2	S2f	te < 100 or tf	te > 100 or ts	Rex (800.343762 MHz)
				ps	ps	s^-1
55	m3	0.836±0.001				2.723±0.086
56	m3	0.835±0.002				2.176±0.104
57	m3	0.825±0.002				3.936±0.122
58	m4	0.842±0.003		2.67±1.04		2.420±0.122
59	m2	0.880±0.002		6.32±1.38		
60	m4	0.800±0.002		3.55±0.84		1.927±0.104
61	m5	0.718±0.009	0.789±0.007		1053.73±45.18	
62	m5	0.821±0.002	0.846±0.002		409.12±39.37	
63	m5	0.785±0.002	0.824±0.003		403.44±55.39	
64	m5	0.733±0.001	0.772±0.002		206.41±23.03	
65	m4	0.835±0.002		7.05±0.94		0.733±0.097
66	m2	0.873±0.011		6.47±1.57		
67	m2	0.856±0.004		3.99±1.15		
68	m2	0.854±0.002		5.28±1.16		
69	m4	0.833±0.001		2.67±1.06		2.593±0.087
70	m4	0.821±0.002		2.37±0.92		1.775±0.070
71	m4	0.850±0.002		3.06±1.06		1.396±0.104
72	m3	0.860±0.001				0.206±0.129
73	m3	0.837±0.002				3.105±0.152
74	m4	0.843±0.002		2.23±1.08		1.747±0.161
75	m4	0.853±0.002		5.90±1.12		0.596±0.191
76	m3	0.863±0.002				1.126±0.102
77	m3	0.861±0.001				2.342±0.124
78	m5	0.836±0.002	0.850±0.002		793.29±122.02	
79	m2	0.869±0.002		2.10±1.26		
80	m4	0.828±0.002		1.59±1.00		2.757±0.118
81	m3	0.863±0.001				1.362±0.095
82	m4	0.854±0.002		2.70±1.14		1.409±0.121
83	m5	0.834±0.003	0.847±0.005		512.86±111.53	
84	m4	0.772±0.001		11.66±0.62		3.396±0.140
85	m4	0.859±0.002		12.34±0.97		1.445±0.117
86	m5	0.760±0.002	0.844±0.002		894.30±23.80	
87	m4	0.825±0.002		14.77±0.80		1.635±0.092
88	m4	0.835±0.001		6.35±1.04		2.396±0.078
89						
90	m1	0.867±0.007				
91	m4	0.810±0.001		3.64±0.95		1.923±0.107
92	m3	0.875±0.002				2.227±0.102
93	m3	0.889±0.004				1.863±0.141
94						
95	m4	0.831±0.016		16.51±2.20		3.681±0.552
96	m4	0.838±0.002		2.31±0.97		3.429±0.097
97	m3	0.863±0.001				2.182±0.102
98	m3	0.812±0.002				5.021±0.178
99	m3	0.840±0.001				2.543±0.179
100	m3	0.862±0.002				1.124±0.127
101	m3	0.839±0.002				2.280±0.097
102	m3	0.830±0.002				3.398±0.087
103	m3	0.853±0.003				2.720±0.132
104	m3	0.853±0.001				1.268±0.102
105						
106	m1	0.871±0.002				
107	m4	0.867±0.002		3.63±1.36		0.445±0.149
108	m4	0.829±0.002		2.14±0.99		1.347±0.130
109	m3	0.838±0.001				1.822±0.069

Residue	Model	S2	S2f	te < 100 or tf	te > 100 or ts	Rex (800.343762 MHz)
				ps	ps	s^-1
110	m3	0.778±0.001				2.386±0.080
111	m3	0.844±0.002				3.820±0.097
112	m3	0.826±0.002				2.776±0.084
113	m3	0.837±0.001				3.083±0.096
114	m3	0.817±0.001				2.131±0.122
115	m0					
116	m3	0.844±0.002				2.367±0.198
117						
118	m2	0.851±0.003		4.67±1.14		
119	m3	0.849±0.002				2.247±0.131
120						
121	m3	0.839±0.005				0.861±0.186
122	m3	0.865±0.003				1.143±0.142
123	m3	0.827±0.002				3.388±0.138
124	m3	0.851±0.002				2.008±0.107
125						
126	m4	0.813±0.003		2.55±0.91		1.173±0.137
127	m4	0.854±0.003		4.43±1.09		2.074±0.167
128	m3	0.807±0.002				4.098±0.133
129	m5	0.856±0.003	0.878±0.003		732.07±94.45	
130	m2	0.853±0.013		8.02±1.34		
131	m4	0.822±0.002		1.60±0.89		3.416±0.106
132	m4	0.822±0.001		3.92±0.90		0.876±0.087
133	m3	0.834±0.002				0.424±0.086
134	m3	0.801±0.002				1.760±0.086
135	m4	0.801±0.004		1.65±0.85		0.625±0.137
136	m3	0.876±0.002				1.062±0.084
137	m4	0.748±0.002		8.05±0.55		1.824±0.080
138						
139	m4	0.760±0.012		4.45±0.87		5.921±0.471
140	m4	0.775±0.002		11.65±0.63		0.464±0.095
141	m5	0.658±0.001	0.765±0.001		537.22±6.02	
142	m5	0.575±0.002	0.803±0.005		842.29±9.49	
143	m5	0.619±0.001	0.738±0.001		497.35±4.22	
144	m2	0.799±0.003		9.92±0.71		
145	m5	0.839±0.003	0.874±0.002		517.74±33.52	
146						
147	m3	0.916±0.001				2.587±0.091
148						
149	m3	0.898±0.001				0.868±0.064
150	m3	0.813±0.001				3.668±0.168
151	m5	0.794±0.008	0.828±0.006		645.36±68.00	
152	m3	0.803±0.002				3.911±0.132
153	m3	0.855±0.009				1.925±0.333
154	m3	0.842±0.002				2.605±0.138
155	m3	0.813±0.002				3.738±0.139
156						
157	m3	0.816±0.002				4.199±0.226
158	m3	0.810±0.002				1.682±0.141
159	m4	0.822±0.003		1.50±0.81		3.266±0.213
160	m3	0.834±0.002				2.644±0.089
161	m4	0.802±0.002		5.57±0.68		3.771±0.128
162						
163	m2	0.831±0.003		14.07±1.07		

Residue	Model	S2	S2f	te < 100 or tf	te > 100 or ts	Rex (800.343762 MHz)
				ps	ps	s^-1
164	m4	0.801±0.002		1.76±0.76		3.683±0.131
165	m3	0.847±0.002				0.833±0.150
166	m3	0.886±0.004				2.891±0.190
167	m3	0.833±0.003				1.572±0.116
168	m3	0.850±0.002				1.478±0.083
169	m3	0.809±0.002				3.461±0.087
170	m4	0.810±0.002		1.65±0.90		1.102±0.116
171	m4	0.880±0.003		3.64±1.50		1.649±0.277
172	m4	0.840±0.002		2.08±0.99		2.080±0.134
173	m3	0.908±0.015				0.812±0.456
174	m3	0.857±0.003				1.661±0.131
175	m5	0.776±0.002	0.817±0.003		883.51±72.78	
176	m3	0.864±0.002				1.945±0.130
177	m3	0.855±0.002				2.427±0.080
178	m3	0.816±0.001				3.520±0.123
179	m3	0.854±0.004				2.586±0.311
180	m3	0.846±0.002				1.539±0.140
181	m3	0.827±0.002				0.808±0.170
182	m3	0.806±0.001				3.549±0.185
183	m3	0.838±0.004				4.281±0.233
184						
185						
186	m3	0.816±0.002				0.990±0.101
187	m3	0.837±0.001				2.408±0.082
188	m3	0.859±0.002				1.979±0.096
189	m3	0.832±0.002				2.478±0.140
190	m4	0.811±0.001		5.21±0.76		1.161±0.059
191						
192	m2	0.851±0.002		7.30±1.18		
193	m4	0.806±0.002		7.19±0.78		2.618±0.152
194	m2	0.813±0.002		5.82±0.84		
195	m4	0.814±0.002		10.10±0.85		1.024±0.090
196	m5	0.726±0.002	0.769±0.002		437.06±27.10	
197	m5	0.851±0.003	0.875±0.003		578.13±72.73	
198	m5	0.715±0.002	0.744±0.001		469.13±30.97	
199	m4	0.837±0.002		7.65±1.06		0.747±0.103
200	m4	0.821±0.001		2.96±0.97		1.898±0.069
201	m4	0.823±0.001		3.57±0.87		2.381±0.075
202						
203	m4	0.810±0.001		3.13±0.77		2.929±0.080
204	m4	0.821±0.002		3.17±0.93		2.253±0.097
205	m4	0.870±0.002		3.78±1.36		3.957±0.207
206	m3	0.843±0.002				5.119±0.137
207	m5	0.866±0.002	0.879±0.009		372.52±104.61	
208	m4	0.724±0.001		2.03±0.58		2.742±0.059
209	m3	0.863±0.001				0.575±0.074
210	m4	0.837±0.002		2.37±1.04		0.737±0.078
211	m2	0.873±0.002		5.51±1.40		
212	m1	0.886±0.002				
213	m3	0.829±0.001				1.518±0.102
214	m5	0.822±0.005	0.838±0.004		1164.56±219.9	
215	m2	0.863±0.001		5.11±1.18		
216	m3	0.877±0.008				1.122±0.362
217	m5	0.790±0.010	0.831±0.007		1429.17±140.7	
218	m5	0.826±0.003	0.856±0.002		582.74±45.14	

βPGM$_{D10N}$:AlF$_4$:G6P complex

Residue	Model	S2	S2f	te < 100 or tf	te > 100 or ts	Rex (950.454467115 MHz)
				ps	ps	s^{-1}
1						
2						
3	m2	0.788±0.004		2.89±0.58		
4	m3	0.837±0.005				1.940±0.225
5	m2	0.803±0.003		3.38±0.64		
6	m1	0.812±0.002				
7	m4	0.797±0.003		1.39±0.62		0.805±0.161
8	m1	0.831±0.002				
9						
10						
11	m1	0.848±0.003				
12	m2	0.832±0.004		1.16±0.70		
13	m2	0.779±0.004		2.17±0.52		
14	m4	0.823±0.004		6.72±0.77		0.681±0.181
15	m5	0.747±0.002	0.787±0.002		280.51±23.69	
16	m5	0.740±0.003	0.778±0.003		321.41±27.88	
17	m5	0.878±0.003	0.891±0.006		252.88±89.21	
18	m5	0.795±0.007	0.827±0.006		1102.05±115.37	
19	m4	0.757±0.002		4.14±0.43		2.320±0.153
20	m4	0.806±0.004		6.54±0.57		2.525±0.298
21	m4	0.810±0.002		2.15±0.59		0.753±0.161
22	m2	0.835±0.002		3.58±0.66		
23	m4	0.811±0.002		2.06±0.59		1.788±0.196
24	m4	0.815±0.003		3.35±0.62		1.504±0.165
25	m4	0.801±0.003		0.91±0.56		5.088±0.191
26	m2	0.803±0.002		4.51±0.55		
27	m3	0.810±0.004				1.266±0.244
28	m3	0.799±0.003				2.613±0.218
29	m4	0.786±0.003		2.41±0.50		2.313±0.204
30	m4	0.781±0.004		6.99±0.49		1.763±0.365
31	m2	0.775±0.002		5.26±0.44		
32	m5	0.731±0.004	0.793±0.003		1101.91±42.02	
33	m5	0.736±0.006	0.779±0.004		846.77±53.38	
34	m5	0.768±0.010	0.799±0.011		438.74±94.99	
35	m5	0.460±0.007	0.555±0.010		763.39±40.07	
36	m4	0.813±0.002		7.28±0.60		0.443±0.143
37	m2	0.740±0.001		5.92±0.40		
38	m5	0.861±0.007	0.876±0.010		352.68±199.18	
39	m5	0.790±0.003	0.810±0.002		338.20±56.89	
40	m1	0.843±0.002				
41	m2	0.827±0.001		2.16±0.75		
42						
43	m2	0.828±0.003		3.76±0.72		
44	m4	0.802±0.004		5.55±0.63		0.383±0.188
45	m4	0.788±0.003		6.32±0.55		0.768±0.174
46	m1	0.824±0.004				
47	m2	0.811±0.002		6.89±0.64		
48	m5	0.802±0.004	0.825±0.003		439.94±57.87	
49	m4	0.849±0.003		7.09±0.84		0.865±0.217
50	m2	0.807±0.001		6.67±0.58		
51	m2	0.823±0.004		3.43±0.67		
52						
53	m4	0.822±0.003		1.83±0.72		1.510±0.143
54	m2	0.833±0.002		2.41±0.71		
55	m1	0.828±0.003				

194

195

Residue	Model	S2	S2f	te < 100 or tf	te > 100 or ts	Rex (950.454467115 MHz)
				ps	ps	s^{-1}
56	m2	0.807±0.003		1.93±0.64		
57	m4	0.824±0.004		2.61±0.66		1.209±0.194
58	m4	0.819±0.004		2.78±0.67		1.985±0.227
59	m4	0.849±0.003		7.28±0.81		0.912±0.139
60	m4	0.778±0.004		7.37±0.55		1.484±0.203
61	m5	0.743±0.010	0.815±0.008		579.43±29.20	
62	m5	0.776±0.004	0.807±0.002		440.30±38.48	
63	m5	0.747±0.002	0.791±0.003		329.08±28.92	
64	m5	0.695±0.002	0.744±0.002		254.29±13.99	
65	m2	0.801±0.002		9.89±0.79		
66	m4	0.826±0.033		3.61±1.13		1.712±1.062
67	m2	0.811±0.004		6.13±0.60		
68	m2	0.813±0.003		6.82±0.66		
69						
70	m5	0.710±0.010	0.779±0.003		6739.16±2140.12	
71						
72	m4	0.819±0.001		3.20±0.69		0.841±0.191
73	m4	0.823±0.003		3.14±0.72		3.043±0.144
74	m4	0.819±0.004		3.08±0.66		0.909±0.195
75	m2	0.814±0.004		7.80±0.65		
76	m3	0.841±0.004				1.783±0.191
77	m2	0.859±0.002		1.45±0.87		
78	m2	0.778±0.002		7.60±0.76		
79	m4	0.842±0.004		4.46±0.85		0.278±0.169
80	m4	0.801±0.004		5.73±0.63		1.766±0.216
81	m2	0.852±0.003		4.19±0.86		
82	m4	0.843±0.004		1.60±0.79		0.524±0.164
83						
84	m5	0.762±0.004	0.790±0.004		200.14±36.84	
85	m2	0.832±0.002		9.42±0.68		
86	m5	0.696±0.002	0.787±0.003		867.20±20.31	
87	m4	0.794±0.002		12.12±0.53		1.011±0.104
88	m2	0.829±0.002		5.22±0.75		
89						
90	m2	0.851±0.009		6.31±0.99		
91	m2	0.806±0.003		4.43±0.62		
92						
93	m2	0.847±0.003		3.88±0.94		
94						
95	m5	0.913±0.007	0.957±0.007		365.26±52.89	
96	m3	0.821±0.002				1.664±0.130
97	m1	0.844±0.004				
98	m3	0.805±0.008				3.015±0.399
99						
100	m2	0.843±0.002		1.50±0.77		
101	m4	0.813±0.002		1.97±0.70		0.702±0.186
102	m3	0.820±0.003				0.995±0.147
103	m3	0.830±0.003				1.130±0.184
104	m3	0.786±0.003				3.303±0.193
105						
106	m2	0.834±0.002		3.95±0.65		
107	m2	0.836±0.002		6.06±0.79		
108	m2	0.762±0.007		4.73±0.56		
109	m2	0.827±0.002		2.79±0.72		
110	m2	0.772±0.001		1.64±0.51		

197

198

199

Residue	Model	S2	S2f	te < 100 or tf	te > 100 or ts	Rex (950.454467115 MHz)
				ps	ps	s^{-1}
111	m1	0.845±0.003				
112	m3	0.811±0.002				0.391±0.093
113	m2	0.811±0.002		3.28±0.69		
114	m1	0.806±0.002				
115						
116	m4	0.805±0.002		3.27±0.58		2.943±0.133
117	m2	0.789±0.003		4.28±0.52		
118	m2	0.822±0.004		2.49±0.71		
119	m2	0.795±0.001		0.96±0.55		
120						
121	m4	0.810±0.003		3.29±0.62		1.168±0.154
122	m3	0.835±0.003				0.461±0.195
123						
124	m4	0.826±0.004		3.59±0.72		1.045±0.228
125	m4	0.821±0.003		5.59±0.70		1.222±0.231
126	m2	0.781±0.002		3.75±0.52		
127	m2	0.836±0.004		8.60±0.73		
128	m4	0.782±0.003		4.35±0.48		3.201±0.202
129	m2	0.822±0.003		8.64±0.70		
130	m2	0.828±0.014		3.25±0.86		
131	m4	0.804±0.002		4.08±0.55		1.740±0.112
132	m2	0.792±0.002		5.26±0.52		
133	m2	0.790±0.003		1.07±0.55		
134	m1	0.794±0.002				
135	m2	0.760±0.002		1.98±0.54		
136	m2	0.831±0.002		4.09±0.77		
137	m4	0.723±0.002		9.44±0.39		1.118±0.110
138						
139	m4	0.803±0.002		10.86±0.58		1.027±0.137
140	m2	0.753±0.002		9.44±0.44		
141	m5	0.665±0.002	0.768±0.002		385.19±7.69	
142	m5	0.582±0.002	0.675±0.002		323.77±7.14	
143	m5	0.611±0.002	0.725±0.001		357.13±4.61	
144						
145						
146						
147	m2	0.881±0.002		4.37±1.00		
148						
149	m1	0.874±0.002				
150	m3	0.816±0.003				1.582±0.230
151	m4	0.829±0.003		2.85±0.75		1.594±0.318
152	m4	0.779±0.002		3.38±0.52		3.242±0.233
153						
154	m3	0.826±0.003				0.477±0.180
155	m4	0.800±0.004		0.91±0.57		2.119±0.218
156	m4	0.809±0.004		1.95±0.59		0.719±0.196
157	m2	0.827±0.005		5.25±0.73		
158	m1	0.747±0.002				
159	m3	0.790±0.005				1.446±0.244
160	m3	0.798±0.002				1.847±0.202
161	m2	0.810±0.001		5.61±0.55		
162						
163	m2	0.756±0.003		5.38±0.52		
164	m3	0.815±0.003				0.370±0.174
165	m4	0.800±0.005		2.47±0.62		0.650±0.260

202

Residue	Model	S2	S2f	te < 100 or tf	te > 100 or ts	κ (950.454467115 M
				ps	ps	s^{-1}
166						
167	m2	0.803±0.003		1.13±0.58		
168	m1	0.813±0.003				
169	m3	0.808±0.003				0.511±0.120
170	m2	0.780±0.002		0.91±0.53		
171	m2	0.871±0.004		4.78±1.00		
172	m4	0.824±0.004		4.84±0.68		0.764±0.204
173	m4	0.856±0.003		3.52±0.86		0.388±0.175
174	m3	0.828±0.007				0.899±0.279
175	m4	0.818±0.003		2.06±0.65		2.144±0.205
176	m4	0.840±0.004		1.46±0.78		3.154±0.206
177	m3	0.830±0.002				1.349±0.158
178	m3	0.797±0.002				1.862±0.167
179						
180	m2	0.830±0.003		6.36±0.70		
181	m1	0.808±0.004				
182	m4	0.804±0.007		1.71±0.63		0.542±0.319
183	m4	0.825±0.006		2.00±0.61		1.755±0.302
184	m2	0.822±0.002		6.58±0.70		
185						
186	m2	0.793±0.003		1.89±0.58		
187	m1	0.827±0.001				
188	m1	0.848±0.002				
189	m4	0.797±0.003		2.52±0.53		1.161±0.148
190	m2	0.782±0.001		6.29±0.49		
191						
192	m2	0.818±0.003		7.88±0.72		
193	m4	0.785±0.003		3.73±0.53		2.225±0.178
194	m2	0.777±0.002		5.51±0.53		
195	m5	0.733±0.002	0.768±0.002		785.79±42.62	
196	m5	0.700±0.003	0.744±0.003		269.76±21.92	
197	m5	0.798±0.005	0.839±0.003		608.79±36.62	
198	m5	0.657±0.002	0.699±0.002		649.44±21.82	
199	m5	0.778±0.002	0.806±0.002		760.62±49.93	
200	m2	0.818±0.002		3.10±0.68		
201	m2	0.816±0.002		3.88±0.63		
202						
203	m4	0.783±0.003		3.58±0.44		1.860±0.126
204	m2	0.795±0.002		5.47±0.55		
205	m4	0.850±0.003		4.04±0.82		2.060±0.200
206	m4	0.838±0.004		3.31±0.72		2.855±0.246
207						
208	m4	0.715±0.002		3.18±0.38		0.766±0.069
209	m4	0.829±0.002		2.53±0.77		0.193±0.099
210	m5	0.749±0.002	0.791±0.002		2049.27±194.03	
211						
212	m2	0.855±0.003		4.55±0.87		
213	m4	0.802±0.003		1.98±0.63		0.744±0.165
214	m2	0.810±0.003		7.73±0.66		
215	m2	0.823±0.003		7.94±0.69		
216						
217	m5	0.751±0.016	0.811±0.011		894.89±87.55	
218						

204

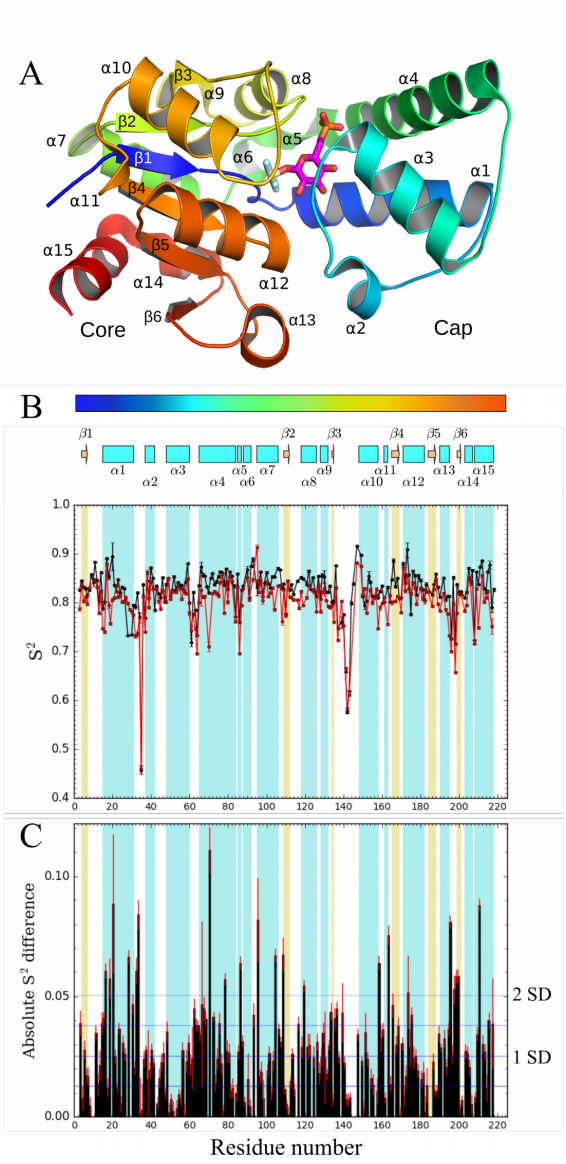


Fig S2

The model free comparison of βPGM$_{WT}$:AlF$_4$:G6P and βPGM$_{D10N}$:AlF$_4$:G6P complexes. **A)** A ribbon representation of the tertiary structure of βPGM (with N- to C-term from blue to red respectively) with α-helices and β-sheets annotated. **B)** Order parameters for βPGM$_{WT}$:AlF:G6P (black) and βPGM$_{D10N}$:AlF$_4$:G6P (red) complexes plotted with error (1 standard deviation (SD)). **C)** The absolute order parameter difference between βPGM$_{WT}$:AlF$_4$:G6P and βPGM$_{D10N}$:AlF$_4$:G6P complexes is presented with error bars at 1 SD (red). Horizontal lines correspond to the standard deviations for the dataset.

7. Chemical shift transition towards unfolded state analysis

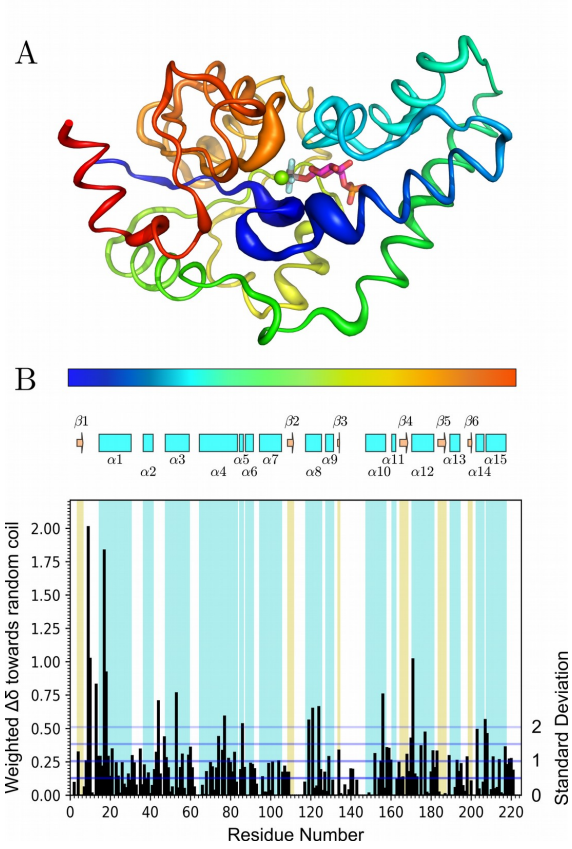


Fig S3

The combined backbone amide chemical shift differences between the assignments of WT and D10N variants of βPGM complexed with AlF_4^- and G6P. Chemical shifts of the $βPGM_{D10N}$:AlF_4:G6P complex were compared to both WT and a random coil chemical shift library (without local sequence compensation) (Schwarzinger et al. 2000). If the chemical shift moved towards random coil, the combined chemical shift was reported as in Williamson (Williamson 2013). Transition to random coil is primarily observed in the active site with large transitions for key residues in the hinge region (I13, A17, G18 and S171). **A)** residues that indicate transition to random coil, between $βPGM_{WT}$:AlF_4:G6P and $βPGM_{D10N}$:AlF_4:G6P complexes, with the polypeptide chain coloured from blue to red (N-term to C-term), magnitude of combined chemical shift change indicated by b-factor putty, and AlF_4 and G6P (pink) ligands included for reference. **B)** The absolute value of the chemical shift change towards random coil-like chemical shift per residue with secondary structure indicated and a color bar that refers to the coloring of the structure in part A. Horizontal lines correspond to the standard deviations for the dataset.

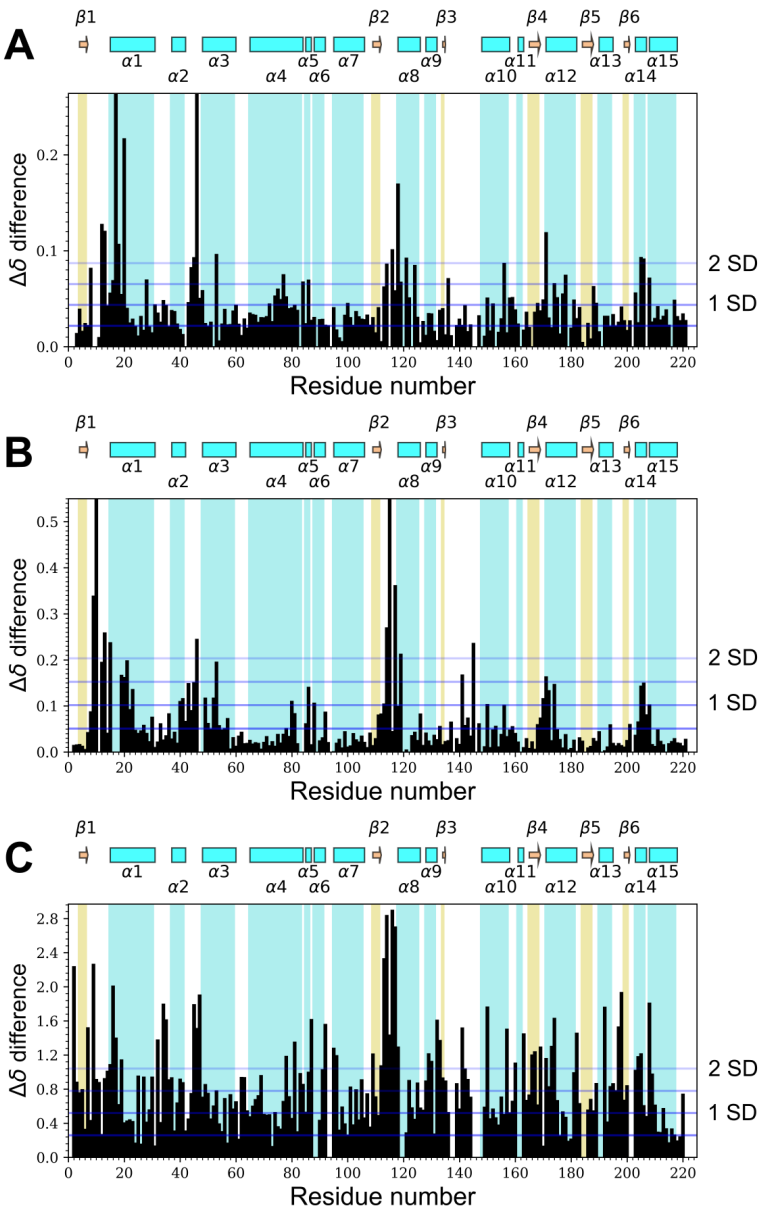


Fig S4.

Chemical shift analysis of combined backbone amide (HN and N) chemical shift perturbation. A) Comparison of βPGM$_{WT}$:AlF$_4$:G6P and βPGM$_{D10N}$:AlF$_4$:G6P complexes, B) comparison of βPGM$_{WT}$:MgF$_3$:G6P and βPGM$_{WT}$:BeF$_3$:G6P complexes, C) comparison of βPGM$_{WT}$:MgF$_3$:G6P and βPGM$_{D10N}$:βG16BP complexes.

8. X-ray crystallography methods

Crystallization and refinement of the βPGM_{D10N}:AlF_4:G6P structure was reported previously (Johnson et al. 2018), however subsequent refinement of the βPGM_{D10N}:AlF_4:G6P structure with carboxamide of residue N10 in both 180° sidechain rotamers is presented here. Initial refinement with the N10 carboxamide oriented such that the carbonyl atom coordinated the 1-OH group of G6P, resulted in a > 3 σ difference map peaks for the nitrogen atoms indicating that the incorrect rotamer had been modelled (**Fig. S5**). Reorientation of the carboxamide such that the nitrogen atom coordinates the 1-oxygen atom of G6P ($O1_{G6P}$) and subsequent refinement results in the disappearance of difference map peaks at this position.

Given the energetic penalty associated with deprotonation of the $O1_{G6P}$ atom when coordinated by the NH_2 group of N10, it would suggest that a conformer where the 1-OH group is not deprotonated would be preferred. This prediction strongly correlates with the solution NMR data presented in Supplementary section 13 (**Fig. S7**), where the βPGM_{D10N}:AlF_4:H_2O:βG1P complex (PDB: 5O6R; (Johnson et al. 2018)) is preferred at equilibrium in a conformation where the N10 sidechain amine coordinates a water molecule, rather than the nucleophilic 1-OH group. In this case, the enzyme preparation has equilibrated G6P with βG1P in the dead-time of the experiment due to residual catalytic activity (Johnson et al. 2018), which permitted the observation of equilibrium populations of the two complexes.

Refinement of another βPGM_{D10N}:AlF_4:G6P crystal at a higher resolution was performed (1.02 Å; **PDB: 6L03**) to further investigate the nature of the TSA binding. The crystal was both obtained and refined using the methods described previously (Johnson et al. 2018). This crystal was a *plate* morphology, data collection statistics presented in **Table S2**. Ligands were omitted until final rounds of refinement to avoid building into biased Fourier maps. In order to satisfy the electron density present, it was necessary to model *ca*. 50 residues across the cap and core domain with split occupancies, with the second occupancy chain translated *ca.* 1 Å away from the first chain. A B-factor weighting of 0.001 was applied in the final stages of refinement to avoid biasing atomic positions to minimise local b-factors. The resulting structure closely reflects the previous structure, and accommodates the ligand in a near identical manner (**Fig. S6)**.

9. Fig. S4 N10 sidechain rotamer indicates deprotonation of G6P 1-OH group.

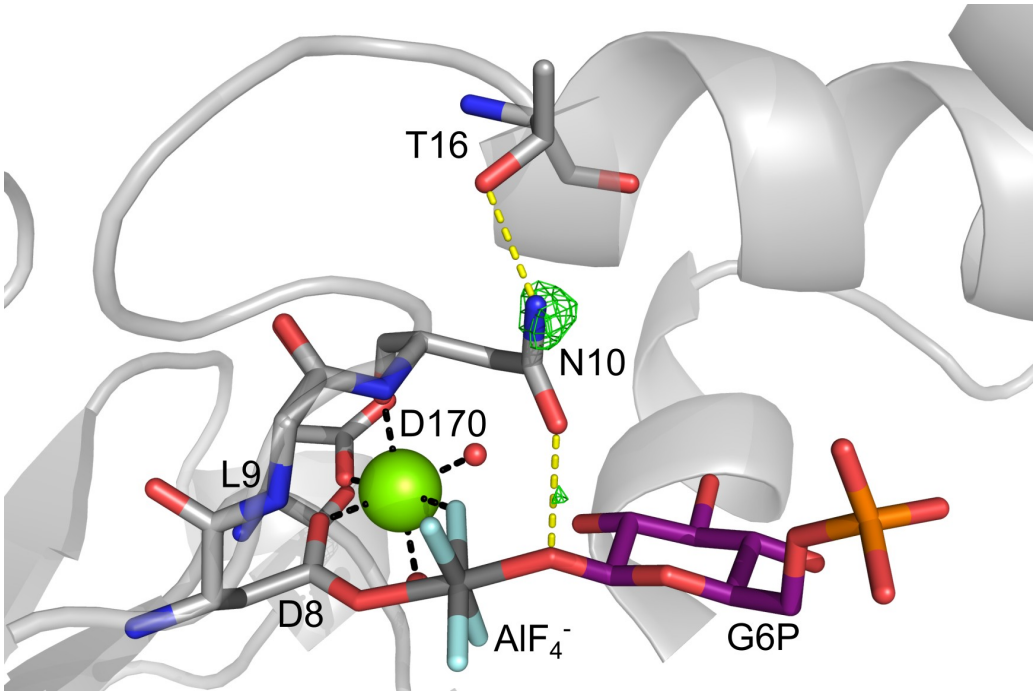


Fig S5.

Difference density (Fo – Fc; green mesh) for the βPGM$_{D10N}$:ALF$_4$:G6P structure (**PDB: 5OK2**; (Johnson et al. 2018)). Selected active site residues are shown as sticks in standard CPK colors, with carbons (grey), aluminum (dark grey). fluorine (light blue), and magnesium (light green). The G6P ligand is shown with purple carbon atoms (for clarity) and structural waters are shown as small red spheres. Yellow dashes indicate hydrogen bonds to and from residue N10 and black dashes show metal ion coordination. The difference density was generated following N10 side chain reorientaion in the final structure (with subsequent re-refinement) and is contoured at 3σ.

10. X-ray crystallography data acquisition and refinement table.

Data acquisition

Complex	βPGM_{D10N}:AlF_4:G6P
PDB code	6L03
Wavelength (\AA)	0.97949
Beamline	I02
Facility	DLS
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c, (Å)	37.52, 54.28, 104.42
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å) [1]	37.52 – 1.02 (1.02 – 1.05)
R_{merge} [1,2]	0.045 (0.917)
R_{pim} [1]	0.021 (0.589)
CC-half	0.999 (0.542)
$< I / \sigma I >$ [1]	16.1 (1.3)
Completeness (%) [1]	97.6 (82.5)
Multiplicity [1]	6.5 (3.7)
Total reflections	106736
Unique reflections	6606
Molecular replacement model	2WF6

Data refinement	
R (%) [3] / R_{free} (%) [4]	15.2 / 17.5
Protein	1961
Ligands	21
Metal ions	2
Water	232
Protein residues (asterix)	219
RMS deviations:	
Length	0.007
Angles	1.507
Average B factor (Å^2)	
Main chain	13.58
Side chains	16.46
Ligands (AlF_4 then G6P)	11.48, 11.44
Metal Ions (Mg^{2+} then Na^+)	10.31, 16.70
Water	23.94
Ramachandran analysis	
Favoured/allowed (\%)	97.65
Disallowed (\%)	0.39
Molprobity score (percentile)	1.27 (88th)

386

387 **Table S2** (previous page).

388 Data acquisition and processing for the new βPGM$_{D10N}$:AlF$_4$:G6P complex.

389 [1] Values for the higher resolution shell are in parenthesis.

390 [2] $R_{merge} = \sum_h \sum_i |I(h) - I(h)_i| / \sum_h \sum_i I(h)_i$, where I(h) is the mean weighted intensity after rejection of

391 outliers.

392 [3] $R = \sum_{hkl} ||F_{obs}| - k|F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor

393 amplitudes.

394 [4] $R_{free} = \sum_{hkl \in T} ||F_{obs}| - k|F_{calc}|| / \sum_{hkl \in T} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure

395 factor amplitudes and T is the test set of data omitted from refinement (5% in this case)

11. Omit map for βPGM_{D10N}:AlF_4:G6P complex (PDB: 6L03)

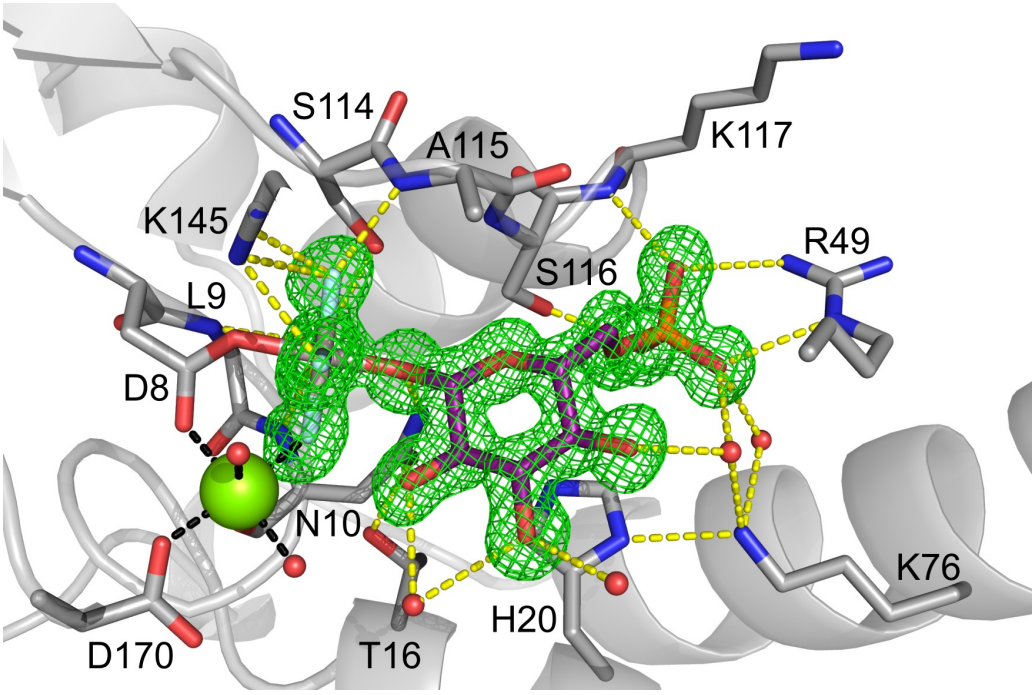


Fig S6

The difference density (Fo – Fc; green mesh) for the βPGM_{D10N}:ALF_4:G6P structure (**PDB: 6L03**). Selected active site residues are shown as sticks in standard CPK colors, with carbons (grey), aluminum (dark grey). fluorine (light blue), and magnesium (light green). The G6P ligand is shown with purple carbon atoms (for clarity) and structural waters are shown as small red spheres. Yellow dashes indicate hydrogen bonds and black dashes show metal ion coordination. The difference density was generated following omission of the AlF_4^- and G6P from the final structure (with subsequent re-refinement) and is contoured at 3σ.

12. Table angles within crystalographically determined AlF_4^- groups.

PDB:	$O\delta1_{D8}$ – Al – Fx (°)	$O1_{G6P}$ – Al – Fx (°)	Al – F_{1-4} improper angle
2WF6	91.22	88.90	1.70 *
5OK2	89.10	90.97	1.32
6L03	88.00	92.10	2.84

Table S3

F_x – Al – $O\delta1_{D8}$ and F_x – Al – $O1_{G6P}$ angles averaged across the four fluorides in each of the the AlF_4^- TSA complexes. Additionally the improper angle that the Al atom makes to the square plane of the 4 fluorides is reported as an average of all possible planes to account for any asymmetry present. The asterisk in the 2WF6 complex denotes that the angle is of opposite direction to the improper angle reported for 5OK2 and 6L03 structures.

13. ^{19}F NMR methods and SIIS determination

1D ^{19}F spectra

One dimensional $_{19}$F spectra to characterize WT and D10N AlF_4:G6P TSAs were acquired using a Bruker 500MHz Avance III spectrometer equipped with a 5mm QCI-F cryoprobe with z-axis gradients (MIB). A spectral width of 120 ppm centered at -140 ppm for ^{19}F was used without proton decoupling. Spectra were accumulations of 1024 – 2048 transients in order to achieve a sufficient signal-to-noise ratio. Spectra were processed with an EM window function with 10 Hz line-broadening and were analyzed in Topspin (Bruker).

^{19}F spectra of βPGM_{WT} and βPGM_{D10N} AlF_4:G6P TSA complexes for initial characterization were acquired at 298 K using 0.5 – 1 mM ^{15}N-βPGM in standard NMR buffer (50mM K+ HEPES pH 7.2/7.2*, 5mM $MgCl_2$, 1mM TSP). Both βPGM_{WT} and βPGM_{D10N} - AlF_4:G6P TSA complexes were made using 5mM $AlCl_3$, 20mM NaF, and ca. 20mM G6P.

Solvent induced isotope shift (SIIS) value determination for WT and D10N complexes.

Both βPGM_{WT} - and βPGM_{D10N} - AlF_4:G6P complexes were made in H_2O and D_2O using ^{13}C-1 labelled G6P in attempt to also characterise the chemical environment of the C1 carbon in each of the complexes. The (U)^{13}C1 labelled G6P was synthesized by incubating 45 mM 100% ^{13}C1-glucose with 14 U of hexokinase and ca. 50 mM ATP in a reaction volume of 4 ml, using a buffer of TRIS 100 mM (pH 8), $MgCl_2$ 50mM, and EDTA 2mM. The reaction was incubated for 4 hours at which point hexokinase (90 kDa) was removed by passing the reaction mixture through a 10 kDa MWCO vivaspin. The filtrate was split into two equal volumes and each was lyophilized overnight. One (U)^{13}C-1 labelled G6P mix was resuspended in 100 µl NMR buffer that contained 99.98 % D2O, while the other was resupended in 100 µl NMR buffer in 100% H_2O.

WT and D10N protein samples were buffer exchanged into standard NMR buffer with either 100% H_2O or 99.9% D_2O, and two $AlCl_3$ and NaF stocks were prepared in both 100% H_2O and 99.98% D_2O. The AlF_4:G6P TSA complexes were made using ca. 1mM WT and D10N ^{15}N-labelled enzyme in either 100% H_2O or 99.7% D_2O standard NMR buffer, supplemented with 5mM $AlCl_3$, 20 mM NaF, and 40mM (U)^{13}C-1 labelled G6P, resulting in 4 samples in total, two 100% H_2O and two *ca.* 99% D_2O. For H_20 samples, a 100% D_2O capillary was included to provide frequency lock in the spectrometer. These four samples were then used to record ^{19}F and ^{13}C 1D spectra of each of the complexes with SIIS values determined following the change in chemical shift of fluorine

resonances in H_2O and in D_2O. ^{19}F spectra were accumulations of 2048 transients with a spectral width of 120 ppm centered at -140 ppm, without proton decoupling. Spectra were processed using an EM window function with 10 Hz line-broadening in Topspin (Bruker), and referenced using TopSpin internal referencing.

In the AlF_4:G6P TSA complex with D10N two separate complexes were observed. The first of these complexes closely reflected the chemical shifts of βPGM_{D10N}:AlF_4:G6P TSA complex that was initially assigned. The second complex was more populated at equilibrium, and likely corresponds to the βPGM_{D10N}:AlF_4:H_2O:βG1P that has previously been observed crystallographically (**PDB: 5O6R**; (Johnson et al. 2018)). The F1 resonance in this complex that coordinates the catalytic Mg ion has moved significantly upfield compared to the F1 resonance in the βPGM_{D10N}:AlF_4:G6P TSA complex, which is consistent with the loss of a hydrogen bond from the 2-OH group. This observation is identical to those previously reported using the MgF_3^- TSA and fluoro-phosphonate βG1P analogs which were both crystallized and characterized by NMR previously (Jin et al. 2014).

SIIS value determination for WT and D10N AlF_4:G6P complexes.

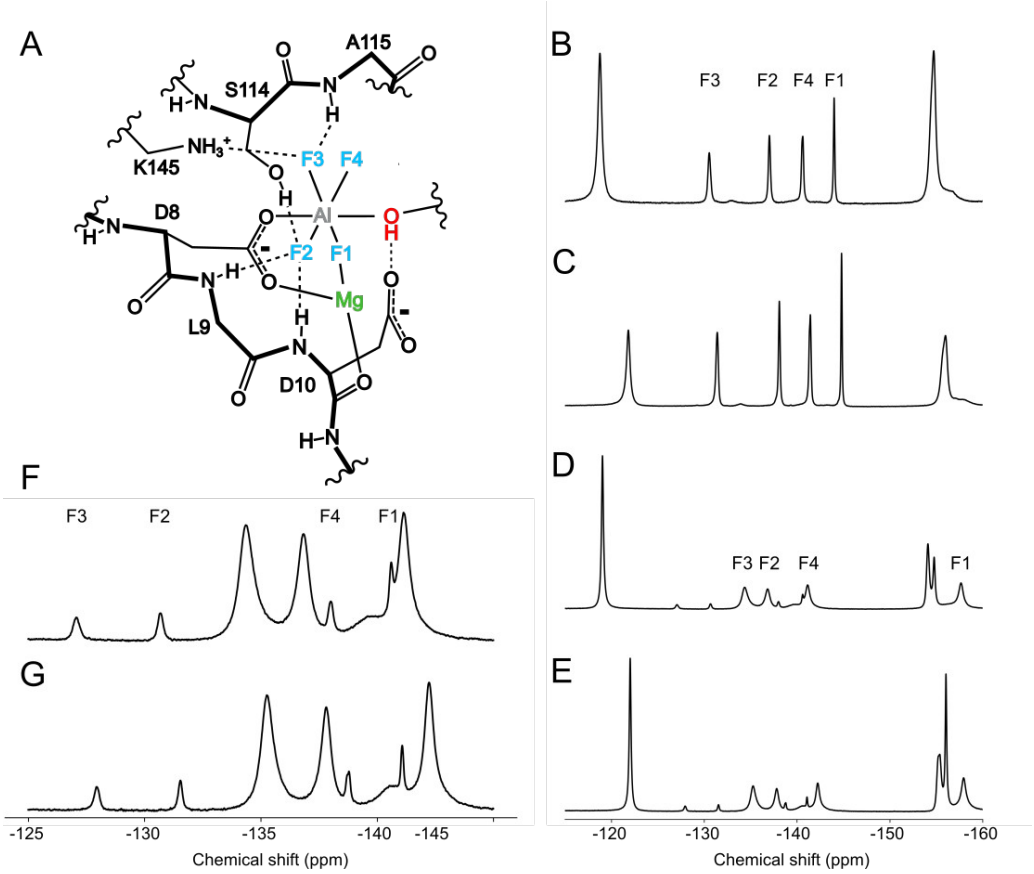


Fig S7

1D ^{19}F NMR spectra of βPGM_{WT} and βPGM_{D10N} complexed with AlF_4^- and G6P/βG1P. **A)** Presents a schematic of the active site of βPGM complexed with AlF_4^- (blue and grey atoms), the catalytic Mg^{2+} ion (green) and a nucleophilic hydroxyl group (red) that could belong to the 1-OH or 6-OH of a phosphorylated glucose, or to a water molecule. Backbone bonds are illustrated using thick lines, side chains (and Mg^{2+} coordination) using thin lines, and hydrogen bonds using dashed lines. Fluorine atoms are labelled in accordance with standard IUPAC nomencalture (Blackburn et al. 2017). **B)** and **C)** correspond to ^{19}F NMR spectra of the βPGM_{WT}:AlF_4:G6P TSA complex in **B)** 100% H_2O NMR buffer and **C)** 100% D_2O NMR buffer. Flourine resonances are labelled according to the reference scheme in A, with assignments determined previously (Baxter et al. 2010). **D)** and **E)** correspond to 1D ^{19}F NMR spectra of a mixed population of βPGM_{D10N}:AlF_4:H_2O:βG1P and βPGM_{D10N}:AlF_4:G6P TSA complexes, with the βG1P complex being favored at equilibrium. The βPGM_{D10N}:AlF_4:H_2O:βG1P complex in **D)** 100% H_2O NMR buffer and **E)** 100% D_2O NMR buffer

is annotated with a transferred assignment, using the upfield shift phenomena observed in Jin *et al.* 2014 (Jin et al. 2014). **F)** and **G)** correspond to a magnified view of the βPGM_{D10N}:AlF_4:G6P TSA complex in the previous NMR spectra, with the the complex in **F)** 100% H_2O NMR buffer and **G)** 100% D_2O NMR buffer with transferred assignments from the WT complex annotated, which themselves were corroborated by DFT chemical shift prediction. Chemical shifts for B and C are presented in **Table S4**, for D and E are presented in **Table S5**, and for F and G are presented in **Table S6** overleaf.

492

A

Fluorine	δ (ppm)	LWHH (Hz)
F1	-144.0	67
F2	-137.0	102
F3	-130.6	139
F4	-140.6	111

B

Fluorine	δ (ppm)	LWHH (Hz)
F1	-144.8	58
F2	-138.1	85
F3	-131.4	123
F4	-141.4	108

C

Fluorine	SIIS (ppm)	ΔLWHH (Hz)
F1	0.8	-10
F2	1.1	-17
F3	0.8	-16
F4	0.8	-3

494 **Table S4**. The chemical shifts and solvent-induced isotope shifts (SIIS) values for the
495 WT:AlF4:G6P TSA complex. **Subtables A** (H_2O) and **B** (D_2O) correspond to the chemical shifts
496 and peak linewidths at half height (FWHH) for the WT:AlF4:G6P TSA complex illustrated in Fig
497 X. Linewidth was extracted from the spectra using the deconvolution tool *dcon* in TopSpin v.4.0
498 **Subtable C** gives the SIIS (ppm) for each of the fluorine resonances defined as $^{19}F(H_2O$ buffer) –
499 $^{19}F(100\%\ D_2O$ buffer), as well as the change in LWHH which is defined as LWHH(D_2O buffer) –
500 LWHH(100% H_2O buffer).

A

Fluorine	δ (ppm)	LWHH (Hz)
F1	-140.7	274
F2	-130.7	77
F3	-127.1	100
F4	-138.0	186

B

Fluorine	δ (ppm)	LWHH (Hz)
F1	-141.1	198
F2	-131.5	69
F3	-128.0	81
F4	-138.7	168

C

Fluorine	SIIS (ppm)	ΔLWHH (Hz)
F1	0.4	-76
F2	0.9	-7
F3	0.9	-19
F4	0.8	-18

502 **Table S5.** The chemical shifts and solvent-induced isotope shifts (SIIS) values for the
503 D10N:AlF4:G6P TSA complex. **Subtables A** (H_2O) and **B** (D_2O) correspond to the chemical shifts
504 and peak linewidths at half height (FWHH) for the D10N:AlF4:G6P TSA complex illustrated in
505 Fig X. **Subtable C** gives the SIIS (ppm) for each of the fluorine resonances defined as $^{19}F(H_2O$
506 buffer) – $^{19}F(100\%\ D_2O$ buffer), as well as the change in LWHH which is defined as LWHH(D_2O
507 buffer) – LWHH(100% H_2O buffer).
508

A

Fluorine	δ (ppm)	LWHH (Hz)
F1	-157.6	315
F2	-136.8	291
F3	-134.4	329
F4	-141.1	260

B

Fluorine	δ (ppm)	LWHH (Hz)
F1	-158.0	340
F2	-137.8	246
F3	-135.3	290
F4	-142.2	244

C

Fluorine	SIIS (ppm)	ΔLWHH (Hz)
F1	0.3	25
F2	1.0	-45
F3	0.9	-39
F4	1.1	-16

510 **Table S6.** The chemical shifts and solvent-induced isotope shifts (SIIS) values for the
511 D10N:AlF4:H_2O:βG1P TSA complex. **Subtables A** (H_2O) and **B** (D_2O) correspond to the chemical
512 shifts and peak linewidths at half height (FWHH) for the D10N:AlF4:H_2O:βG1P TSA complex
513 illustrated in Fig X. **Subtable C** gives the SIIS (ppm) for each of the fluorine resonances defined as
514 $\delta^{19}F(H_2O$ buffer) – $\delta^{19}F(100\%\ D_2O$ buffer), as well as the change in LWHH which is defined as
515 LWHH(D_2O buffer) – LWHH(100% H_2O buffer).

14. Obtaining the active site models for the WT:AlF_4^-:G6P and D10N:AlF_4^-:G6P complexes

A quantum mechanical (QM) cluster model was constructed starting from the X-ray crystal structure of β-phosphoglucomutase inhibited with Glucose-6- phosphate (G6P) and aluminium tetrafluoride (AlF_4^-) (2WF6: 1.4Å). This contains a transition state analogue (TSA) of βG16BP, whereby AlF_4^- is used in place of the transferring phosphate (PO_3^-) group to 'trap' the transition state configuration. Amino acid residues not contributing to the stabilization of the active site through key hydrogen bonding interactions were removed. Specifically, we included G6P, AlF_4^- and a catalytic Mg^{2+} ion, along with 20 amino acid residues (8-12, 16-17, 20, 45-48, 113-116, 145, 169-171) and 6 explicit water molecules (2014, 2077, 2127, 2210, 2211, 2250) (Fig**.** 1 (main manuscript)). The resulting active site QM cluster model contained 386 atoms. All truncated amino acid residues were capped with methyl groups, with the carbon atom held fixed during geometry optimization to mimic the structural rigidity provided by the deleted amino acid residues. This resulted in a total of 14 fixed carbon atoms (Fig**.** 1 (main manuscript)). The G6P phosphorous atom was also held fixed, in its crystallographically determined coordinates. All fixed atoms are denoted with an asterisk (*) in Fig. 1 (main manuscript). Geometry optimization was performed with Gaussian09 ((Frisch et al., n.d.)) using the B3LYP hybrid functional formulation of Kohn-Sham Density Functional Theory (KS-DFT) (Becke 1993; Lee, Yang, and Parr 1988; Vosko, Wilk, and Nusair 1980; Stephens et al. 1994). A 6-31G basis set was used for all atoms except fluorine, which was treated with a 6-31+G(d) basis set. A better basis set for fluorine was chosen so as to improve the agreement of calculated ^{19}F NMR chemical shifts with experiment. The structure was considered optimized when the force on all nuclei fell below 1 μHartree/Bohr. The SCF was considered converged when the density matrix residual was less than 10^{-7}. To create a model for the mutant complex, the βPGM_{WT}:AlF_4^-:G6P active site model was manually altered at residue 10 from an Asp to an Asn. The resulting βPGM_{D10N}:AlF_4^-:G6P active site model was then reoptimized as above. Coordinates for the βPGM_{WT}:AlF_4^-:G6P and βPGM_{D10N}:AlF_4^-:G6P active site models are available on request (**j.waltho@sheffield.ac.uk**).

15. NMR Chemical shift calculations

NMR shielding tensors for ^{19}F nuclei in both the βPGM_{WT}:AlF_4^-:G6P and βPGM_{D10N}:AlF_4^-:G6P active site models were computed from the coupled-perturbed Hartree-Fock equation and gauge-invariant atomic orbitals (GIAO) derived from the DFT electron densities using standard algorithms implemented in Gaussian09 (Frisch et al., n.d.). A 6-31+G(d) basis set was used for the fluorine atoms. Calculated shielding tensors were plotted against experimental ^{19}F chemical shift values to determine calculated ^{19}F chemical shift values. See **Fig. 2 (main manuscript)** for a comparison of calculated and experimental ^{19}F chemical shifts.

16. Obtaining the active site models for the WT:PO_3^-:G6P and D10N:PO_3^-:G6P complexes

To assess the geometrical effects of having a more polarized species than AlF_4^- in our βPGM_{WT}:AlF_4^-:G6P and βPGM_{D10N}:AlF_4^-:G6P active site models, AlF_4^- was manually replaced with PO_3^- in each model. Geometry optimization was performed with Gaussian09 ((Frisch et al., n.d.)) using the B3LYP hybrid functional formulation of Kohn-Sham Density Functional Theory (KS-DFT) (Vosko, Wilk, and Nusair 1980; Becke 1993; Lee, Yang, and Parr 1988; Stephens et al. 1994). In this case, all atoms were held fixed, except the newly introduced PO_3^- species, and the **1-OH** hydrogen atom. This allowed us to assess the effect of changing the polarization and geometry of only the reacting species, given the exact same active site geometry. A 6-31G basis set was used for all atoms except the PO_3^- oxygen atoms, which were treated with a 6-31+G(d) basis set. The structure was considered optimized when the force on all nuclei fell below 1 μHartree/Bohr. The SCF was considered converged when the density matrix residual was less than 10^{-7}.

Fixed point	Closest S^2 parameter	WT S^2	D10N S^2
At CO of F7	NH of D8	0.83	0.83
At CA of I13	NH of I13	0.80	0.78
At CO of D15	NH of T16	0.80	0.75
At CA of E18	NH of E18	0.85	0.80
At CO of Y19	NH of H20	0.90	0.81
At CA of F21	NH of F21	0.84	0.81
At CO of L44	NH of K45	0.81	0.80
At CA of R49	NH of R49	0.85	0.85
At CO of L112	NH of A113	0.84	0.81
At CA of K117	NH of K117	-	0.79
At CO of S144	NH of K145	0.84	-
At CA of P146	NA	-	-
At CO of L168	NH of E169	0.81	0.81
At CA of Q172	NH of Q172	0.84	0.82

Table S7. Points of truncation in the active site model compared to NMR derived backbone order parameters. No residues showed significant mobility proximal to the points of truncation.

17. Obtaining the transition state model for the WT:PO_3^-:G6P transition state complex

An active site model for the transition state (TS) of the phosphoryl transfer of the 1-phosphate group was obtained from the βPGM_{WT}:AlF_4^-:G6P active site model, described above. AlF_4^- was manually replaced with PO_3^-, and the model further truncated. Specifically, we included β-glucose 1,6-biphosphate (βG16BP), and a catalytic Mg^{2+} ion, along with 10 amino acid residues (8-10, 46-47, 114-115, 145, 169-170) and 2 explicit water molecules (2210, 2211) (**Fig S8**). The resulting active site QM cluster model contained 163 atoms. All fixed atoms are denoted with an asterisk (*) in **Fig. S8**. The TS search was performed with Gaussian09 ((Frisch et al., n.d.)) using the B3LYP hybrid functional formulation of Kohn-Sham Density Functional Theory (KS-DFT) (Vosko, Wilk, and Nusair 1980; Stephens et al. 1994; Lee, Yang, and Parr 1988; Becke 1993). A cc-pVDZ basis set was used for all atoms, excepting atoms for which more care was given due to reaction importance. Specifically, oxygen atoms in formally negatively charged amino acid residues were treated with aug-cc-pVDZ (169,170, **non-transferring PO_3^-**), and those in residues directly involved in bond-making/bond-breaking were treated with aug-cc-pVTZ (8,10, **transferring PO_3^-, O-Sugar**). This procedure gave a converged TS model with a harmonic vibrational value of 158i cm^{-1} corresponding to motion along the reaction coordinate. However, in freezing certain Cartesian coordinates (**Fig. 3 in main manuscript**), there were a small number of non-relevant imaginary frequencies (67i, 32i, 27i, 15i, 8i cm^{-1}). Coordinates for the TS active site model are available on request (**j.waltho@sheffield.ac.uk**). Snapshots of the vibrational mode corresponding to motion along the reaction coordinate were taken at regular $O\delta1_{D8}$-P-O_3^- intervals (**0.14 Å**), and the energy evaluated at each point (**Fig. 3 in main manuscript**), using the same level of theory and basis set for each atom as in the TS search. An animation of the reaction trajectory is provided as a GIF in the supplementary content.

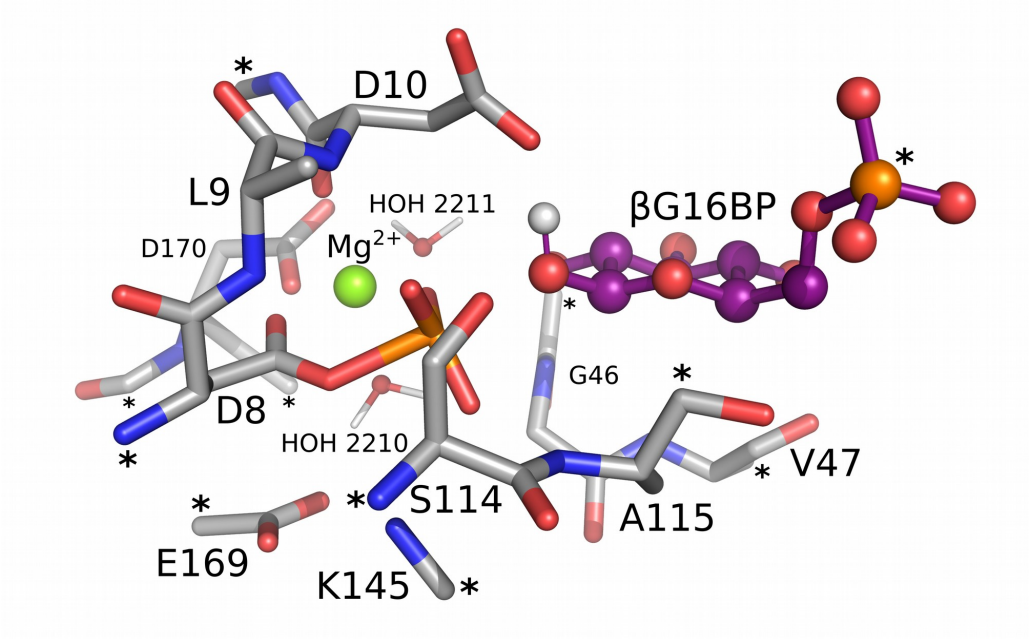


Fig S8. The 163 atom active site model for the phosphoryl transfer reaction between βG16BP and residue D8. Selected active site residues (and waters) are shown as sticks in standard CPK colours, with carbons (grey), phosphorus (orange), oxygen (red), nitrogen (blue), and magnesium (light green). The G6P ligand is shown with purple carbon atoms (for clarity). All fixed atoms are denoted with an asterisk (*).

Residue	WT	D10N
8	0.826±0.003	0.831±0.002
9	0.857±0.004	-
10	0.842±0.008	-
46	0.851±0.002	0.824±0.004
47	0.801±0.001	0.811±0.002
114	0.817±0.001	0.806±0.002
115	-	-
145	0.839±0.003	-
169	0.809±0.002	0.808±0.003
170	0.810±0.002	0.780±0.002

Table S8. TSA derived backbone order parameter (S^2) values for the active site residues included in the QM cluster model. While not perfect, the S^2 values determined in the βPGM_{WT}:AlF_4:G6P and

643 βPGM$_{WT}$:AlF$_4$:G6P TSA complexes indicate that the active site is held relatively rigidly (on the ps-
644 ns timescale) when the protein adopts transition state architecture. This would suggest that minimal
645 error would be introduced by truncating and fixating the residues as shown in **Fig. S8**.

18. QTAIM Charge Evaluation

Atomic charges were calculated using the Quantum Theory of Atoms in Molecules (QTAIM), with the AIMAll software package (Keith 2017). Electronic charge density for each atomic basin was determined using either the Proaim or Promega integration method implemented in the AIMAll package, while the integral of the Laplacian of each atomic basin was kept below $1x10^{-3}$ Hartrees. The theory of Interacting Quantum Atoms (IQA) (Blanco, Martin Pendas, and Francisco 2005; Francisco, Pendás, and Blanco 2006; Pendás, Francisco, and Blanco 2005, 2005), incorporated into the AIMAll software package, partitions the energy of a molecule into intra- and interatomic components, providing a rigorous path to quantification of both electrostatic and covalent contributions to atomic interactions (**see Equations S1-5**). In the context of the present work, atomic charges and interatomic electrostatic and covalent energies are calculated for selected atomic pairs (**Spreadsheet in Supporting Information, Fig. S9**),

Equation S1 shows how the system's energy is fully described by only intra-atomic (mono-) and interatomic (pairwise) energy contributions,

$$E_{IQA} = \sum_{A}^{n} E_{self}^{A} + \sum_{A}^{n} \sum_{B<A}^{n-1} V_{\text{inter}}^{AB} \qquad [S1]$$

where n is the number of atoms of the total system. The self and inter-atomic components are further decomposed as follows,

$$E_{self}^{A} = V_{ne}^{AA} + V_{ee}^{AA} + T^{A} \qquad [S2]$$

where T^{A} is the atomic kinetic energy, V_{ne}^{AA} represents intra-atomic electron-nuclear interactions while V_{ee}^{AA} represents intra-atomic electron-electron interactions.

$$V_{\text{inter}}^{AB} = V_{nn}^{AB} + V_{ne}^{AB} + V_{en}^{AB} + V_{ee}^{AB} \qquad [S3]$$

675 V_{nn}^{AB} accounts for nuclear-nuclear interactions, while V_{ne}^{AB} accounts for the interaction between the

676 nucleus of atom *A* and the electrons of atom *B*, and V_{en}^{AB} accounts for interaction between the

677 nucleus of atom *B* and the electrons of atom *A*. V_{ee}^{AB} accounts for electron-electron interactions and

678 can be written as the sum of Coulomb and exchange-correlation interactions.

679

680 $$V_{\mathrm{ee}}^{AB} = V_{Coulomb}^{AB} + V_{xc}^{AB} \qquad [S4]$$

681

682 We have now separated out the exchange-correlation interaction from the four classical electrostatic

683 interactions (summarized as V_{cl}^{AB}), giving,

684

685 $$V_{\mathrm{inter}}^{AB} = V_{cl}^{AB} + V_{xc}^{AB} \qquad [S5]$$

686

687 where V_{xc}^{AB} describes exchange-correlation energy, which can be seen as a measure of covalency

688 numerically dominated by the exchange part of V_{xc}^{AB}.

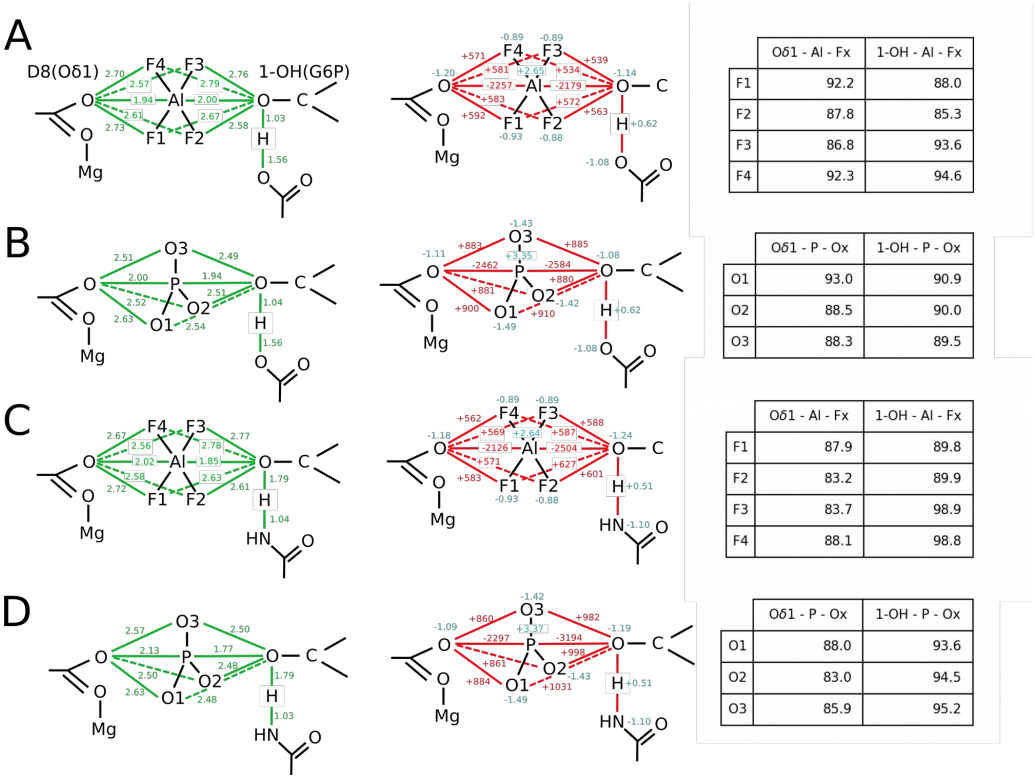
689

690 It is this pairwise interaction term V_{inter}^{AB} that we calculate for selected atomic pairs.

691

692 When performing a full IQA energy decomposition on large QM model systems, the number of
693 atom-atom pairwise interactions computed scales as N^2 (N = no. of atoms). When considering
694 motion across a reaction coordinate, this number of interactions must then be multiplied by the
695 number of snapshots computed. All of these terms are summed, as outlined above, in each snapshot,
696 to produce the total energy profile of the reaction of interest. In our case, nine snapshots are
697 computed across the reaction coordinate, each with a system size of 163 atoms. Correcting for
698 double-counting, we arrive at a total of 163 intra-atomic, as well as pairwise electrostatic (Vcl) and
699 pairwise covalent (Vxc) inter-atomic terms, for each snapshot. These terms must be ranked to
700 elucidate the subset of terms that are the main contributors to the overall size and shape of the
701 energy profile. To do so, we have developed the in-house program ANANKE, that uses the Relative
702 Energy Gradient (REG) method described in previous publications (Alkorta, Thacker, and Popelier
703 2018; Thacker and Popelier 2017, 2018).

19. Key electrostatic distances and interactions in the active site.



	Oδ1 - Al - Fx	1-OH - Al - Fx
F1	92.2	88.0
F2	87.8	85.3
F3	86.8	93.6
F4	92.3	94.6

	Oδ1 - P - Ox	1-OH - P - Ox
O1	93.0	90.9
O2	88.5	90.0
O3	88.3	89.5

	Oδ1 - Al - Fx	1-OH - Al - Fx
F1	87.9	89.8
F2	83.2	89.9
F3	83.7	98.9
F4	88.1	98.8

	Oδ1 - P - Ox	1-OH - P - Ox
O1	88.0	93.6
O2	83.0	94.5
O3	85.9	95.2

Fig S9.

Distances (left), electrostatic interactions (middle) and angles (right) between selected atoms in the QM models of **A)** βPGM_{WT}:AlF_4:G6P complex, **B)** the βPGM_{WT}:PO_3:G6P complex, **C)** βPGM_{D10N}:AlF_4:G6P complex, and **D)** the βPGM_{WT}:PO_3:G6P complex. A full table of interactions is available as a spreadsheet in the Supporting Information.

20. References

Alkorta, Ibon, Joseph C.R. Thacker, and Paul L.A. Popelier. 2018. “An Interacting Quantum Atom Study of Model SN2 Reactions (X–···CH3X, X = F, Cl, Br, and I).” Journal of Computational Chemistry 39 (10): 546–56. https://doi.org/10.1002/jcc.25098.

Baxter, N J, M W Bowler, T Alizadeh, M J Cliff, A M Hounslow, B Wu, D B Berkowitz, N H Williams, G M Blackburn, and J P Waltho. 2010. “Atomic Details of Near-Transition State Conformers for Enzyme Phosphoryl Transfer Revealed by MgF$\textsubscript{3}$\ textsuperscript{$-$} Rather than by Phosphoranes.” *Proc Natl Acad Sci U S A* 107: 4555–60. https://doi.org/10.1073/pnas.0910333106.

Becke, Axel D. 1993. “Density-Functional Thermochemistry. III. The Role of Exact Exchange.” *The Journal of Chemical Physics* 98 (7): 5648–52. https://doi.org/10.1063/1.464913.

Bieri, Michael, Edward J d’Auvergne, and Paul R Gooley. 2011. “RelaxGUI: A New Software for Fast and Simple NMR Relaxation Data Analysis and Calculation of Ps-Ns and $µ$s Motion of Proteins.” *Journal of Biomolecular NMR* 50 (2): 147–55. https://doi.org/10.1007/s10858-011-9509-1.

Blackburn, G. Michael, Jacqueline Cherfils, Gerard P. Moss, Nigel G.J. Richards, Jonathan P. Waltho, Nicholas H. Williams, and Alfred Wittinghofer. 2017. “How to Name Atoms in Phosphates, Polyphosphates, Their Derivatives and Mimics, and Transition State Analogues for Enzyme-Catalysed Phosphoryl Transfer Reactions (IUPAC Recommendations 2016).” *Pure and Applied Chemistry* 89 (5): 653–75. https://doi.org/10.1515/pac-2016-0202.

Blanco, M A, A Martin Pendas, and E Francisco. 2005. “Interacting Quantum Atoms: A Correlated Energy Decomposition Scheme Based on the Quantum Theory of Atoms in Molecules.” *Journal of Chemical Theory and Computation* 1 (6): 1096–1109. https://doi.org/10.1021/ct0501093.

Clore, M G, A Szabo, A Bax, L E Kay, P C Driscoll, and A M Gronenborn. 1990. “Deviations from the Simple Two-Parameter Model-Free Approach to the Interpretation of Nitrogen-15 Nuclear Magnetic Relaxation of Proteins.” *J Am Chem Soc* 112: 4989–91.

d’Auvergne, E J, and P R Gooley. 2008a. “Optimisation of NMR Dynamic Models I. Minimisation Algorithms and Their Performance within the Model-Free and Brownian Rotational Diffusion Spaces.” *J Biomol NMR* 40: 107–19. https://doi.org/10.1007/s10858-007-9214-2.

———. 2008b. “Optimisation of NMR Dynamic Models II. A New Methodology for the Dual Optimisation of the Model-Free Parameters and the Brownian Rotational Diffusion Tensor.” *J Biomol NMR* 40: 121–33. https://doi.org/10.1007/s10858-007-9213-3.

d’Auvergne, Edward J., and Paul R. Gooley. 2003. “The Use of Model Selection in the Model-Free Analysis of Protein Dynamics.” *Journal of Biomolecular NMR* 25 (1): 25–39. https://doi.org/10.1023/A:1021902006114.

———. 2006. “Model-Free Model Elimination: A New Step in the Model-Free Dynamic Analysis of NMR Relaxation Data.” *Journal of Biomolecular NMR* 35 (2): 117–35. https://doi.org/10.1007/s10858-006-9007-z.

———. 2007. “Set Theory Formulation of the Model-Free Problem and the Diffusion Seeded Model-Free Paradigm.” *Molecular BioSystems* 3 (7): 483. https://doi.org/10.1039/b702202f.

Delaglio, Frank, Stephan Grzesiek, Geerten W Vuister, Guang Zhu, John Pfeifer, and Ad Bax. 1995. “NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes.” *Journal of Biomolecular NMR* 6 (3): 277–93. https://doi.org/10.1007/BF00197809.

Francisco, E., A. Martín Pendás, and M. A. Blanco. 2006. “A Molecular Energy Decomposition Scheme for Atoms in Molecules.” *Journal of Chemical Theory and Computation* 2 (1): 90–102. https://doi.org/10.1021/ct0502209.

Frisch, M J, G W Trucks, H B Schlegel, G E Scuseria, M A Robb, J R Cheeseman, G Scalmani, et al. n.d. “Gaussian∼09 {R}evision {E}.01.”

Griffin, J. L., M. W. Bowler, N. J. Baxter, K. N. Leigh, H. R. W. Dannatt, a. M. Hounslow, G. M. Blackburn, C. E. Webster, M. J. Cliff, and J. P. Waltho. 2012. “Near Attack Conformers Dominate -Phosphoglucomutase Complexes Where Geometry and Charge Distribution Reflect Those of Substrate.” *Proceedings of the National Academy of Sciences* 109 (18): 6910–15. https://doi.org/10.1073/pnas.1116855109.

Halle, Bertil. 2009. “The Physical Basis of Model-Free Analysis of NMR Relaxation Data from Proteins and Complex Fluids.” *Journal of Chemical Physics* 131 (22): 1–23. https://doi.org/10.1063/1.3269991.

Halle, Bertil, Thomas Andersson, Sture Forsén, and Björn Lindman. 1981. “Protein Hydration from Water Oxygen-17 Magnetic Relaxation.” *Journal of the American Chemical Society* 103 (3): 500–508. https://doi.org/10.1021/ja00393a004.

Halle, Bertil, and Göran Carlström. 1981. “Hydration of Ionic Surfactant Micelles from Water Oxygen-17 Magnetic Relaxation.” *Journal of Physical Chemistry* 85 (14): 2142–47. https://doi.org/10.1021/j150614a037.

Halle, Bertil, and Håkan Wennerström. 1981. “Interpretation of Magnetic Resonance Data from Water Nuclei in Heterogeneous Systems.” *The Journal of Chemical Physics* 75 (4): 1928–43. https://doi.org/10.1063/1.442218.

Hyberts, Sven G, Alexander G Milbradt, Andreas B Wagner, Haribabu Arthanari, and Gerhard Wagner. 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): 315–27. https://doi.org/10.1007/s10858-012-9611-z.

Hyberts, Sven G, Scott A Robson, and Gerhard Wagner. 2013. “Exploring Signal-to-Noise Ratio and Sensitivity in Non-Uniformly Sampled Multi-Dimensional NMR Spectra.” *Journal of Biomolecular NMR* 55 (2): 167–78. https://doi.org/10.1007/s10858-012-9698-2.

Jin, Yi, Debabrata Bhattasali, Erika Pellegrini, Stephanie M Forget, Nicola J Baxter, Matthew J Cliff, Matthew W Bowler, David L Jakeman, G Michael Blackburn, and Jonathan P Waltho. 2014. “α-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): 12384–89. https://doi.org/10.1073/pnas.1402850111.

Johnson, Luke A., Angus Robertson, Nicola Jane Baxter, Clare Trevitt, Claudine Bisson, Yi Jin, Henry Wood, et al. 2018. "Van Der Waals Contact between Nucleophile and Transferring Phosphorus Is Insufficient to Achieve Enzyme Transition State Architecture." *ACS Catalysis*. https://doi.org/10.1021/acscatal.8b01612.

Keith, Todd A. 2017. "AIMAll (Version 17.11.14),." TK Gristmill Software, Overland Park KS, USA.

Lakomek, Nils Alexander, Jinfa Ying, and Ad Bax. 2012. "Measurement of 15N Relaxation Rates in Perdeuterated Proteins by TROSY-Based Methods." *Journal of Biomolecular NMR* 53 (3): 209–21. https://doi.org/10.1007/s10858-012-9626-5.

Lee, Chengteh, Weitao Yang, and Robert G Parr. 1988. "Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density." *Phys. Rev. B* 37 (2): 785–89. https://doi.org/10.1103/PhysRevB.37.785.

Lipari, G, and A Szabo. 1982a. "Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .1. Theory and Range of Validity." *J Am Chem Soc* 104: 4546–59.

———. 1982b. "Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .2. Analysis of Experimental Results." *J Am Chem Soc* 104: 4559–70.

Niklasson, Markus, Renee Otten, Alexandra Ahlner, Cecilia Andresen, Judith Schlagnitweit, Katja Petzold, and Patrik Lundström. 2017. "Comprehensive Analysis of NMR Data Using Advanced Line Shape Fitting." *Journal of Biomolecular NMR* 69 (2): 93–99. https://doi.org/10.1007/s10858-017-0141-6.

Pendás, A. Martín, E. Francisco, and M. A. Blanco. 2005. "Two-Electron Integrations in the Quantum Theory of Atoms in Molecules with Correlated Wave Functions." *Journal of Computational Chemistry* 26 (4): 344–51. https://doi.org/10.1002/jcc.20173.

Reed, Michelle A C, Andrea M Hounslow, K H Sze, Igor G Barsukov, Laszlo L P Hosszu, Anthony R Clarke, C.Jeremy Craven, and Jonathan P Waltho. 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): 1189–1201. https://doi.org/https://doi.org/10.1016/S0022-2836(03)00625-9.

Schwarzinger, S., G. J.A. Kroon, T. R. Foss, P. E. Wright, and H. J. Dyson. 2000. "Random Coil Chemical Shifts in Acidic 8 M Urea: Implementation of Random Coil Shift Data in NMRView." *Journal of Biomolecular NMR* 18 (1): 43–48. https://doi.org/10.1023/A:1008386816521.

Stephens, P. J., F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. 1994. "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields." *Journal of Physical Chemistry®* 98 (45): 11623–27. https://doi.org/10.1021/j100096a001.

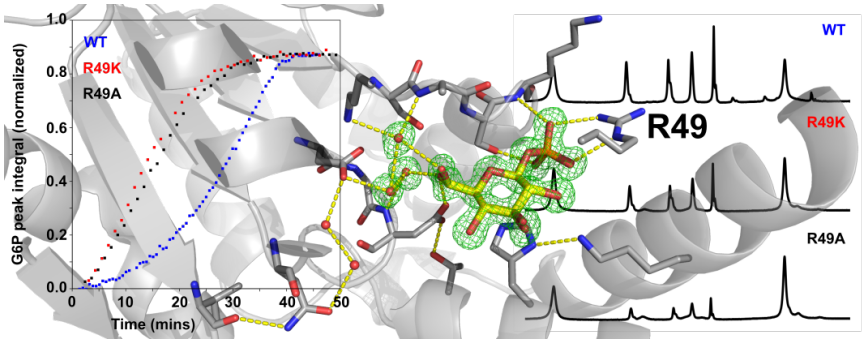
Thacker, Joseph C.R., and Paul L.A. Popelier. 2017. “The ANANKE Relative Energy Gradient (REG) Method to Automate IQA Analysis over Configurational Change.” *Theoretical Chemistry Accounts* 136 (7): 1–13. https://doi.org/10.1007/s00214-017-2113-z.

———. 2018. “Fluorine Gauche Effect Explained by Electrostatic Polarization Instead of Hyperconjugation: An Interacting Quantum Atoms (IQA) and Relative Energy Gradient (REG) Study.” *Journal of Physical Chemistry A* 122 (5): 1439–50. https://doi.org/10.1021/acs.jpca.7b11881.

Vosko, S. H., L. Wilk, and M. Nusair. 1980. “Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis.” *Canadian Journal of Physics* 58 (8): 1200–1211. https://doi.org/10.1139/p80-159.

Williamson, Mike P. 2013. “Using Chemical Shift Perturbation to Characterise Ligand Binding.” *Progress in Nuclear Magnetic Resonance Spectroscopy* 73: 1–16. https://doi.org/10.1016/j.pnmrs.2013.02.001.

A.3 Paper III: Arg - phosphate interaction in β-phosphoglucomutase improves substrate affinity, but introduces inhibition



Contribution:I performed much of the crystallography, I designed the experiments and analysed and interpreted the NMR data, I wrote the manuscript with contributions from HPW and JPW.

Arg - phosphate interaction in β-phosphoglucomutase improves substrate affinity, but introduces inhibition.

Angus J. Robertson,[‡] Henry P. Wood,[‡] Nikita Vekaria,[§] Alex L. Wilson,[§] Clare R. Trevitt,[‡] Andrea M. Hounslow,[‡] Claudine Bisson,[‖] Matthew J. Cliff,[§] and Jonathan P. Waltho[*,‡,§]

‡*Department of Molecular Biology and Biotechnology, The University of Sheffield*

§*Manchester Institute of Biotechnology, The University of Manchester*

‖*Department of Biological Sciences, Birkbeck, University of London*

E-mail: j.waltho@sheffield.ac.uk

Running header

Arg - phosphate interaction in β-phosphoglucomutase improves substrate affinity, but introduces inhibition.

Abstract

Under biological conditions, phosphate mono- (R-P-O-R) and di-ester (R-O-P-O-R) bonds have half-lives of millions of years, a stability which plays a crucial role in maintaining cell viability. Phosphoryl transfer enzymes have evolved with some of the largest rate accelerations known to biology, with typical catalytic rate enhancements (k_{cat}/k_{non}) of 10^{21}. β-phosphoglucomutase (βPGM) [EC 5.4.2.6] from *Lactococcus lactis* is a well-characterized magnesium-dependent phosphoryl transfer enzyme of the haloacid dehydrogenase superfamily. βPGM has two phosphate binding sites necessary for its mutase activity, but is inhibited by the initial substrate of the reaction (β-glucose 1-phosphate (βG1P)) - the structural mechanism of which is currently unclear. Here we show that removal of the guianidinium group of an arginine residue (R49) in the non-catalytic (*distal*) phosphate binding site alleviates a βG1P-dependent lag-phase prior to steady state catalysis. Additionally, this *distal* site shows redundancy in phosphate binding as proximal cationic groups are recruited to maintain charge balance in R49K/A enzyme variants, and further indicates a mechanism for ligand association/dissociation. In the catalytic site, metal fluoride transition state analogue (TSA) complexes of R49K/A enzyme variants show minimal perturbation to the electronic environment around the transferring phosphate mimic when the protein adopts a fully closed TSA conformer. Furthermore, we structurally characterize βG1P bound to βPGM in a closed, non-catalytically competent, manner which provides a structural basis for the βG1P inhibition of the open-βPGM enzyme modeled previously. Together, this study highlights the delicate antagonism present between substrate affinity and inhibition in an archetypal phosphoryl transfer enzyme.

Introduction

Under biological conditions, phospho mono- (R-P-O-R) and di-ester (R-O-P-O-R) bonds have half-lives of millions of years.[1] This inherent stability is essential to ensure the high-fidelity storage of genetic information within DNA. However, many core biological pro-

cesses, including DNA processing, metabolic cycles and cell signaling, depend on the efficient transfer of phosphate groups between metabolites. Phosphoryl transfer enzymes have evolved with some of the largest rate accelerations known to biology, with typical catalytic rate enhancements (k_{cat}/k_{non}) of 10^{21}.[1–3] Phosphoglucomutases catalyze phosphoryl group transfer to/from sugar molecules and reversibly produce glucose 6-phosphate, an important precursor for glycolysis and energy production in both prokaryotes and eukaryotes. β-phosphoglucomutase (βPGM) [EC. 5.4.2.6] from *Lactococcus lactis* is a well-characterized[4–11] magnesium-dependent phosphoryl transfer enzyme, which catalyzes the reversible isomerization of β-glucose 1-phosphate (βG1P) to glucose 6-phosphate (G6P) via a β-glucose 1,6-bisphosphate (βG16BP) intermediate using a ping-pong bi-bi reaction mechanism[6] (Fig. 1). As part of this mutase reaction, the enzyme adopts two different catalytically competent states, the substrate-free state and the phospho-enzyme state (βPGM^{P}, phosphorylated at residue D8), which have different substrate specificities. The βPGM^{P} state preferentially binds βG1P and G6P substrates while the substrate-free enzyme is more specific for the reaction intermediate βG16BP.[12]

The active site of βPGM is located at the interface between the helical cap domain (T16-V87) and the α/β core domain (M1-D15, S88-K216) and opening and closing of the cap domain relative to the core domain occurs during the catalytic cycle.[6] The active site contains a magnesium(II) ion binding site and two phosphate binding sites. One phosphate binding site, termed the *proximal* site, is adjacent to the Mg^{2+} ion at the catalytic center ($Mg_{cat}{}^{2+}$), while the other phosphate binding site, termed the *distal* site, is *ca.* 10 Å removed from the catalytic center. The *distal* phosphate binding site has a role in anchoring ligands in the active site via interactions with several conserved residues (R49, S116, K117), together forming a positive electrostatic region in the substrate-free enzyme. The interplay between the two phosphate binding sites allows βPGM to bind substrates, intermediates and products in two orientations to facilitate mutase activity - the ability to both transfer and remove a phosphoryl group from a substrate using the same active site residues.[9,10]

The binding of phosphate to the active site of βPGM is mediated through recognition by a substrate affinity domain.[4,14] An Arg residue (R49) makes an ionic bond with the phosphate group of substrates in several ground state analogue (GSA) and transition state analogue (TSA) complexes with (and without) metal fluoride phosphate mimics.[9–11,15,16] Arg - phosphate interactions have been reported to provide substantial binding energies in the range 11-13 kcal mol^{-1} in several enzymes,[17–19] with some cases indicating an associated conformational change necessary for catalysis.[20–23]

Here we structurally characterize single mutations to R49 in the *distal* phosphate binding site using metal fluoride-based transition state analogue complexes[7–10] and determine a *ca.* 4.1 kcal mol^{-1} stabilization of the transition state analogue complex when the distal phosphate group is coordinated by Arg rather than Lys. Furthermore, these R49 enzyme variants alleviate the βG1P-dependent component of the kinetic lag-phase prior to steady state catalysis in βPGM[12] which likely results from the reduced phosphate binding capacity in the *distal* site. A mutation that weakens $Mg_{cat}{}^{2+}$ binding in the *proximal* site facilitated the structural characterization of βG1P bound to Mg^{2+}-free enzyme in a closed but catalytically non-competent complex. Taken together, these results present a structural mechanism for the substrate inhibition displayed by βPGM, and highlight the delicate antagonism present between substrate affinity and inhibition.

Results

Structure of substrate-free R49 variants

The distal phosphate group of substrate makes hydrogen bonds with the guanidinium group of R49 the backbone NH of K117, and sidechain OH of S116 in the *distal* site. Since arginine - phosphate interactions have been reported to provide substantial binding energies in the range 11-13 kcal/mol,[17–19] with some cases indicating associated conformational change,[20–23] R49K and R49A variants of βPGM (βPGM_{R49K} and βPGM_{R49A} respectively)

were generated. Neither variant showed any deleterious effect to the expression or overall fold of the recombinant proteins (compared to wild-type βPGM (βPGM_{WT})) when observed using 2D NMR (Fig. S1). To investigate the active site of the substrate-free βPGM_{R49K} and βPGM_{R49A} conformers, substrate-free βPGM_{R49K} and βPGM_{R49A} complexes were crystallized and their structures were determined to 1.6 Å and 2.0 Å resolution respectively (**PDB: 6HDH, 6HDI** respectively; Table 2; Fig. 7). Both structures closely overlaid with previously deposited open-βPGM_{WT} structures (Table 4, 5) with a Mg^{2+} ion in the catalytic center ($Mg_{cat}{}^{2+}$) of the enzyme. Comparison of the *distal* phosphate binding site shows minimal structural perturbation to residues surrounding the mutation, while a Poisson-Boltzman energy surface indicates a reduced positive charge in the *distal* phosphate binding site of the βPGM_{R49A} variant (Fig. 3). In the substrate-free βPGM_{R49K} structure, the K sidechain occupies a similar position to the R sidechain, with the amine nitrogen of K49 located 0.9 Å away from the Nϵ atom of R49 in the substrate-free βPGM_{WT} structure. In the substrate-free βPGM_{R49A} structure, the removal of the R49 sidechain eliminates the possibility for a direct interaction between phosphate and residue 49, and no definable water molecule position was identifiable in the resulting void. Taken together, this suggests that binding of substrate to βPGM_{R49K} should be impaired but less so than binding to βPGM_{R49A}.

Kinetics and binding of R49 variants

Previous kinetic characterization of the βPGM_{R49K} and βPGM_{R49A} variants indicated that the mutations substantially disrupt enzyme activity compared to WT, lowering k_{cat} values by 300 and 200 fold and K_m values by 40 and 3000 fold, respectively.[4] To corroborate this, the standard glucose 6-phosphate dehydrogenase coupled assay was used to monitor conversion of βG1P to G6P using acetylphosphate (AcP; 20 mM) as a priming agent.[11] The data for the βPGM_{R49K} variant fitted to a K_m of 600 ± 100 μM (*cf.* 92 ± 6 μM for βPGM_{WT}[11]) and a k_{cat} of 2.1 ± 0.3 (*cf.* 24.5 ± 0.7 s^{-1} for βPGM_{WT}[11]). However, in the βPGM_{R49A} variant a linear substrate concentration dependence of the reaction velocity was observed over the

standard[12] concentration range (Fig. 4). Together this indicates that it was not possible to determine the K_m or k_{cat} value reliably for both βPGM_{R49K} and βPGM_{R49A} using this method.

To better saturate the βPGM_{R49K} and βPGM_{R49A} variants with βG1P, 1D ^{31}P NMR was used to follow the interconversion of βG1P (10 mM) and G6P using the method described previously.[11] 20 mM AcP was used as a priming agent to initiate the reaction and steady state k_{obs} values of 11.6 $\pm$ 1 s^{-1} and 5.6 $\pm$ 0.5 s^{-1} were determined for R49K and R49A respectively (Fig. 4). This represented a 7 and 11-fold reduction in catalytic rate constant compared to the βPGM_{WT} value of 70 $\pm$ 30 s^{-1} reported previously using this method.[11] It was also noticeable that the lag-phase associated with βPGM_{WT} was substantially perturbed by the mutation (see below).

Standard methods to investigate βG1P binding to the substrate-free βPGM enzyme are complicated owing to mutase and phosphatase activity of βPGM. Instead, the relative affinities of G6P in the AlF_4:G6P TSA complexes were determined for both βPGM_{R49K} and βPGM_{R49A}. The equivalent affinity in βPGM_{WT} is 9 $\pm$ 1 μM.[10] The level of TSA saturation (Fig. 6) was determined using 2D $^{1}H^{15}N$ TROSY NMR of the R49A and R49K variants (1 mM) complexed with 5 mM AlF_4^- and 20 mM G6P (under these conditions, the βPGM_{WT}:AlF_4:G6P complex is fully saturated). Both R49 variants demonstrated partial TSA complex formation (characterized by the downfield shift of the K117 backbone amide resulting from hydrogen bonding to the 6-phosphate of G6P), in slow exchange (<10 s^{-1}) with the holo-enzyme where AlF_4^- is bound in the catalytic center. K_d values of *ca.* 11 mM for R49K and 49 mM for R49A were determined using the holo-enzyme and TSA peaks in the TROSY spectra for the side-chain indole amide resonance of W24. These K_d values are some 3 orders of magnitude larger than for βPGM_{WT}, consistent with a substantial reduction in binding affinity for phosphorylated ligand as a result of mutation of the R49 residue.

Structure of R49K/A TSAs

To investigate any communication of reduced phosphate binding affinity across the ligand from *distal* to *proximal* phosphate binding sites at the point of chemical transfer, metal fluorides were used to trap transition state analogue (TSA) complexes of the phosphoryl transfer process involving G6P.[7–10] The βPGM$_{\mathrm{R49K}}$ and βPGM$_{\mathrm{R49A}}$ variants were crystallized in complex with AlF_4^- and G6P using standard crystallization conditions.[9] The resulting βPGM$_{\mathrm{R49A}}$:AlF_4:G6P and βPGM$_{\mathrm{R49K}}$:AlF_4:G6P structures were determined to 1.2 Å resolution (**PDB: 6HDJ, 6HDK** respectively; Table 3). Each variant produced a closed complex with a near identical protein conformation and mode of ligand binding to the equivalent complex in the WT enzyme (PDB: 2WF6; Table 4, 5). Furthermore, the interactions in the *proximal*, site of the βPGM$_{\mathrm{R49K}}$ and βPGM$_{\mathrm{R49A}}$ variants were also equivalent to those in the WT complex (Fig. 5).

Coordination of the phosphate group of G6P in the *distal* site was different in each of the enzyme variants due to the R49 mutation. In the βPGM$_{\mathrm{R49K}}$:AlF_4:G6P complex, the 6-phosphate group of G6P occupies the same position in the *distal* site, with the K49 sidechain replacing R49 and making a monodentate hydrogen bonding interaction with the 6-phosphate. There is a 0.1 Å reduction in hydrogen bond length between the backbone amide of K117 and the phosphate oxygen, which is consistent with the downfield shifted backbone amide peak relative to the βPGM$_{\mathrm{WT}}$:AlF_4:G6P complex (Fig. 6). In the βPGM$_{\mathrm{R49A}}$:AlF_4:G6P complex, A49 does not have the the capacity to charge balance the 6-phosphate group of G6P in the *distal* site. Instead, the side chain amine of K117 from the core domain on the opposite side of the active site is recruited from a solvent exposed position on the enzyme surface to the *distal* phosphate binding site (Fig. 5). Thus, K117 makes two interactions with the 6-phosphate group in the *distal* site, one through the backbone amide (2.9 Å) and the second through the side chain amine (2.8 Å). When observed crystallographically, there is no lengthening of the hydrogen bond between the K117 backbone amide and the phosphate oxygen of G6P despite the observed chemical shift change in the two TROSY

spectra (from βPGM$_{WT}$ to βPGM$_{R49A}$, Fig. 5). Consequently, the observed upfield shift of the K117 backbone amide may be as a result of the strain induced by maintaining the charge balance in the *distal* site. Together, these structures show that the R49 *distal* site mutations can achieve the transition state architecture and indicate an impaired coordination of the phosphate group of the substrate in the *distal* phosphate binding site. This sub-optimal coordination of the *distal* phosphate is a likely cause of the reduced k_{obs} values for the R49 variants (through reducing K_m), highlighting the importance of a bidentate interaction to a phosphate group in the *distal* site.

In order to ensure that the formation of a TSA was not an artifact of using a non-isosteric TSA in AlF_4^-, the βPGM$_{R49K}$ and βPGM$_{R49A}$ variants were crystallized in complex with MgF_3^- and G6P to probe the transition state using both an isosteric and isoelectronic TSA. The resulting βPGM$_{R49K}$:MgF_3:G6P and βPGM$_{R49A}$:MgF_3:G6P structures were determined to 1.2 Å and 1.3 Å resolution respectively (**PDB: 6HDL, 6HDM** respectively; Table 3). These structures showed a high degree of similarity to both the AlF_4^- TSA complexes and to the equivalent βPGM$_{WT}$:AlF_4:G6P structure (Fig. 5; Table 4, 5). Notably in the βPGM$_{R49A}$:MgF_3:G6P complex the side chain amine of K117 is again recruited to the *distal* site where it charge balances the phosphate group of G6P. Taken together, these TSA complexes demonstrate that the enzyme variants are capable of stably forming transition state protein architecture, with minimal perturbation from the TSA structures in the wild type enzyme.

^{19}F NMR of R49K/A TSAs

Given the perturbation of ligand binding in the βPGM$_{R49K}$ and βPGM$_{R49A}$ variants, it was investigated whether the *proximal* (catalytic) site was affected by mutations affecting the *distal* site, namely, if the different chemical environment was relaid across the G6P substrate. 1D ^{19}F NMR was used to characterize the chemical environments of the catalytic centers in the AlF_4:G6P TSA complexes for the βPGM$_{R49K}$ and βPGM$_{R49A}$ variants under the same

conditions used for the $^{1}H^{15}N$-TROSY spectra (Fig. 6). The observed ^{19}F spectra mimicked the saturation observed by $^{1}H^{15}N$-TROSY NMR, corroborating the K_d values determined previously. Despite deleterious mutation to the *distal* site, the ^{19}F NMR peaks show minimal chemical shift perturbation in either the R49A or R49K variant, demonstrating near-identical chemical environments in the *proximal* site surrounding the AlF_4^- TSA (Fig. 6, Table 1).

To ensure that the minimal perturbation was not biased by strong Lewis basicity of the central Al^{3+} atom of the AlF_4^- moiety, 1D ^{19}F spectra of the MgF_3^- complex were acquired for both variants (as the MgF_3^- group is a looser, but more accurate analogue of phosphoryl transfer[24]). These ^{19}F spectra again show minimal chemical shift perturbation to ^{19}F resonances (Fig. 6, Table 1). The only significant perturbation ($\geq$ 1 ppm) is to the F2 resonance in the βPGM_{R49A}:MgF_3:G6P complex, where the F2 fluoride ion is coordinated by the backbone NH of L9, D10, and the sidechain OH of S114. Given the upfield shift and the relatively poor K_d for G6P binding to βPGM_{R49A}, these data are consistent with a subtle loosening of TS coordination at the F2 position.

In order to correlate the structures observed by X-ray crystallography and the ^{19}F NMR peaks observed experimentally, both AlF_4 and MgF_3 TSA complexes with R49A and R49K enzyme variants were modeled by DFT as described previously.[25] The resulting Chemical shifts agree with the assignment of experimental peaks and predict minimal perturbation of chemical shift based upon the reference crystal structures (Table 1). Together these data show that mutation of R49 and perturbation of the key bidentate interaction in the *distal* site is not communicated to the chemical environment of proximal site when the protein adopts the TSA conformation. Furthermore, the distortion of the K117 sidechain observed in the TSA structures presents a model of the catalytic form of the enzyme at steady state.

βPGM_{D170N} binds βG1P in a closed Mg^{2+}-free complex

In the R49K/A variants, it was observed that the lag-phase prior to steady state catalysis was perturbed by the removal of the guanidinium group in the *distal* phosphate binding

Table 1: ^{19}F measurements for βPGM variant complexes either by direct observation or calculated using DFT. ^{19}F resonances numbered in accordance with IUPAC recommendations[26]

Complex	F1	F2	F3	F4
Experimental ^{19}F shifts				
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	-137.9	-131.3	-140.7
βPGM_{WT}:MgF_3:G6P	-159.0	-147.0	-151.9	
βPGM_{R49K}:MgF_3:G6P	-159.2	-147.3	-152.1	
βPGM_{R49A}:MgF_3:G6P	-158.7	-148.3	-151.8	
Calculated ^{19}F shifts				
βPGM_{WT}:AlF_4:G6P	-147.0	-142.4	-133.2	-140.6
βPGM_{R49K}:AlF_4:G6P	-147.0	-142.4	-133.2	-140.7
βPGM_{R49A}:AlF_4:G6P	-147.1	-142.4	-133.2	-140.6
βPGM_{WT}:MgF_3:G6P	-148.6	-142.8	-144.7	
βPGM_{R49K}:MgF_3:G6P	-148.7	-142.7	-144.6	
βPGM_{R49A}:MgF_3:G6P	-148.9	-142.6	-145.0	

site (Fig. 4). Previous kinetic characterization of the βPGM_{WT}-catalyzed conversion of βG1P to G6P has identified a lag-phase prior to steady-state catalysis which results from two components.[12] The first component is a chemical equilibration, where it was modelled that catalysis was retarded until sufficient βG16BP intermediate was generated to efficiently prime the enzyme for catalysis (by phosphorylation of residue D8). This is termed here the βG16BP-dependent component of the lag-phase. The second component was modeled as βG1P binding to un-phosphorylated βPGM (with a K_i of 122 ± 8 μM[12]) and is termed here the βG1P-dependent component. Currently, there is no structural evidence to explain the substrate inhibition by βG1P or how the lag-phase can be alleviated.

When observed by NMR at an elevated βG1P concentration (compared to standard conditions[12]), a reduction in the βG1P-dependent component of the lag-phase was observed for both βPGM_{R49K} and βPGM_{R49A} variants, while the βG16BP-dependent component persisted (Fig. 4). These observations demonstrate a key involvement of the guanidinium group of R49 in the βG1P-dependent component of the lag-phase. βPGM_{WT} displays no observable activity in the absence of Mg^{2+},[5,11] and the reaction rate increases linearly up to at least 2 mM $MgCl_2$.[12] Given this relatively low affinity for the essential cation, it was tested whether the βG1P-dependent component of the lag-phase arises from βG1P binding to apo-

βPGM enzyme (i.e. Mg_{cat}^{2+} free enzyme) mediated by this arginine - phosphate interaction. Crystallization of βPGM_{WT}:βG1P complexes is complicated by the rapid re-equilibration of βG1P with G6P. Hence, a D170N variant of βPGM (βPGM_{D170N}) was investigated since it should be less active and the Mg^{2+} binding affinity is expected to be weaker since the Mg_{cat}^{2+} ion in the active site is coordinated by the side chain carboxylate of D170, as well as the side chain carboxylate of D8, the backbone carbonyl of D10 and two to three structural waters (depending on crystal structure used: **PDB: 1ZOL**,[5] **2WHE**[9]). When observed by 2D $^{1}H^{15}N$-TROSY NMR the substrate-free βPGM_{D170N} variant showed no deleterious effect to the overall fold compared to substrate-free βPGM_{WT} (Fig. SS1). However, due to the intermediate exchange present in the active site of the βPGM_{WT} enzyme,[9] some active site residues cannot be compared between these two proteins.

To investigate the active site of substrate-free βPGM_{D170N}, the protein was crystallized and the structure determined to 1.4 Å resolution (**PDB: 6HDF**; Table 2; Fig. 7). The substrate-free βPGM_{D170N} structure closely resembles other open-βPGM conformers (Table 5, 4). It has a poorly coordinated Na^{+} ion in the Mg_{cat}^{2+} site of both monomers in the asymmetric unit and the sidechain carbonyl group of N170 coordinates the Na^{+} ion, in place of the carboxylate group of D170 in the βPGM_{WT} structure (Fig. 7). Together this indicates that the βPGM_{D170N} variant serves as a good model of an open-βPGM enzyme with a reduced affinity for Mg_{cat}^{2+} that would permit further investigation of βG1P binding.

The binding of substrate to the βPGM_{D170N} variant was initially investigated by preparing the βPGM_{D170N}:MgF_3:G6P complex in solution and crystallizing using standard conditions.[9,27] The structure was determined to 1.2 Å resolution and was found to be the βPGM_{D170N}:βG1P complex (**PDB: 6HDG**; Table 2; Fig. 7). The observation of a different substrates in βPGM crystals compared to the initial substrate composition has been reported in both transition state analogue (TSA),[28] and ground state analogue (GSA)[11,16] complexes, and is a result of the mutase activity of βPGM equilibrating βG1P and G6P in the crystallization drop. The βPGM_{D170N}:βG1P complex adopts the closed protein confor-

mation associated with the metal fluoride TSA complexes (Table 4, 5). The electron density map shows that there is no $Mg_{cat}{}^{2+}$ or Na^{+} ion bound in the active site of the enzyme. This is associated with a *ca.* 180° rotation of the χ_1 angle of N170, which moves the carboxamide group out of the metal binding site to form a hydrogen bond with the backbone carbonyl of V188. In the active site, βG1P is bound with the 6-OH towards the *proximal* site (Fig. 7). The 6-OH occupies two positions separated by a *ca.* 120° rotation of the C5-C6 bond. This multiple occupancy facilitates hydrogen bonding with two of the three water molecules that are bound in the *proximal* site, each of which occupies near identical positions to the transferring phosphate oxygen atoms as mimicked in $MgF_3{}^{-}$ TSA complexes,[7,9,10] and in the βPGM:α-galactose 1-phosphate complex.[15] There is further similarity with the TSA structures in that the C6-O6 bond is aligned with the Oδ1 atom of D8 and there is the engagement of key catalytic residues D10 and T16 associated with full domain closure.[11]

The 1-phosphate of βG1P in this structure is bound in the *distal* phosphate binding site and, as in the TSA[10] and GSA[11,16] complexes, makes hydrogen bonds with the mainchain NH of K117, and the side chain OH of S116, as well as a bidentate interaction with the sidechain guanidinium group of R49. In order to dissect the contributions of the phosphate group and the sugar ring to induce full closure of the enzyme, crystals of βPGM_{WT} were grown in standard crystallization conditions[9,27] supplemented with 50 mM phosphate and 50 mM glucose. Crystals were briefly cryoprotected (*ca.* 30s) prior to flash freezing in their original mother liquor (with the addition of 25% ethylene glycol) either with or without the supplemented phosphate and glucose. Crystals where phosphate and glucose were not included in the cryoprotectant resulted in open βPGM_{WT} structures with no ligands bound. Crystals where phosphate and glucose were included in the cryoprotectant were comparably open structures, but resulted in the presence of a phosphate ion in the *distal* phosphate binding site (**PDB: 6H93**; Table 2). The phosphate ion makes a bidentate hydrogen bonding interaction with the sidechain guanidinium group of R49, and in one monomer, also interacts with the side chain amine groups of K117 and K76 (Fig 8). However, there is no evidence

to suggest that the binding of phosphate (and glucose) in isolation causes domain closure to the extent observed in the βPGM$_{\mathrm{D170N}}$:βG1P complex (Table 4, 5), nor cause spontaneous phosphorylation of the open enzyme.[5,9,27,29] Together, this demonstrates the key role of a covalent bond between phosphate and glucose, particularly when the reported K_i is so poor ($K_i = 122 \pm 8$ μM[12]).

Discussion

Here we have structurally characterized βG1P bound to βPGM in a non-catalytically competent, closed complex, which provides a structural basis for the βG1P inhibition of the substrate-free enzyme postulated previously in kinetic models.[12] Furthermore, it is demonstrated that single mutations in the *distal* phosphate site can alleviate the βG1P-dependent component of the lag-phase prior to steady state catalysis, implicating a role for the specific bidentate hydrogen bonding interaction between phosphate in the *distal* site and the terminal guanidinium group of residue R49 in the cap domain. MgF_3^- and AlF_4^- transition state analogue complexes with G6P demonstrate minimal perturbation to the *proximal* phosphate binding site (at the point of phosphoryl transfer) in response to removal of the R49 guanidinium group in the *distal* phosphate binding site. This minimal communication between the two sites indicates that the role of the *distal* site is primarily to recruit ligand into the active site and contribute to domain closure prior to the chemical step, while playing a minimal role in the chemical step itself.

Both MgF_3^- and AlF_4^- TSA crystal structures of the βPGM$_{\mathrm{R49K}}$ and βPGM$_{\mathrm{R49A}}$ variants complexed with G6P demonstrated a redundancy in the phosphate coordination in the *distal* phosphate binding site. In the βPGM$_{\mathrm{R49K}}$ variant, the positive charge was maintained in the *distal* site, but with a reduced hydrogen bonding capacity. This manifests as a reduced stability for substrate bound complexes, but given that the charge is maintained in this conservative mutation, the small reduction in k_{obs} is readily rationalizable. In the

βPGM$_{\mathrm{R49A}}$ variant, though charge balance was nominally removed from the *distal* site on the cap-domain side, there was still a substantial level of activity. In both of the TSA structures with G6P, K117 from the core-domain (which is solvent exposed in βPGM$_{\mathrm{WT}}$ and βPGM$_{\mathrm{R49K}}$ complexes), is repositioned in order to coordinate the *distal* phosphate group of G6P via its sidechain amine group. If this occurs in solution, which is consistent with the chemical shift changes in the ^{1}H^{15}N-TROSY spectra, then this presents a redundancy in phosphate binding capability in the *distal* phosphate binding site of βPGM. This conformer also indicates that a competitive binding interaction between K117 and R49 may exist in βPGM$_{\mathrm{WT}}$. This alternative binding partner for phosphate groups in the *distal* site could present a pathway to ligand dissociation from the active site, prior to either reorientation of βG16BP or product release of G6P or βG1P.

Guanidinium - phosphate interactions have been reported to provide substantial binding energies in the range 11-13 kcal/mol for glycerol 3-phosphate dehydrogenase (GPDH),[17] for triose phosphate isomerase (TIM),[18] and for orotidine 5-monophosphate decarboxylase.[19] This phosphodianion binding has been associated with a protein conformational change and active site assembly in other systems such as orotidine 5-monophosphate decarboxylase[20,21] and for GPDH.[22] Furthermore, it has been demonstrated that the energetic cost of disconnecting groups of either substrate or enzyme GPDH[17,21,30] and TIM[30,31] was directly reflected in the reduction in observed reaction rate. This suggests that the transition state of the reaction in the re-assembled complex closely reflected transition states of the native reaction (reviewed[32]). In βPGM, binding of the phosphodianion to the *distal* site is insufficient to close the enzyme. Furthermore, the inclusion of both phosphate and glucose is insufficient to re-assemble the inhibited βG1P-bound complex. This is relatively unsurprising as the K_i for the βG1P-dependent contribution to the lag-phase is reportedly 122 $\pm$ 8 μM[12] and disconnection of substrate components would likely incur an energetic penalty, further destabilizing binding.

The bidentate interaction between the sugar-associated phosphate group and the guani-

dinium group of R49 is well conserved across most substrate bound complexes in βPGM. The *ca.* 1000 fold reduction in K_d for AlF_4:G6P TSA complex (*ca.* 4.1 kcal mol^{-1}) in the R49K/A variants approximates to the energy associated with the loss of a single hydrogen bond (*ca.* 1.5 – 4 kcal mol^{-1}).[33] If TSA stabilization is equated to transition state (TS) stabilization, then these observations correlate well with the significant contribution of an Arg group binding to phosphate in GPDH. In that case, mutation to alanine (without the compensation seen in βPGM) resulted in a 9.1 kcal mol^{-1} destabilization of the TS for enzyme catalyzed reduction of DHAP.[21] The ^{19}F NMR of the TSA complexes in βPGM indicate that the electrostatic environment[25,34] surrounding the TSA in the *proximal* site is not perturbed by mutation of R49 in the *distal* site. When this is taken with *ca.* 4.1 kcal mol^{-1} destabilization of the AlF_4^-:G6P TSA on the R49K/A variants, it suggests that the reduction in k_{obs} at 10 mM βG1P is a result of an increase in K_m value, rather than a decrease in k_{cat}. This increase in K_m apparently also translates to an increase in K_i value for βG1P inhibition as no βG1P-dependent lag-phase was observed in either of the R49K/A variants.

Structural evidence to support the inhibition of βPGM by βG1P is presented, where the closed βPGM_{D170N}:βG1P complex closely resembles fully closed TSA structures of phosphoryl transfer, with key residues in the active site adopting catalytic orientations, but without a phosphoryl group to transfer and without a metal ion in the active site. The sidechain of residue N170 is rotated away from the *proximal* phosphate site, which potentially implicates D170 dissociation from the active site as a mechanism to release the Mg_{cat}^{2+} ion in the WT enzyme. This may be important in the dissociation of the reaction intermediate βG16BP, since it has a high affinity to holo-βPGM (Mg_{cat}^{2+} bound; $K_m = 0.8 \pm 0.2$ μM[12]). When a protonated general acid base variant (βPGM_{D10N}) was used to trap the βG16BP intermediate in the active site of βPGM , the resulting βPGM_{D10N}:βG16BP complex displayed a relatively weak (7.1 $\pm$ 0.6 mM) Mg_{cat}^{2+} binding affinity.[11] Given that the βG16BP ligand has a higher binding affinity than βG1P or G6P ligands for substrate-free βPGM, it is important that this state does not become a kinetic trap. It is tempting to speculate that

βPGM uses one (or both) of the above ligand dissociation pathways - poor $Mg_{cat}{}^{2+}$ binding affinity, and competitive phosphate binding in the *distal* site - to avoid such a kinetic trap.

Taken together, these observations illustrate some of the elegant mechanisms that enzymes employ in order to achieve the significant rate enhancements necessary for life. Here we see a *ca.* 10 fold rate enhancement through the use of a guanidinium group (WT) over an amine group (R49 variants) to coordinate the phosphate in the *distal* site. This rate enhancement, however, is at the expense of introducing a source of inhibition to catalysis – inhibition of the substrate-free enzyme by its initial substrate βG1P.

Methods

βPGM

The *pgmB* gene from *Lactococcus lactis*, together with the *pgmB* gene containing the D170N, R49A and R49K mutations were expressed using pET22b+ vectors in *E. coli* strain BL21(DE3) using protocols outlined previously.[11] All NMR experiments were recorded in standard NMR buffer; 50 mM K^+ HEPES buffer (pH 7.2) containing 5 mM $MgCl_2$, 2 mM NaN_3, 1 mM TSP, and 10% (vol/vol) D_2O at 298K, unless otherwise stated. Site directed mutagenesis and DNA sequencing was performed by GenScript (HK) to generate the R49K, R49A, and D170N mutants.

Reagents

Reagents and buffers including glucose-6-phosphate and lithium potassium acetylphosphate were purchased from Sigma or Melford labs. β-glucose-1-phosphate (βG1P) was prepared in house from Maltose (Sigma) using Maltose phosphorylase (Sigma) as follows: 1M maltose in 0.5M phosphate buffer pH 7 was reacted with 1.5units/ml of Maltose Phosphorylase at 30°C overnight. βG1P production was confirmed by ^{31}P NMR. Maltose phosphorylase (90 kDa) was then removed by passing the solution through a 5 kDa MWCO vivapsin, and the

reaction mixture was used without further purification. The resulting concentrations were *ca.* 150 mM βG1P, 150 mM glucose, 350 mM maltose and 350 mM phosphate.

Reaction kinetics by glucose 6-phosphate dehydrogenase coupled assay

Steady-state kinetic assays for substrate-free βPGM_{R49K} and βPGM_{R49A} were conducted at 294 K using a FLUOstar OMEGA microplate reader (BMG Labtech) in standard kinetic buffer (200 mM K^+ HEPES buffer (pH 7.2) containing 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 oxidized and concomitant NAD^+ reduction is monitored by the increase in absorbance at 340 nm (NADH extinction coefficient = 6220 M^{-1} cm^{-1}). Enzyme stock concentrations were determined using a NanoDrop One C spectrophotometer (Thermo Scientific) and diluted accordingly (βPGM extinction coefficient = 19940 M^{-1} cm^{-1}).

Reaction kinetics by ^{31}P NMR

^{31}P NMR spectroscopy observed reaction kinetics for βPGM-catalyzed reactions were followed at 298K on a Bruker 500 MHz Avance III HD spectrometer (operating at 202.48 MHz for ^{31}P) equipped with a 5-mm Prodigy BBO cryoprobe (School of Chemistry, University of Manchester). One-dimensional ^{31}P spectra without proton decoupling were recorded within 1 minute with 16 transients and a 2s recycle delay to give signal-to-noise ratios for 10 mM βG1P of greater than 100:1. The turnover of 10 mM βG1P to G6P by βPGM_{WT} (0.1 - 1 μM), βPGM_{R49A} βPGM_{R49K} (10 - 50 μM) were measured in standard kinetic buffer (200 mM K^+ HEPES buffer (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3) with the addition of 10% D2O 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 spectra. Integral values of the G6P peak

following baseline correction and 2 Hz Lorentzian apodization were plotted against time to give kinetic profiles. The linear portion of the data was fitted using a linear least-squares fitting algorithm to derive the observed catalytic rate constant.

NMR characterization of R49 variants complexed with AlF_4 and G6P

AlF_4:G6P TSA complexes with βPGM_{R49A} or βPGM_{R49K} were made by addition of 5mM AlF_4 and 20mM G6P to 1mM enzyme in standard NMR buffer (50 mM K^+ HEPES pH 7.2, 5 mM $MgCl_2$, 2 mM NaN_3, and 1 mM TSP). ^{19}F and $^{1}H^{15}N$-TROSY experiments were recorded with a Bruker Avance III 500 MHz spectrometer using a 5-mm QCI-F cryo-probe equipped with z-axis gradients (Manchester Insitiute of Biotechnology). ^{19}F 1D spectra were acquired without proton decoupling and were processed using an EM window function with 10 Hz linebroadening. $^{1}H^{15}N$-TROSY spectra were acquired using a pulse sequence with echo/anti-echo gradient selection and were processed without linear prediction in either dimension. Data were processed using Topspin and direct referencing to TSP at 0.0 ppm was applied for TROSY spectra, while indirect referencing (Bruker standard referencing) was used for ^{19}F spectra.

X-ray crystallography

βPGM was prepared at a concentration of 15 mg mL^{-1} in 50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 1 mM NaN_3. For crystallization, the enzyme solution was mixed 1:1 with the precipitant (26-30% (wt/vol) PEG 4000, 200 mM Na acetate, and 100 mM Tris (pH 7.5)) and placed in sitting-drop crystallization plates with 700 μl reservoir volume.

For the R49K and R49A crystal structures, 0.6 mM enzyme was mixed with 20 mM NaF, 5 mM $AlCl_3$, and 10 mM G6P (in that order) and incubated for >10 min prior to the crystallization trial. The initial crystallization conditions for the βPGM_{D170N}:βG1P

structure were 0.6 mM enzyme mixed with 20 mM NaF, 5mM $MgCl_2$, and 10 mM G6P. Notably D170N did not form the intended MgF_3:G6P TSA, instead, the partly inhibited enzyme inter-converted G6P to βG1P, with only βG1P observed in the active site of the closed enzyme. Crystallization trials yielded thin plate-like crystals after several days.

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[35,36] with resolution cut-offs applied using CC-half values and the structures were determined by molecular replacement with MolRep[37] using previously modelled βPGM PDB structures as a search models. Model building was carried out in COOT[38] and either a restrained refinement with isotropic temperature factors (resolution worse than 1.5Å) or anisotropic temperature factors (resolutions better than 1.5Å) was performed using REFMAC5[39] in the CCP4i suite[40]. Ligands were not included until the final stages of refinement to avoid biasing Fourier maps. Structure validation was carried out in COOT and MolProbity[41], superpositions were generated using PyMOL (The PyMOL Molecular Graphics System, version 1.8/2.2 Schrödinger, LLC), maps were generated using FFT[42] and domain movements were calculated using DynDom[43].

KS-DFT ^{19}F chemical shift calculations

Our model for the transition state analogue (TSA) of the βG16BP hydrolysis reaction carried out by βPGM was obtained using Kohn-Sham Density Functional Theory (KS-DFT). We used the B3LYP functional formulation of KS-DFT.[44–47] A 6-31G Pople basis set was used to represent single-particle wavefunctions for all atoms excepting fluorine atoms for which more care was given due to NMR calculation sensitivity to basis set. For these 4 atoms, 6-31+G(d) was used, as diffuse and polarization functions are required to adequately describe the electron distribution around each atom.[48]

The active site (cluster) model, initially obtained from a high-resolution X-ray structure (PDB: 2WF5[7]), was constructed so as to maintain all key hydrogen bonding capable of

stabilizing the transition state. More specifically, we included residues F7, D8, L9, D10, G11, V12, I13, T16, A17, H20, W24, K45, L44, G46, V47, S48, R49, E50, D51, S52, Y80, A113, S114, A115, S116, K117, K145, F151, L168, E169, D170, S171. 19 explicit water molecules as well as a catalytic magnesium ion were also retained. Where opportune, we truncated amino acid residues with a methyl group. Geometry optimization of the entire (602 atoms) model was considered unfeasible with current computational resources, and so we chose to optimize those residues most proximal to the fluorine atoms of the AlF_4^- group, for which we were interested in calculating shielding tensors. To this extent, the following groups were optimized; substrate, AlF_4^-, catalytic Mg^{2+} ion, 3 explicit water molecules, selected residues (D8, D10, S114, A115, S116, K145), with the remaining groups fixed at their crystallographic coordinates in the X-ray crystal structure. We optimized the geometry of the resulting active site model using standard algorithms,[49] as implemented in the Gaussian09 software package.[50]

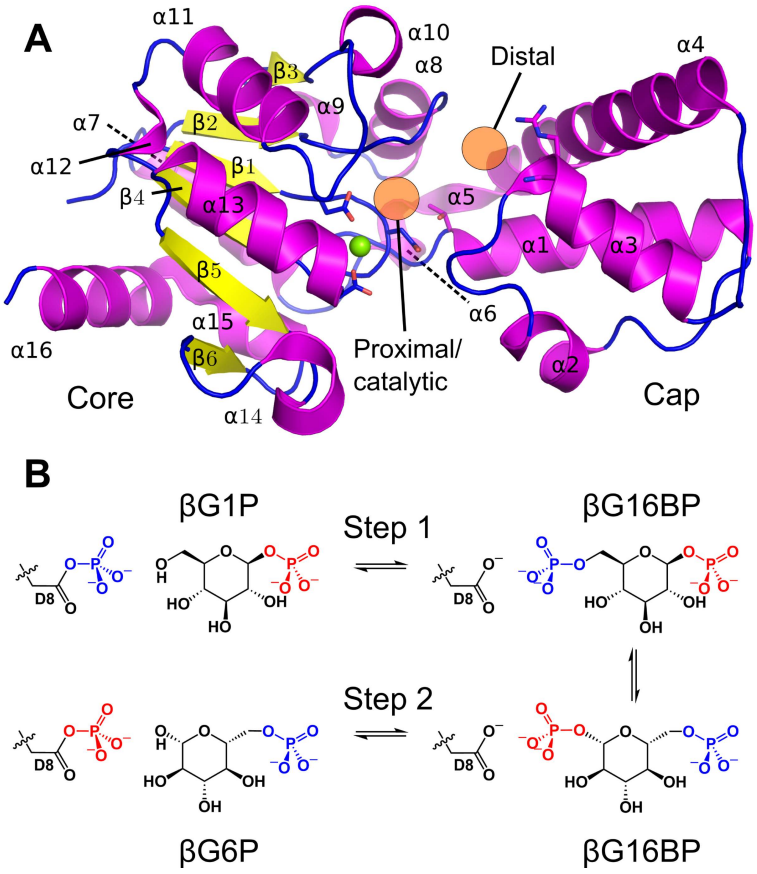


Figure 1: Overview of βPGM structure (A), and the reaction scheme catalyzed by βPGM (B). **A)** The Rossman fold of βPGM is illustrated with a HADSF four-helix-bundle cap domain (type C1[13]). α-helices are colored in purple, β-sheets in yellow, and loops are illustrated in blue. The Mg_{cat}^{2+} ion is depicted as a green sphere, while the *proximal* (catalytic) and *distal* phosphate binding sites are illustrated as orange circles. **B)** The βPGM reaction scheme 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 termed Step 1 . Here the transferring phosphate (blue) occupies the *proximal* site and the 1-phosphate (red) of βG1P occupies the *distal* site. Following formation of βG16BP, the βG16BP intermediate dissociates from the enzyme and re-binds in an orientation such that the 1-phosphate (red) occupies the proximal site, and the 6-phosphate (blue) occupies the *distal* site. The phosphoryl transfer reaction between βG16BP and D8 is termed Step 2 and generates βPGM^{P} and G6P.

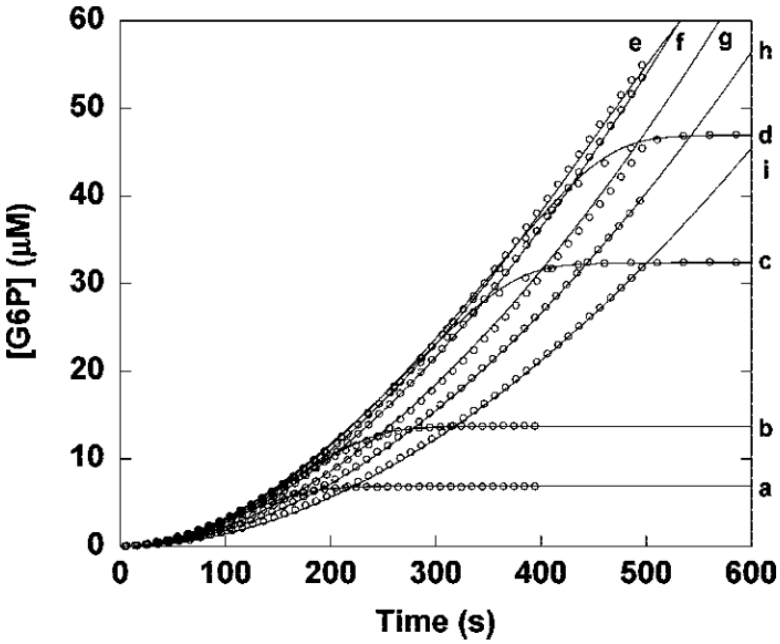


Figure 2: Kinetic profile image adapted from[12] illustrating the effect of increasing βG1P concentration on a pre-steady state kinetic lag-phase observed using a glucose 6-phosphate dehydrogenase coupled assay. 4 nM βPGM enzyme was used for concentrations of βG1P (a) 7.5, (b) 15, (c) 35, (d) 50, (e) 70, (f) 100, (g) 160, (h) 230, and (i) 330 μM with 50 μM αG16BP used as a priming agent. 5 U/mL G6PDH was used to convert 0.5 mM NAD^+ in 50 mM K^+ HEPES buffer pH 7.2, with 2mM $MgCl_2$ at 25 °C.

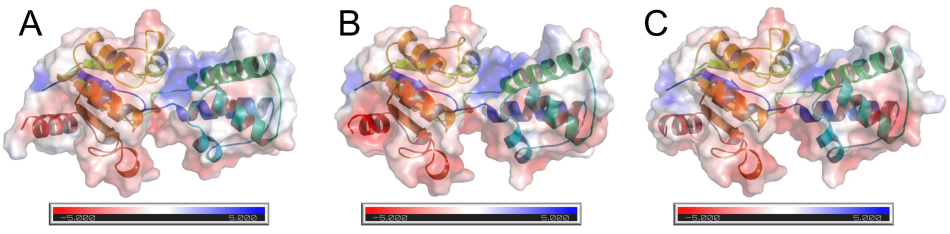


Figure 3: Electrostatic surface shown for open- βPGM_{WT} (A), βPGM_{R49K} (B), and βPGM_{R49A} (C) variants. Surface prepared using pdb2pqr and APBS in pymol 2.2 (Schrödinger). The open R49-variant structures overlay closely with the βPGM_{WT} structure with non-H atom RMSDs of 0.169 and 0.196 for βPGM_{R49K} and βPGM_{R49A} respectively. A reduced positive charge is observed in the distal phosphate binding site of the R49A variant which is expected given the loss of a positively charged guanidinium group.

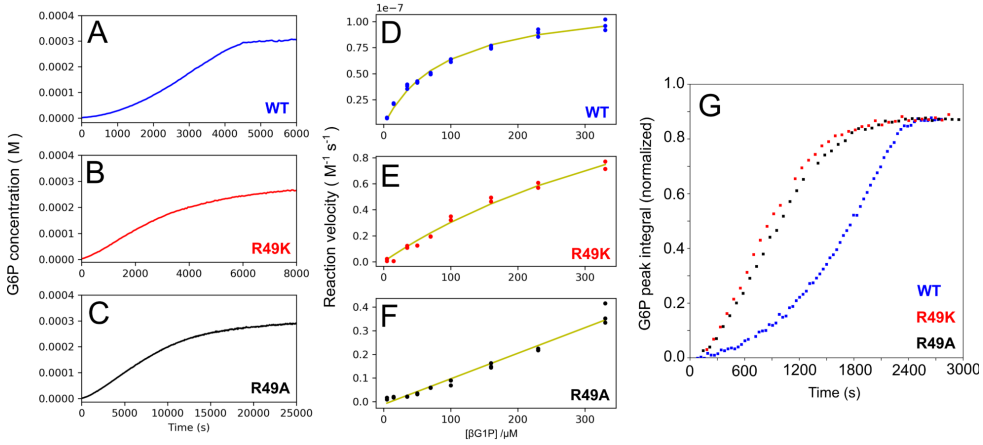


Figure 4: Kinetic profiling of βPGM WT, R49K and R49A variants. **(A-C)** The βPGM catalyzed enzymatic conversion of 330 μM βG1P to G6P followed by spectrophotometric coupled assay for **(A)** βPGM_{WT} [5nM], **(B)** βPGM_{R49K} [60nM] , **(C)** βPGM_{R49A} [60nM]. **(D-E)** Reaction velocity as a function of βG1P concentration for **(D)** βPGM_{WT} (data presented in[11]) , **(E)** βPGM_{R49K}, **(F)** βPGM_{R49A}, with line of best fit drawn in yellow. Both R49K and R49A variants of βPGM displayed a linear concentration dependence with R=0.98. **(G)** G6P integral against time for each of the three variants when the reaction is followed by ^{31}P 1D NMR. In this case the βG1P concentration was 10 mM and the enzyme concentrations used were 0.1 μM (βPGM_{WT}), 0.5μM (βPGM_{R49K}), and 1.0 μM (βPGM_{R49A}).

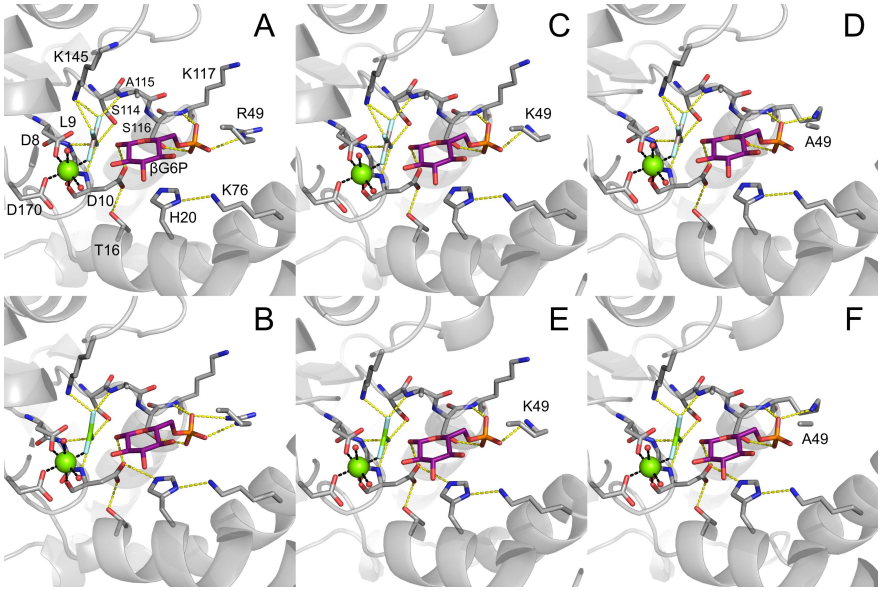


Figure 5: Crystal structures of TSA complexes for R49K and R49A variants. **(A)** βPGM$_{\mathrm{WT}}$:AlF$_4$:G6P complex (PDB: 2WF6;[9]), **(B)** βPGM$_{\mathrm{WT}}$:MgF$_3$:G6P complex (PDB: 2WF5;[9]), **(C)** βPGM$_{\mathrm{R49K}}$:AlF$_4$:G6P complex (PDB: 6HDJ), **(D)** βPGM$_{\mathrm{R49A}}$:AlF$_4$:G6P complex (PDB: 6HDK), **(E)** βPGM$_{\mathrm{R49K}}$:MgF$_3$:G6P complex (PDB: 6HDL), **(F)** βPGM$_{\mathrm{R49A}}$:MgF$_3$:G6P complex (PDB: 6HDM). Selected active site residues are illustrated alongside water molecules (small red spheres). Atoms are drawn using the coloring; gray=C, blue=N, red=O, green=Mg, orange=P. Hydrogen bonds are drawn as yellow dashed lines, metal ion coordination as black dashed lines, and the G6P ligand illustrated with purple carbon atoms for clarity.

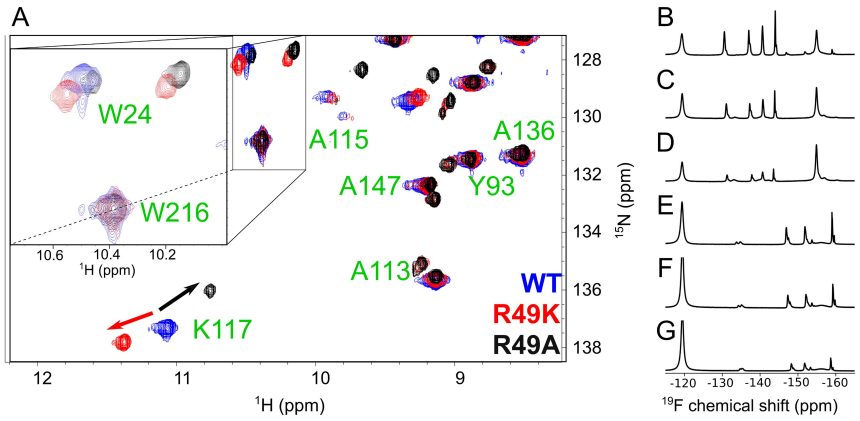


Figure 6: NMR spectra of WT, R49K, and R49A enzyme variants in TSA complexes with G6P. **(A)** The 2D $^{1}H^{15}N$-TROSY NMR spectra of all three variants overlaid, with an inlay illustrating the sidechain indole amides of W24 and W216. There is a hydrogen bond between the 6-phosphate oxygen of G6P and the backbone amide of K117, which causes such a marked downfield shift of the backbone amide resonance. The chemical shift of K117 backbone amide in each of the complexes demonstrates a tighter hydrogen bond between phosphate-oxygen and amide in the βPGM_{R49K}:AlF_4:G6P complex (red arrow), but a weaker hydrogen bond in the βPGM_{R49A}:AlF_4:G6P complex (black arrow). (B-G) ^{19}F 1D NMR spectra of each of the variants complexed with metal fluorides and G6P. **(B)** βPGM_{WT}:AlF_4:G6P, **(C)** βPGM_{R49K}:AlF_4:G6P, **(D)** βPGM_{R49A}:AlF_4:G6P, **(E)** βPGM_{WT}:MgF_3:G6P, **(F)** βPGM_{R49K}:MgF_3:G6P, **(G)** βPGM_{R49A}:MgF_3:G6P. In all cases, βPGM concentration was 1 mM , with the addition of 5 mM $MgCl_2$, 15 mM NaF, 20 mM G6P, and 3 mM $AlCl_3$ (where applicable).

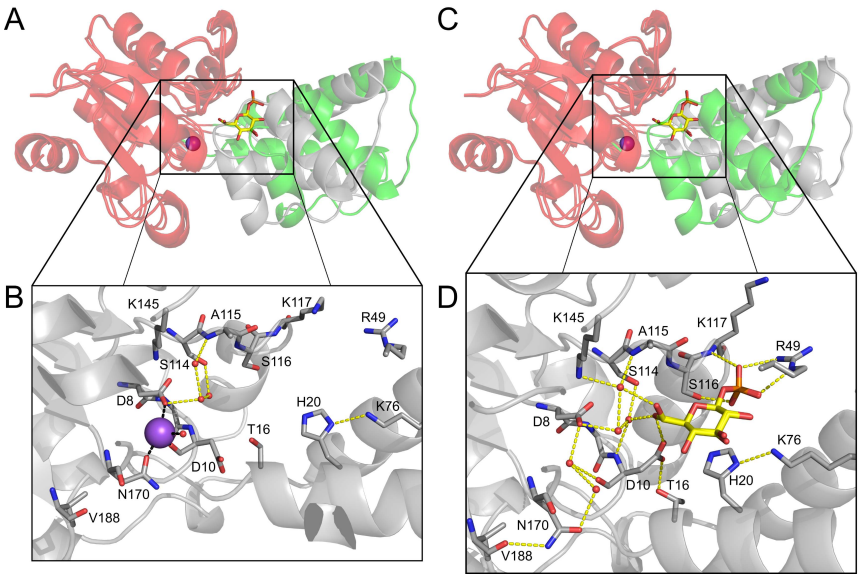


Figure 7: Crystal structures of βPGM$_{\mathrm{D170N}}$ in both open (A and B), and βG1P-complexed (C and D), states. In **A)** and **C)**, the open-βPGM$_{\mathrm{D170N}}$ and βPGM$_{\mathrm{D170N}}$:βG1P complexes are overlaid (aligned on core domain, non-H atom RMSD=0.423) with the core domain (left) shaded in red, and the cap domain (right) shaded in green for either the open-βPGM$_{\mathrm{D170N}}$ complex **(A)** or the βPGM$_{\mathrm{D170N}}$:βG1P complex **(C)**. In **B)** and **D)** selected active site residues are illustrated alongside water molecules (small red spheres). Atoms are drawn using the coloring; gray=C, blue=N, red=O, purple=Na, orange=P. Hydrogen bonds are drawn as yellow dashed lines, metal ion coordination as black dashed lines, and the βG1P ligand illustrated with yellow carbon atoms for clarity.

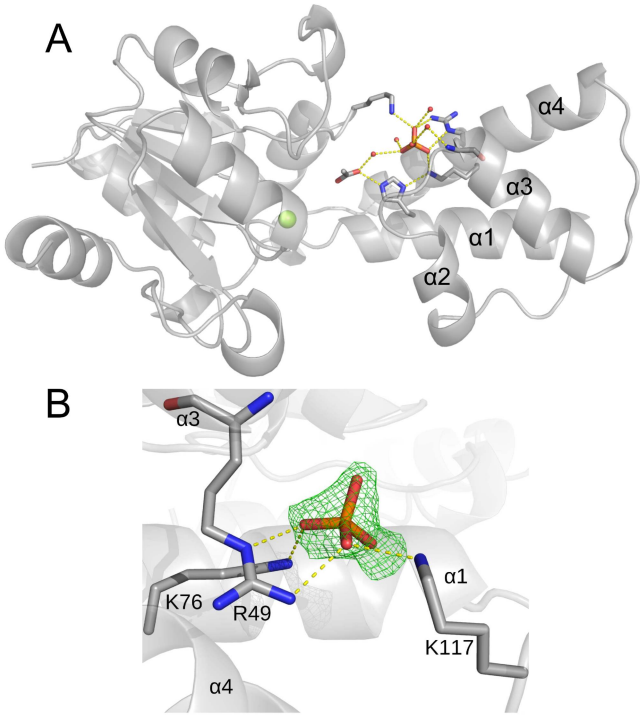


Figure 8: Binding of un-phosphorylated glucose to open-βPGM. The structure of the βPGM$_{\mathrm{WT}}$:Pi complex is presented in **A)** with selected hydrogen bonds to the phosphate group and active site residues indicated by dashed yellow lines. In both A and B atoms are drawn using the coloring; gray=C, blue=N, red=O, purple=Na, orange=P. **B)** The phosphate coordinated in the *distal* site with omit map density (green) contoured at 3 σ, hydrogen bonds are indicated by dashed yellow lines.

Table 2:

[1]Values for the higher resolution shell are in parenthesis.

[2] $R_{merge} = \Sigma_{hkl}\Sigma_i \mid I_i - I_m \mid / \Sigma_{hkl}\Sigma_i I_i$.

[3] $R_{pim} = \Sigma_{hkl} \sqrt{1/n\text{-}1}\Sigma_{i=1} \mid I_i - I_m \mid / \Sigma_{hkl}\Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively.

[4] $R = \sum_{hkl} || F_{obs} \mid - k \mid F_{calc} || \, I \sum_{hkl} \mid F_{obs} \mid$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.

[5] $R_{free} = \sum_{hklIT} || F_{obs} \mid - k \mid F_{calc} || \, I \sum_{hklIT} \mid F_{obs} \mid$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes and T is the test set of data omitted from refinement (5% in this case).

[6] For structures where there are two proteins in the asymmetric unit, the value for chain A will be given first, then the value for chain B.

[7] Only the ligands that are the subject of investigation are presented. Other ligands such as ethylene glycol and acetate, etc. that are part of the crystallization solution are not presented.

[8] Frequently the only observable metal ion was Mg^{2+}, however in some cases Na^{+} ions were also observed. Where this was the case, the B-factor for Na^{+} ions were given first, followed by Mg^{2+} ions.

Data Acquisition

Complex	open-βPGM$_{D170N}$	βPGM$_{D170N}$:βG1P	βPGM$_{WT}$: Pi	open-βPGM$_{R49K}$	open-βPGM$_{R49A}$
PDB code	6HDF	6HDG	6H93	6HDH	6HDI
Wavelength (Å)	0.92819	0.92819	0.97950	0.97624	0.97625
Beamline	I04-1	I04-1	i04	i03	i03
Facility	DLS	DLS	DLS	DLS	DLS
Space group	$P2_1$	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$P2_1$
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 (°)	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)
R_{merge} [1,2]	0.091 (1.243)	0.126 (1.986)	0.085 (1.426)	0.052 (0.770)	0.131 (0.829)
R_{pim} [1,3]	0.055 (0.701)	0.037 (0.634)	0.035 (0.583)	0.030 (0.487)	0.077 (0.488)
CC-half	0.997 (0.462)	0.999 (0.562)	0.999 (0.467)	0.998 (0.491)	0.992 (0.525)
$< I/\sigma I >$ [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	2WHE	2WF5	2WHE	2WHE	2WHE

Data Refinement

Complex	open-βPGM$_{D170N}$	βPGM$_{D170N}$:βG1P	βPGM$_{WT}$: Pi	open-βPGM$_{R49K}$	open-βPGM$_{R49A}$
PDB code	6HDF	6HDG	6H93	6HDH	6HDI
R(%) [4] / R_{free} (%) [5]	17.2 / 22.3	14.9 / 17.8	17.9 / 23.3	18.2 / 21.8	20.5 / 27.3
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
RMS deviations:					
Length	0.01	0.01	0.12	0.01	0.01
Angles	1.50	1.51	1.49	1.50	1.51
Average B factors(Å^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					
Favored/allowed (%)	98.19	97.80	98.39	98.61	97.71
Disallowed (%)	0.00	0.00	0.00	0.00	0.00
Molprobity score (percentile)	0.76 (100^{th})	0.86 (100^{th})	0.73 (100^{th})	0.97 (100^{th})	1.16 (100^{th})

Table 3:

[1]Values for the higher resolution shell are in parenthesis.

[2] $R_{merge} = \Sigma_{hkl}\Sigma_i \mid I_i - I_m \mid / \Sigma_{hkl}\Sigma_i I_i$.

[3] $R_{pim} = \Sigma_{hkl} \sqrt{1/n-1}\Sigma_{i=1} \mid I_i - I_m \mid / \Sigma_{hkl}\Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively.

[4]$R = \sum_{hkl} || F_{obs} \mid - k \mid F_{calc} || \; I \sum_{hkl} \mid F_{obs} \mid$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.

[5]$R_{free} = \sum_{hkl \; I \; T} || F_{obs} \mid - k \mid F_{calc} || \; I \sum_{hkl \; I \; T} \mid F_{obs} \mid$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes and T is the test set of data omitted from refinement (5% in this case).

[6] For structures where there are two proteins in the asymmetric unit, the value for chain A will be given first, then the value for chain B.

[7] Only the ligands that are the subject of investigation are presented. Other ligands such as ethylene glycol and acetate, etc. that are part of the crystallization solution are not presented.

[8] Frequently the only observable metal ion was Mg^{2+}, however in some cases Na^{+} ions were also observed. Where this was the case, the B-factor for Na^{+} ions were given first, followed by Mg^{2+} ions.

Data Acquisition

Complex	βPGM$_{R49K}$:AlF_4:G6P	βPGM$_{R49A}$:AlF_4:G6P	βPGM$_{R49K}$:MgF_3:G6P	βPGM$_{R49A}$:MgF_3:G6P
PDB code	6HDJ	6HDK	6HDL	6HDM
Wavelength (Å)	0.97625	0.97625	0.97629	0.97625
Beamline	i03	i03	i03	i03
Facility	DLS	DLS	DLS	DLS
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions				
a, b, c, (Å)	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 (^{o})	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]	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)
R_{merge} [1,2]	0.084 (1.082)	0.099 (1.019)	0.068 (1.345)	0.052 (0.263)
R_{pim} [1,3]	0.033 (0.460)	0.040 (0.480)	0.027 (0.591)	0.022 (0.138)
CC-half	0.999 (0.554)	0.999 (0.530)	0.999 (0.515)	0.999 (0.944)
$< I/\sigma I >$ [1]	11.2 (1.5)	10.1 (1.4)	14.2 (1.3)	21.1 (6.2)
Completeness (%) [1]	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	70516	60728	73452	53048
Molecular replacement model	2WF6	2WF6	2WF5	2WF5

Data Refinement

Complex	βPGM$_{R49K}$:AlF_4:G6P	βPGM$_{R49A}$:AlF_4:G6P	βPGM$_{R49K}$:MgF_3:G6P	βPGM$_{R49A}$:MgF_3:G6P
PDB code	6HDJ	6HDK	6HDL	6HDM
R(%) [4] / R_{free} (%) [5]	14.3 / 16.6	13.6 / 16.7	13.2 / 14.4	12.6 / 14.8
Protein [6]	1739	1706	1774	1802
Ligands [7]	21	21	20	20
Metal ions [8]	1	3	2	2
Water	179	212	278	213
Protein residues[6]	219	219	219	219
RMS deviations:				
Length	0.01	0.01	0.01	0.01
Angles	1.50	1.50	1.50	1.47
Average B factors(A^2)				
Main chain [6]	14.1	13.7	14.0	13.3
Side chains [6]	17.6	17.7	16.9	16.5
Ligands [7]	9.1, 8.8	9.0, 9.5	11.6, 10.9	10.0, 9.1
Metal Ions [8]	8.6	23.1, 8.9	15.6, 9.3	14.4, 8.4
Water	26.1	28.4	27.5	24.1
Ramachandran analysis				
Favored/allowed (%)	98.21	97.74	97.40	98.70
Disallowed (%)	0.00	0.00	0.00	0.00
Molprobity score (percentile)	0.76 (100^{th})	0.82 (100^{th})	1.07 (98^{th})	1.19 (95^{th})

Table 4: A structural comparison of all complexes discussed in the text using and DynDom[43] to generate angles (°) of rotation to map cap domains between different complexes and variants of βPGM. Comparisons where no dynamic domains were found are denoted with an angle of 0.0.

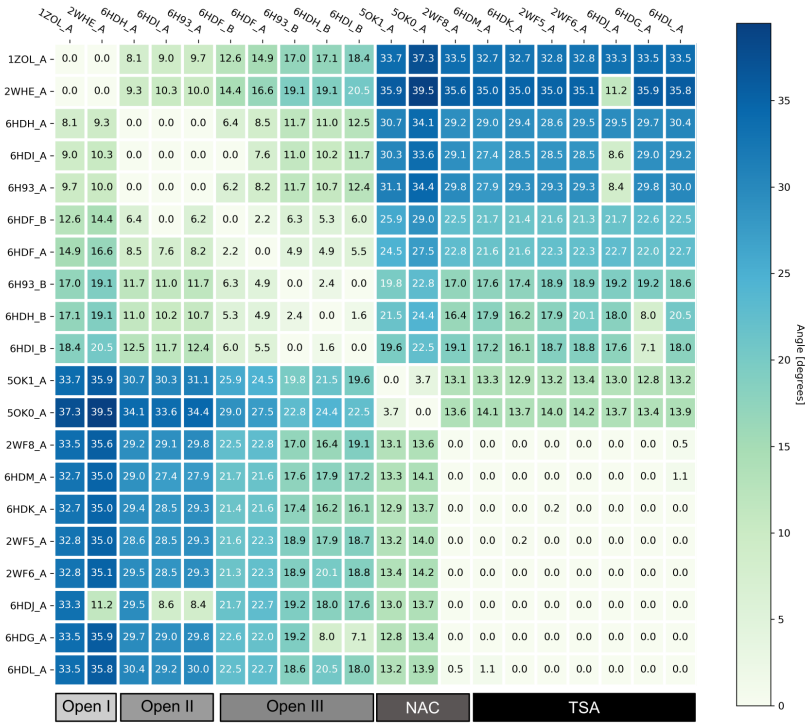
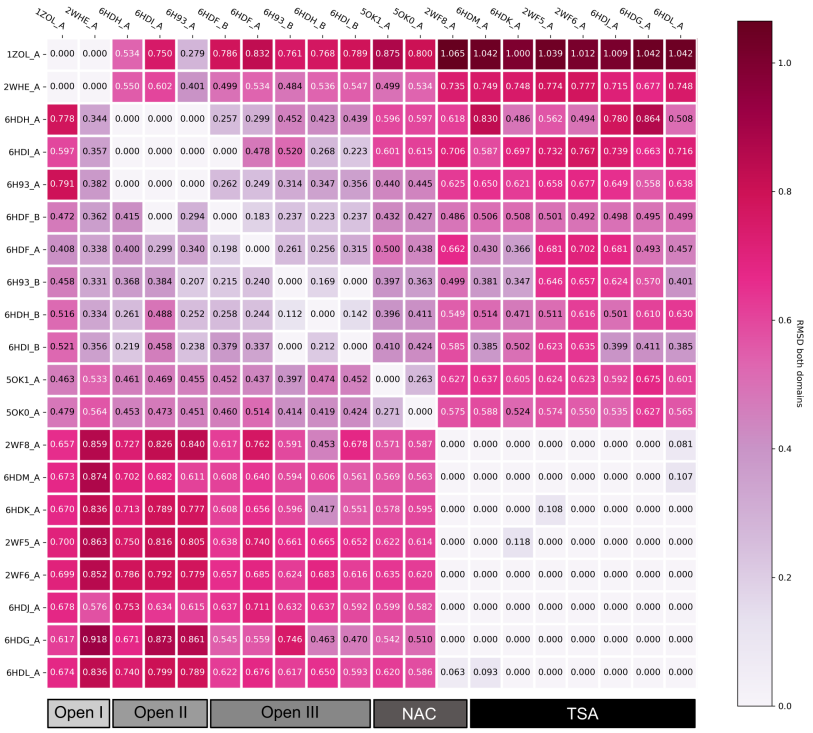


Table 5: A structural comparison of all complexes discussed in the text using and DynDom[43] to generate non-H atom RMSD values for aligned cap and core domain between different complexes in βPGM. The top right side of the matrix indicates the RMSD value for the core domain (Å), while the bottom left side of the matrix indicates the RMSD for the cap domain (Å). The diagonal is necessarily indicates no difference. Comparisons where no dynamic domains were found are denoted with an RMSD of 0.0.



References

(1) 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. *Proc Natl Acad Sci U S A 100*, 5607–5610.

(2) Lassila, J. K., Zalatan, J. G., and Herschlag, D. (2011) Biological Phosphoryl-Transfer Reactions: Understanding Mechanism and Catalysis. *Annual Review of Biochemistry 80*, 669–702.

(3) Kamerlin, S. C., Sharma, P. K., Prasad, R. B., and Warshel, A. (2013) Why nature really chose phosphate. *Quarterly Reviews of Biophysics 46*, 1–132.

(4) Lahiri, S. D., Zhang, G., Dai, J., Dunaway-Mariano, D., and Allen, K. N. (2004) Analysis of the Substrate Specificity Loop of the HAD Superfamily Cap Domain. *Biochemistry 43*, 2812–2820.

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

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

(7) Baxter, N. J., Olguin, L. F., Golic, 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 MgF_3^- formation in an enzyme active site. *Proceedings of the National Academy of Sciences (USA) 103*, 14732–14737.

(8) 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*, 3952–3958.

(9) 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 MgF_3^- rather than by phosphoranes. *Proceedings of the National Academy of Sciences (USA) 107*, 4555–4560.

(10) 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) α-Fluorophosphonates reveal how a phosphomutase conserves transition state conformation over hexose recognition in its two-step reaction. *Proceedings of the National Academy of Sciences (USA) 111*, 12384–12389.

(11) 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*, 8140–8153.

(12) 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*, 1575–88.

(13) Huang, H. et al. (2015) Panoramic view of a superfamily of phosphatases through substrate profiling. *Proceedings of the National Academy of Sciences (USA) 112*, E1974–E1983.

(14) 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 β-Phosphoglucomutase Catalysis. *Biochemistry 48*, 1984–1995.

(15) Tremblay, L. W., Zhang, G., Dai, J., Dunaway-Mariano, D., and Allen, K. N. (2005) Chemical Confirmation of a Pentavalent Phosphorane in Complex with β-Phosphoglucomutase. *Journal of the American Chemical Society 127*, 5298–5299, PMID: 15826149.

(16) 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 β-phosphoglucomutase complexes where geometry and charge distribution reflect those of substrate. *Proceedings of the National Academy of Sciences (USA) 109*, 6910–6915.

(17) 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*, 4575–4582, PMID: 18376850.

(18) 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*, 11325–11326, PMID: 11697989.

(19) 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*, 15708–15709, PMID: 16277505.

(20) Desai, B. J., Wood, B. M., Fedorov, A. A., Fedorov, E. V., Goryanova, B., Amyes, T. L., Richard, J. P., Almo, S. C., and Gerlt, J. A. (2012) Conformational Changes in Orotidine 5-Monophosphate Decarboxylase: A Structure-Based Explanation for How

the 5-Phosphate Group Activates the Enzyme. *Biochemistry 51*, 8665–8678, PMID: 23030629.

(21) Reyes, A. C., Amyes, T. L., and Richard, J. P. (2016) Enzyme Architecture: Self-Assembly of Enzyme and Substrate Pieces of Glycerol-3-Phosphate Dehydrogenase into a Robust Catalyst of Hydride Transfer. *Journal of the American Chemical Society 138*, 15251–15259, PMID: 27792325.

(22) Reyes, A. C., Koudelka, A. P., Amyes, T. L., and Richard, J. P. (2015) Enzyme Architecture: Optimization of Transition State Stabilization from a CationPhosphodianion Pair. *Journal of the American Chemical Society 137*, 5312–5315, PMID: 25884759.

(23) Kulkarni, Y. S., Liao, Q., Byléhn, F., Amyes, T. L., Richard, J. P., and Kamerlin, S. C. (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*, 3854–3857.

(24) Jin, Y., Richards, N. G., Waltho, J. P., and Blackburn, G. M. (2017) Metal Fluorides as Analogues for Studies on Phosphoryl Transfer Enzymes. *Angewandte Chemie - International Edition 56*, 4110–4128.

(25) Jin, Y., Molt, R. W., Waltho, J. P., Richards, N. G. J., and Blackburn, G. M. (2016) ^{19}F NMR and DFT Analysis Reveal Structural and Electronic Transition State Features for RhoA-Catalyzed GTP Hydrolysis. *Angewandte Chemie* 3379–3383.

(26) Blackburn, G. M., Cherfils, J., Moss, G. P., Richards, N. G. J., Waltho, J. P., Williams, N. H., and Wittinghofer, A. (2017) How to name atoms in phosphates, polyphosphates, their derivatives and mimics, and transition state analogues for enzyme-catalysed phosphoryl transfer reactions (IUPAC Recommendations 2016). *Pure and Applied Chemistry 89*, 653–675.

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

(28) Lahiri, S. D., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2003) The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction. *Science 299*, 2067–2071.

(29) Lahiri, S. D., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2002) Caught in the Act: The Structure of Phosphorylated β-Phosphoglucomutase from Lactococcus lactis,. *Biochemistry 41*, 8351–8359, PMID: 12081483.

(30) Go, M. K., Amyes, T. L., and Richard, J. P. (2010) Rescue of K12G Triosephosphate Isomerase by Ammonium Cations: The Reaction of an Enzyme in Pieces. *Journal of the American Chemical Society 132*, 13525–13532, PMID: 20822141.

(31) Zhai, X., Amyes, T. L., and Richard, J. P. (2014) Enzyme Architecture: Remarkably Similar Transition States for Triosephosphate Isomerase-Catalyzed Reactions of the Whole Substrate and the Substrate in Pieces. *Journal of the American Chemical Society 136*, 4145–4148, PMID: 24588650.

(32) Amyes, T. L., and Richard, J. P. (2013) Specificity in Transition State Binding: The Pauling Model Revisited. *Biochemistry 52*, 2021–2035, PMID: 23327224.

(33) Perrin, C. L., and Nielson, J. B. (1997) Strong hydrogen bonds in chemistry and biology. *Annual Review of Physical Chemistry 48*, 511–544, PMID: 9348662.

(34) 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*, 1–31.

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

(36) Kabsch, W. (2010) *XDS*. *Acta Crystallographica Section D 66*, 125–132.

(37) Vagin, A., and Teplyakov, A. (1997) *MOLREP*: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography 30*, 1022–1025.

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

(39) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D 53*, 240–255.

(40) Winn, M. D. et al. (2011) Overview of the *CCP*4 suite and current developments. *Acta Crystallographica Section D 67*, 235–242.

(41) 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 66*, 12–21.

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

(43) Hayward, S., and Berendsen, H. J. (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*, 144–154.

(44) Becke, A. D. (1993) Densityfunctional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics 98*, 5648–5652.

(45) Lee, C., Yang, W., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B 37*, 785–789.

(46) Vosko, S. H., Wilk, L., and Nusair, M. (1980) Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics 58*, 1200–1211.

(47) Stephens, P. J., Devlin, F. J., Chabalowski, C. F., and Frisch, M. J. (1994) Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *The Journal of Physical Chemistry 98*, 11623–11627.

(48) Kendall, R. A., Jr., T. H. D., and Harrison, R. J. (1992) Electron affinities of the firstrow atoms revisited. Systematic basis sets and wave functions. *The Journal of Chemical Physics 96*, 6796–6806.

(49) Li, X., and Frisch, M. J. (2006) Energy-Represented Direct Inversion in the Iterative Subspace within a Hybrid Geometry Optimization Method. *Journal of Chemical Theory and Computation 2*, 835–839, PMID: 26626690.

(50) Frisch, M. J. et al. Gaussian09 Revision E.01. 2009; Gaussian Inc. Wallingford CT.

Supporting Information

Included below are supporting images ...

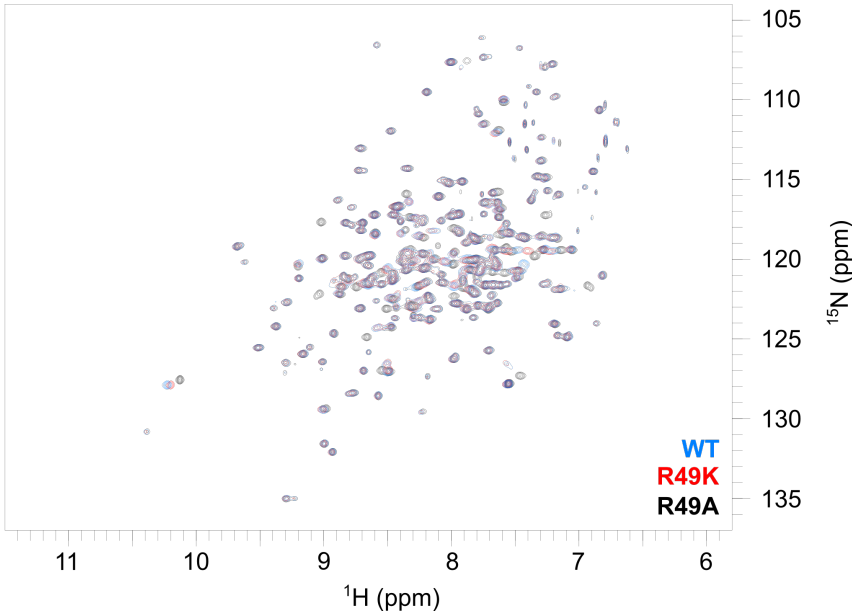


Figure S1: Overlay of the $^{1}H^{15}N$-TROSY spectra of βPGM_{WT} (blue), βPGM_{R49K} (red), βPGM_{R49A} (black) illustrating minimal perturbation to the overall fold chemical environment for the observed residues.

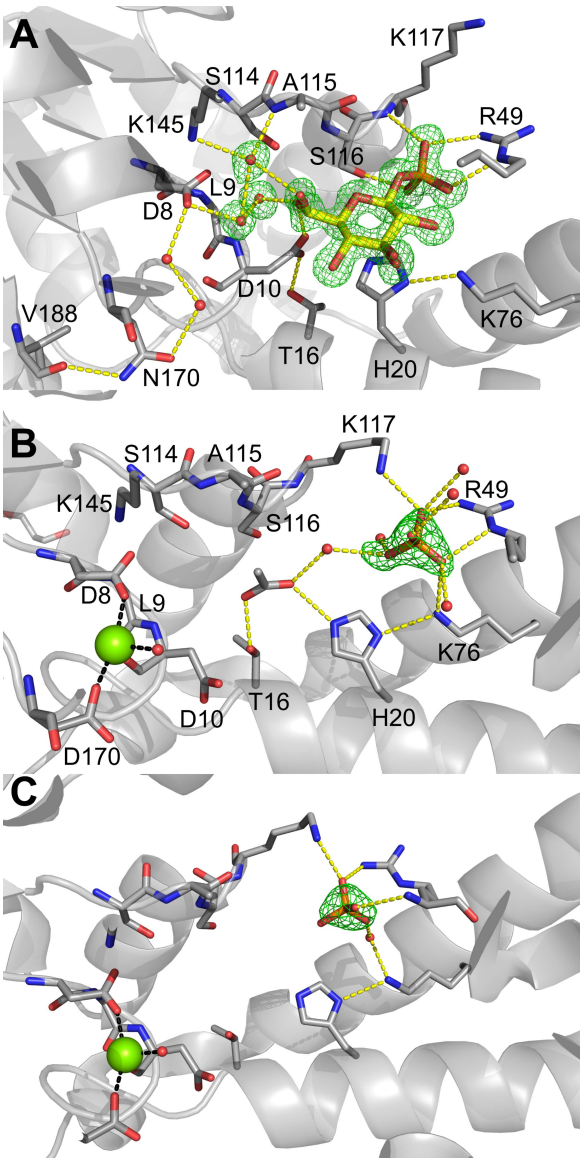


Figure S2: Omit maps for D170N:βG1P complex **(A)** and the βPGM_{WT}:Pi complex chain A **(B)** and B **(C)** contoured at 3 σ.

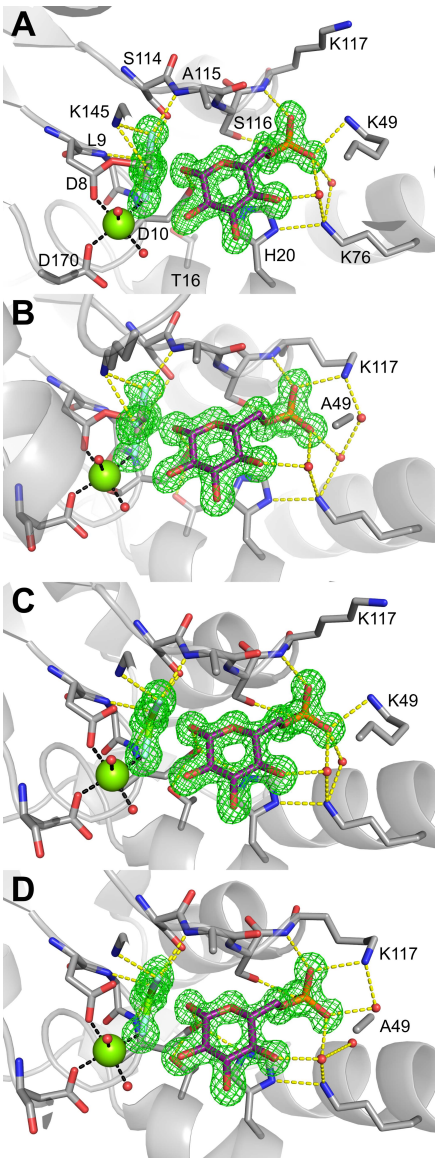
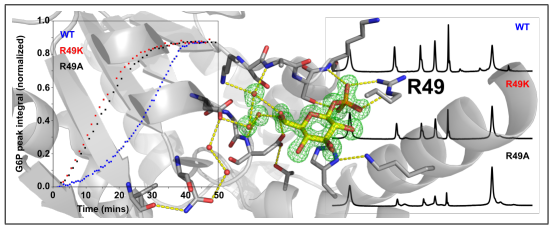


Figure S3: Omit maps for βPGM$_{\mathrm{R49K}}$:AlF$_4$:G6P **(A)**, βPGM$_{\mathrm{R49A}}$:AlF$_4$:G6P **(B)**, βPGM$_{\mathrm{R49K}}$:MgF$_3$:G6P **(C)**, βPGM$_{\mathrm{R49A}}$:MgF$_3$:G6P **(D)** complexes contoured at 3 σ.

Graphical TOC Entry



A.4 Paper IV: Mechanisms of phosphatase activity in good and bad phosphatases of the HAD superfamily

Contribution: I expressed and purified proteins with different isotope enrichment schemes. I performed the crystallography on βPGM with some technical assistance from CB. I performed the NMR experiments on βPGM and I analysed and interpreted the acquired data. I wrote the manuscript with early contributions from NJB and JPW.

Notably, much of the characterization of PSP was performed by Dr. Joanna Griffin (SIIS value determination of TSA complexes) Griffin, 2011, and Dr. Luke Johnson (mutagenesis and ^{19}F NMR) (Johnson, 2015). Furthermore, the crystallography of PSP with MgF_3 TSA complexes was performed by Dr. Matthew Bowler. This manuscript is currently a work in progress and where additional data are required, this is indicated by "XXXX", or notes to the reader in square brackets.

Mechanisms of phosphatase activity in good and bad phosphatases of the HAD superfamily

Angus J. Robertson[a], Joanna Griffin[a], Luke A. Johnson[a], Nikita Vekaria[b], Nicola J. Baxter[a,b], Andrea M. Hounslow[a], Claudine Bisson[c], Matthew J. Cliff[b], Matthew W. Bowler[d], Jonathan P. Waltho[a,b]

[a]*Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Firth Court, Western Bank, Sheffield, S10 2TN, United Kingdom*
[b]*Manchester Institute of Biotechnology and School of Chemistry, The University of Manchester, Manchester, M1 7DN, United Kingdom*
[c]*Department of Biological Sciences, ISMB, Birkbeck College, University of London, Malet Street, London, UK. WC1E 7HX*
[d]*European Molecular Biology Laboratory, Grenoble Outstation, 71 Avenue des Martyrs, CS 90181, F-38042 Grenoble, France*

Abstract

Phosphoryl transfer enzymes play a key role in biology, with vital roles in metabolism, cell signaling, and manipulation of genetic material. These enzymes can be broadly categorized into phosphatases, phosphotransferases (eg. kinases), and mutases. While in phosphatase enzymes the phospho-enzyme state is destabilized such that spontaneous autodephosphorylation is promoted, mutase enzymes need to stabilize a phospho-enzyme state in order to efficiently perform a ping-pong bi-bi reaction mechanism. To investigate how specific phosphatase vs. mutase activity has diverged, two enzymes from the well characterized haloacid dehalogenase (HAD) superfamily were selected. Here we show that a specific phosphatase (Phosphoserine phosphatase (PSP) from *Methanococcus jannaschii*) employs several mechanisms that promote phosphatase activity compared to a mutase (β-phosphoglucomutase (βPGM) from *Lactococcus lactis*), which actively employs mechanisms to prevent such activity. These themes can be roughly partitioned into three areas; translation of catalytic machinery, dislocation of solvent from the transferring phosphate group, and rotation of the phosphate group on a catalytically relevant timescale. All three themes act to ensure that βPGM acts as a mutase not a phosphatase, and that PSP acts

as a phosphatase, not a phosphotransferase and together present tools for the future design of enzymes in either class.

Keywords: Phosphatase, phosphomutase, Metal fluoride, Transition state analog, phospho-enzyme hydrolysis

1. Introduction

β-Phosphoglucomutase from *Lactococcus lactis* is a magnesium-dependent phosphoryl transfer enzyme (βPGM; EC. 5.4.2.6) in the haloacid dehalogenase superfamily (HADSF) which has been well-characterized physiologically [1–3], kinetically [4–7] and mechanistically [4, 6, 8–13]. βPGM catalyzes the reversible isomerization of β-glucose 1-phosphate (βG1P) to glucose 6-phosphate (G6P) via a β-glucose 1,6-bisphosphate (βG16BP) intermediate using a ping-pong bi-bi reaction mechanism (Fig. 1, S6). The active site of βPGM is located at the interface between the helical cap (T16-V87) domain and the α/β core domain (M1-D15, S88-K216). Cap opening and closing relative to the core domain occurs during the catalytic cycle [6], which exposes the active site to solvent and facilitates release of the substrates and the βG16BP intermediate [5]. There are two phosphate binding sites (proximal and distal to the catalytic Mg^{2+} ion) which allow βPGM^{P} to bind either βG1P or G6P as substrates, and βPGM to bind the βG16BP intermediate in either orientation, thus facilitating mutase activity (Fig. 1). βPGM uses a general acid base (GAB; residue D10) to both align and activate substrate for phosphoryl group transfer, and the engagement of the GAB has been observed following cap domain closure using both metal fluoride transition state analogue complexes [9, 11], as well as the native βG16BP intermediate [13]. βPGM has a secondary activity as a phosphatase which is *ca.* 1000x slower than the mutase activity [7, 9, 13].

Phosphoserine phosphatase from *Methanococcus jannaschii* (PSP; EC 3.1.3.3) is another HADSF member that has been well characterized [14, 15] and is similar in structure to βPGM but it is solely a phosphatase. The active site of PSP is located at the interface between the helical cap (N18-T76) domain and the α/β core domain (M1-V17, P77-K211). The catalytic DXD motif of the HADSF is present in both PSP and βPGM, and substrate selectivity is primarily achieved through differences in the cap domain, as is typical of HADSF members [4, 9, 16–19]. The reaction cycles in both PSP and βPGM necessitate a phospho-enzyme state (phosphorylated aspartate

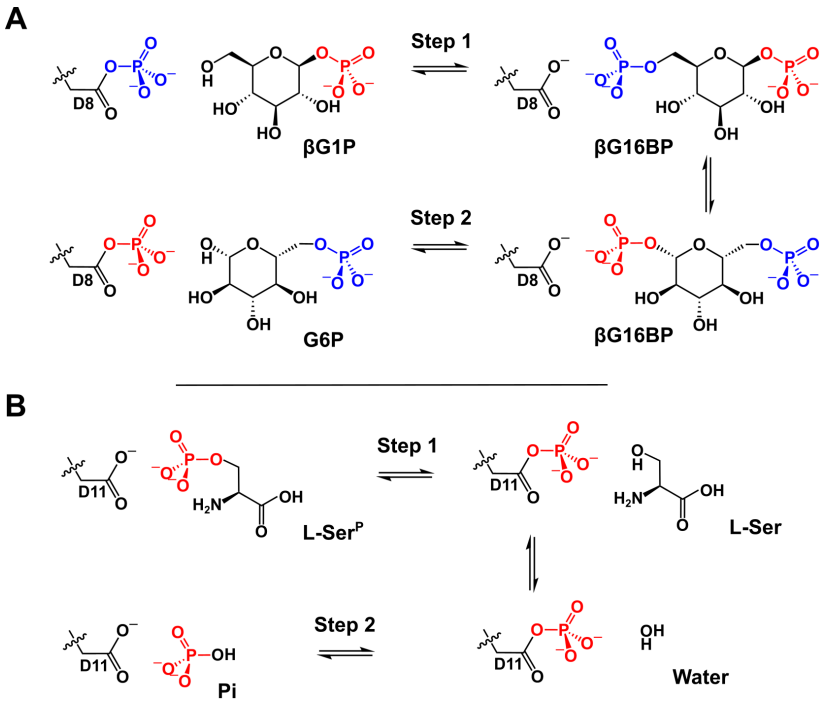


Figure 1: Reaction scheme for **(A)** the enzymatic conversion of β-glucose 1-phosphate (βG1P) to glucose 6-phosphate (G6P) via a β-glucose 1,6-bisphosphate (βG16BP) intermediate by βPGM and **(B)** the dephosphorylation L-serine-O-phosphate (L-Ser$^{\mathrm{P}}$) to L-serine (L-Ser) and subsequent hydrolysis of the aspartyl phosphate group catalyzed by PSP. **A** illustrates the phosphoryl transfer reaction between the phospho-enzyme (βPGM$^{\mathrm{P}}$, phosphorylated at residue D8) and βG1P (termed Step 1) with the transferring phosphate illustrated in blue. The equivalent reaction between βPGM$^{\mathrm{P}}$ and G6P is termed Step 2 with the transferring phosphate illustrated in red. **B** illustrates the phosphoryl transfer reaction between L-Ser$^{\mathrm{P}}$ and the enzyme (to generate phospho-enzyme PSP$^{\mathrm{P}}$, phosphorylated at residue D11) with L-Ser as product (termed Step 1). The hydrolysis of PSP$^{\mathrm{P}}$ is termed Step 2 and generates inorganic phosphate.

residue D11 and D8 respectively). Previous work indicates that phospho-enzyme hydrolysis in βPGM is GAB independent [13] which contrasts with structural investigations of PSP where the GAB is positioned to activate a water molecule for attack on the phospho-enzyme analogue [15].

Structural investigations of species along the reaction coordinate in both βPGM and PSP have made extensive use of metallofluoride moieties to trap both ground state analogue (GSA) and transition state analogue (TSA) complexes [20, 21]. The stable GSA complex BeF_3^- in both βPGM and PSP is a close mimic of phospho-enzyme, where the Be atom forms a covalent bond with the carboxylate Oδ1 atom of residue D8 in place of the P atom, and the three F atoms substitute for the three non-bridging O atoms (Fig. 2) [11, 15]. Typical inorganic Be–F bonds (1.5 - 1.6 Å) in BeF_3^- are of a similar length to P–O bonds in phosphates, producing surrogates with comparable geometry, near obligate tetrahedral organization and the same net charge [11, 15, 22]. Metal fluorides have also been used to mimic the transferring phosphate in phospho-enzyme hydrolysis reactions in PSP [15], as both MgF_3^- and AlF_4^- assemble on the catalytic aspartate residue with a similar geometry and charge to the native phosphoryl group.

It has been argued previously that βPGM is such a poor phosphatase relative to PSP because D10 is in the *out* position (rotated away from the catalytic aspartate (D8)) when there is no substrate engaged and moves to the *in* position (χ_1 angle rotated *ca.* 180° towards D8) on domain closure. However, a comparison of the BeF_3^- complexes (PDB: 2WF9 [11]) and the βG16BP complex (PDB: 5OK0, 5OK1 [13]) indicates the *in* – *out* transition is not necessarily coupled to domain closure until the enzyme adopts the near-transition-state conformation. This raises the question as to why D10 is not recruited in the open form of βPGM given the analogous DXD HADSF motif in both βPGM and PSP.

Here we utilize metal fluoride GSA and TSA complexes in both PSP and βPGM systems to examine differences in the behaviour of the GAB in the two enzymes during phospho-enzyme hydrolysis. The results show that there are three components that act to ensure that βPGM acts as a mutase not a phosphatase. The components are the translation of key parts of the catalytic machinery away from the transferring phosphoryl group, the disruption of solvation of the transferring phosphoryl group, and the destabilization of phosphoryl group binding as evidenced by its rotation on the ms timescale. We also show that the water molecules in the active site of the phosphatase are constrained into the positions occupied by polar groups of

the substrate during the phospho-enzyme hydrolysis step. Finally, we establish that the native reaction is associated with the freezing out of motions across the enzyme which is not apparent when phosphatase reactivity is a side reaction.

2. Results

2.1. Investigation of phospho-enzyme hydrolysis in βPGM

In order to compare more closely the properties of PSP and βPGM during their phosphatase reactions, various complexes of βPGM_{WT} with metal fluorides were crystallized to investigate the TS protein architecture associated with βPGM^{P} hydrolysis. The structure of the βPGM_{WT} MgF_3 complex was determined to a resolution of 1.8 Å (**PDB: 6H8X**; Table S1; Fig. S7), and the βPGM_{WT} AlF_4 complex to a resolution of 2.0 Å (**PDB: 6H8W**; Table S1; Fig. 3). In both structures, the cap and core domains are in an open arrangement that is broadly similar to the βPGM_{WT}:BeF_3 complex (PDB: 2WHE, [11]) and the substrate-free βPGM_{WT} structures (PDB: 1ZOL [5], 2WHE [10]) reported previously (Table S2, S3).

In the βPGM:MgF_3^- complex, fluoride F1 coordinates the catalytic magnesium, F2 hydrogen bonds with the backbone amide of D10 and to S114 Oγ2, and F3 hydrogen bonds with the backbone NH group of A115 and the side chain amine group of K145 (Fig. S7). In contrast, the hydrogen bond from the backbone amide of L9 to F2 is lost compared to the βPGM:MgF_3:G6P structure (PDB: 2WF5, [8]), and, surprisingly, there is no observable Fo-Fc density (at 3 σ) or significant 2Fo-Fc density (at 1.5 σ) to indicate the presence of a nucleophilic water attacking the MgF_3^- group. In the AlF_4 complex (Fig. 3) a square planar AlF_4^- moiety occupies the *proximal* site of the enzyme with almost identical hydrogen bonds to those made by the AlF_4^- TSA in the βPGM:AlF_4: G6P structure (PDB: 2WF6 [10]). A water molecule (W1) acts as a sixth ligand to the Al^{3+} atom with an additional water molecule (W2) hydrogen bonding to W1.

In both the βPGM:MgF_3 and βPGM:AlF_4 structures, the general acid-base (GAB; D10) is (primarily) rotated away from the catalytic aspartate residue (D8) termed the *out* position, compared to the *in* position which aligns substrate hydroxyl group for nucleophilic attack. This indicates that the GAB does not readily engage with an attacking water molecule (W1) in the active site of βPGM^{P} hydrolysis analogues. However, in the the βPGM_{WT}:AlF_4 structure, there is evidence in the difference Fourier map

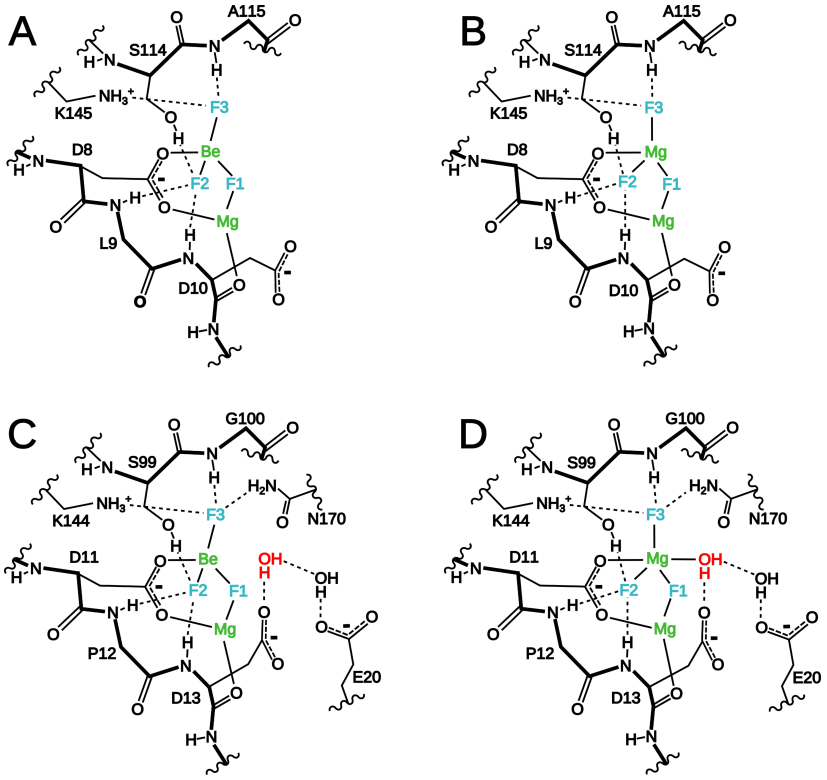


Figure 2: Schematics of the active site of βPGM and PSP complexes with metal fluorides from deposited crystal structures. **A)** βPGM_{WT}:BeF_3^-(PDB: 2WFA), **B)** βPGM_{WT} :MgF_3^- (PDB: 6H8X), **C)** PSP_{WT}:BeF_3^- (PDB: 1J97), **D)** PSP_{WT}:MgF_3^- (PDB: 1L7N). Backbone covalent bonds are drawn with a thick line, sidechains (and metal coordination bonds) are drawn with a thin line, and hydrogen bonds are drawn with a dashed line. The metal fluorides, catalytic Mg^{2+}, and nucleophilic water are drawn with color for clarity.

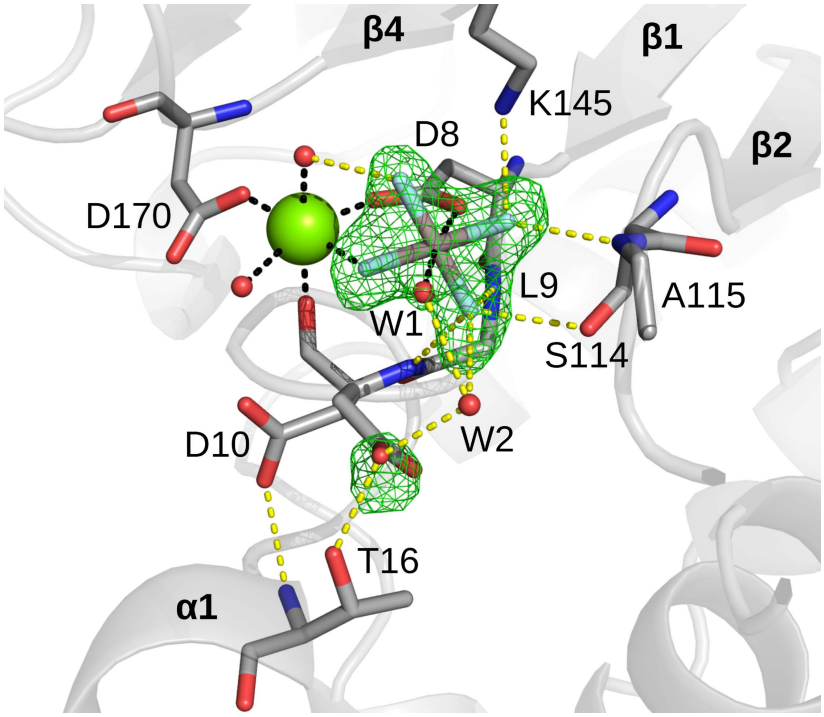


Figure 3: Active site of βPGM$_{WT}$:AlF$_4$ complex with Fo-Fc difference density contoured at 3σ. Standard CPK colors are used for atoms. Metal ion coordination is illustrated using black dashes, hydrogen bonds are drawn using yellow dashes, and structural waters are drawn as red spheres. A structural overview of the active site of βPGM is presented in Fig. S6

immediately adjacent to D10 that the sidechain populates the *in* rotamer (occupancy of *ca.* 0.2). Here, the sidechain carboxylate oxygen hydrogen bonds to W2 (and the sidechain hydroxyl group of T16) rather than aligning W1 for attack on the phosphate surrogate (AlF_4^-). This indicates that the GAB in βPGM is both rotated and translated away from the active site, where rotation from *out* to *in* in the open state is insufficient to align water for nucleophilic attack on the phosphoryl group.

To investigate whether the stabilization of the *out* D10 rotamer over the *in* rotamer affected the hydrolysis rate of phospho-enzyme (as was suggested previously [9, 11, 13]), the T16A variant of βPGM (βPGM$_{T16A}$) was

generated. The transition state of phospho-enzyme hydrolysis was investigated using MgF_3^- and AlF_4^- TSAs (**PDB: 6H8Z, 6H8Y** respectively; Table S1). The resulting structures were highly similar to the wild-type structures (Table S2, S3) with similarly poor water definition around the transferring phosphate analogs. In solution, the resting state for βPGM_{T16A} in the presence of phosphorylating agent (acetylphosphate; AcP) was determined to be phosphorylated following the incubation of the enzyme with 50 mM AcP and observation by $^1H^{15}N$-BEST-TROSY NMR (Fig. S8). Hydrolysis of the phospho-βPGM_{T16A} enzyme was followed using 1D ^{31}P NMR and the rate of hydrolysis was determined to be 0.05 ± 0.003 s^{-1}, which is in close agreement with previously reported rates for βPGM_{WT} (0.06 ± 0.006 s^{-1}) and βPGM_{D10N} (0.02 ± 0.002) using this method [13], and also by other groups [7]. Thus, despite the stabilization of the *out* rotamer over the *in* rotamer in βPGM_{T16A}, there appears to be no significant change in hydrolysis rate on removal of the T16 hydroxyl group. It was previously predicted that phospho-enzyme hydrolysis in βPGM_{WT} was independent of the GAB as mutation to asparagine did not change the βPGM^P hydrolysis rate [13]. Furthermore, phosphorylation of substrate-free βPGM by AcP within pre-formed crystals demonstrated no GAB involvement (see SI section 3 for details). Taken together these observations demonstrate that βPGM^P hydrolysis is GAB independent and point towards a translation of the GAB away from the transferring phosphoryl group as a mechanism to prevent alignment of water for nucleophilic attack.

2.2. Structural investigation of phosphoryl group transfer in PSP

Previously both AlF_4^- and MgF_3^- TSA complexes of the hydrolysis of the PSP phospho-enzyme (PSP^P ; phosphorylated at residue D11) were observed when both Mg^{2+}, Al^{3+}, and fluoride were included in the crystallization condition [14, 15]. The previously reported PSP^P hydrolysis TSA complex was a mix of octahedral and trigonal bipyramidal coordination. In order to simplify this complex, PSP was crystallized in the presence of XXXX mM $MgCl_2$ and XXXX mM NaF, and the structure was determined to XXXX Å resolution (**PDB: XXXX**, Table XXXX; Fig. 4). This confirms that the trigonal species observed previously was MgF_3^- as opposed to AlF_3, and displays an almost identical active site in terms of geometry and water positioning compared to the previously reported structure.

The phosphatase activity of PSP was investigated using MgF_3^- TSA, and the structure of the PSP:MgF_3:L-Ser complex was determined to XXXX Å

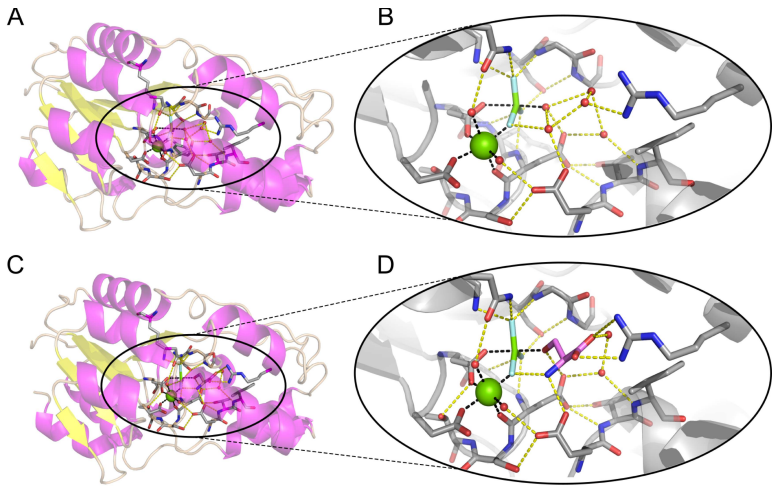


Figure 4: The structure of PSP complexed with (A, B) MgF_3, or with (C, D) MgF_3:L-Ser. A ribbon representation of PSP is used (in A and C) showing the core (left) and cap (right) domains with α-helices colored purple, β-sheets yellow, using standard standard CPK colors for atoms. The L-ser is colored with purple carbon atoms in (D) for clarity. Metal ion coordination is illustrated using black dashes, hydrogen bonds are drawn using yellow dashes, and structural waters are drawn as red spheres.

resolution (**PDB: XXXX**, Table S1; Fig. 4). In this structure, the cap and core domains XXXX overlay with the PSP^P hydrolysis analogue indicating minimal domain closure when the L-Ser substrate is bound compared to water. The L-Ser substrate occupies the active site with MgF_3^- mimicking the phosphate group being transferred from the side-chain hydroxyl of L-Ser to a sidechain carboxylate oxygen of residue D11 (Fig. 4). The positioning of the L-Ser polar groups closely reflects the positioning of water molecules in the binding pocket which presents a potential mechanism for bi-specificity of this enzyme. Namely, that both L-Ser and water molecules are both bound and specifically oriented in the active site prior to catalysis of either phosphatase or hydrolysis. This suggests that water molecules are accommodated as-substrate in PSP, with specific geometries and orientations.

2.3. Phospho-enzyme hydrolysis in PSP probed by ^{19}F NMR

To probe this relationship between the accommodation of polar groups (either of L-Ser, or of water molecules) in the active site of PSP, 1D ^{19}F NMR spectra were recorded of the PSP complexed with metal fluoride complexes (MF_x) BeF_3^-, MgF_3^-, and AlF_4^-, both with, and without the L-Ser ligand (Table. S7). Here phospho-enzyme (PSP:BeF_3) and phosphoryl group transfer (PSP:MgF_3 and PSP:AlF_4) analogues report on the chemical environment surrounding the phosphate analogue [23, 24]. A significant chemical shift change (14.8 ppm) of the F1 resonance was observed upon addition L-Ser to the PSP:MgF_3 TSA complex (Fig. S13). This downfield shift correlates with the crystallographic observation where the hydrogen bond donor to F1 in the PSP:MgF_3 complex (H_2O, 3.2 Å) is replaced by the amide group of L-Ser at 2.8 Å. However, the ^{19}F NMR spectra of both PSP:MgF_3:L-Ser and PSP:AlF_4:L-Ser complexes are only partly saturated with the L-Ser substrate.

To investigate the binding affinity of L-Ser to PSP:MF_x complexes, L-Ser was titrated into the PSP:AlF_4 TSA complex, as the AlF_4^- TSA resulted in the highest affinity TSA complex in βPGM [12]. The K_d value for the binding of L-Ser to PSP:AlF_4 TSA complex was determined to be 13.2 ± 2.1 mM (Fig. S16). The PSP:MgF_3:L-Ser complex was *ca.* 80% saturated at 10 mM L-Ser, which indicates a comparable K_d to the PSP:AlF_4:L-Ser complex, and a general low affinity for the L-Ser to MF_x complexes in PSP. Upon addition of 5 mM L-Ser to a PSP:BeF_3 complex (Fig. S19), no change in ^{19}F NMR spectrum was observed suggesting that a stable product complex

is not populated. This is readily rationalizable as the enzyme must dissociate L-Ser prior to the phospho-enzyme hydrolysis reaction.

Despite the poor affinity for the PSP:MF_x TSA complexes, narrow linewidths were observed in all of the complexes indicating that the MF_x moiety is stably coordinated in the active site (Table 6). In order to validate this prediction, and corroborate the crystallographic observation that polar groups of L-Ser closely reflect water positions in the absence of substrate, SIIS measurements [25] of the MgF_3 and AlF_4 complexes were performed (SI section 5). A subtle increase in SIIS value was observed on addition of L-Ser to PSP:MgF_3 TSA complex (Fig. S14, S15), while a more significant increase was observed on addition of L-Ser to the PSP:AlF_4 TSA complex (Fig. S17, S18). Across all of the complexes F1 displays only a small SIIS which indicates that that fluoride is strongly coordinated by the Mg^{2+} ion, while water molecules primarily hydrogen bond to F2, F3, and F4 (when applicable) as is suggested in the crystal structure (Table S8).

To investigate the effect of destabilizing water coordination on the phospho-enzyme hydrolysis reaction in PSP, the E20A variant of PSP (PSP_{E20A}) was generated. Sidechain atom Oγ1 of E20 hydrogen bonds to a non-nucleophilic water in the active site, which hydrogen bonds to both the nucleophilic water molecule and the fluoride in the F1 position in PDB: 1L7N [15] (Fig. 2). 1D-^{19}F NMR spectra were recorded of 1 mM PSP_{E20A} in standard PSP NMR buffer with the addition of 20 mM NaF to form the PSP_{E20A}:MgF_3^- complex (Fig. 5). Remarkably similar lineshapes of MgF_3^-:H_2O TSAs are observed for wild-type PSP and PSP_{E20A} variants, however a *ca.* 30 Hz linebroadening of the F1 peak is observed in PSP_{E20A} (Table 6). This linebroadening correlates with a subtle reduction in the stabilization of the F1 position, but apparently no consequence on the overall rotation of the MgF_3^- moiety in the active site of PSP.

A principal difference between the active sites of βPGM and PSP in the inclusion of an addition hydrogen bonding partner to the F3 position in PSP (Fig. 2). To investigate the contribution of this extra hydrogen bond to overall rotation of the transferring phosphate mimic in the active site of PSP, the N170A variant (PSP_{N170A}) was generated. 1D-^{19}F NMR spectra were recorded of 1 mM PSP_{E20A} in standard PSP NMR buffer with the addition of 20 mM NaF to form the PSP_{N170A}:MgF_3^- complex (Fig. 5). Lineshapes of the PSP_{N170A}:MgF_3^-:H_2O TSA reflect those of wild-type PSP for F1 and F2 positions, however a *ca.* 200 Hz linebroadening of the F3 peak is observed for the complex with PSP_{E20A} (Table 6). Notably, it is only the

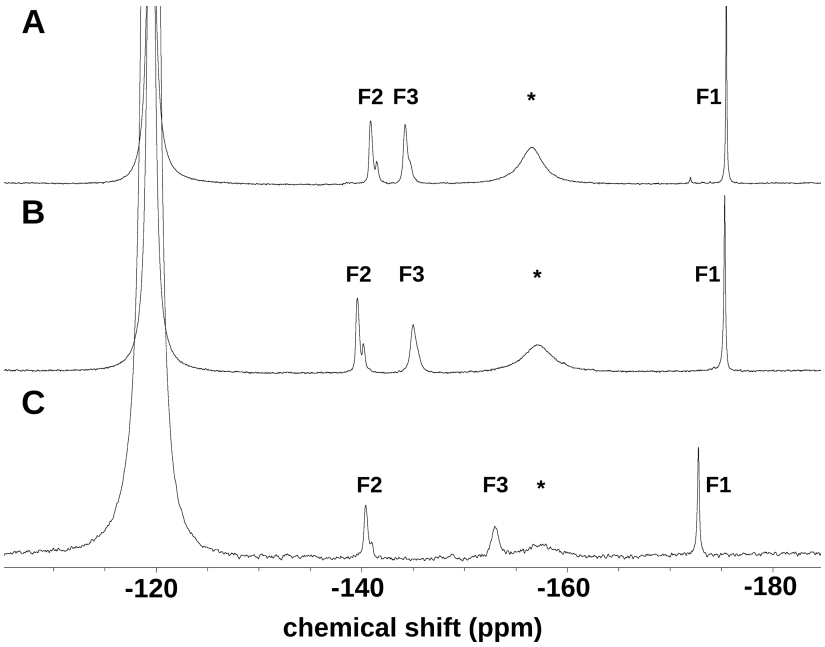


Figure 5: ^{19}F 1D NMR spectra of **A)** WT, **B)** PSP_{E20A}, and **C)** PSP_{N170A}, complexed with MgF_3^- and water to form a phospho-enzyme hydrolysis analog. Fluorine assignments are labeled in black according to Fig. 2. The peak denoted with an asterisk corresponds to MFx species free is solution (see ref. [10, 11]), while the leftmost peak corresponds to free fluoride. The chemical shifts of the fluorides are **A)** -175.5, -140.9, -144.2, **B)** -175.3, -139.6, -145.0, and **C)** -172.8, -140.4, -153.0, for fluorides F1, F2 and F3 respectively in each of the complexes.

F3 resonance that is linebroadened, indicating a local perturbation that is not propagated to other fluorides in the MF_x TSA complex which also shows that the fluorine resonances are not exchanging with each other. Together this indicates that the water network in the active site of PSP is relatively robust, which further implies that the active site has evolved to select for water *as-substrate* in addition to phospho-L-Ser.

2.4. Phospho-enzyme hydrolysis in βPGM probed by ^{19}F NMR

To investigate the chemical environment surrounding phospho-enzyme hydrolysis in βPGM, analogous phospho-enzyme (βPGM:BeF_3) and phosphoryl group transfer (βPGM:$MgF_3$and βPGM:AlF_4) complexes were pre-

pared to those in PSP, and 1D ^{19}F NMR spectra were recorded. Unlike the ^{19}F NMR spectrum of the PSP:MgF_3 complex, the NMR spectrum of the βPGM:MgF_3 complex resulted in only two resolvable (protein associated) peaks (Fig. 6). The sharpest peak at -173.4 ppm corresponds to a fluoride occupying the F1 position, while the second peak appears much broader at -147.0 ppm and indicates that F2 and F3 are likely averaged to a single peak. This poor stabilization of the MgF_3 moiety in the active site of βPGM may be in response to the weak Lewis acidity of the Mg^{2+} ion compared to Al^{3+}, or the fact that MgF_x can exist in both octahedral and trigonal bipyramidal geometries.

The 1D ^{19}F NMR spectrum of the βPGM:AlF_4 complex displays a reduced linebroadening effect and the chemical shifts of the four ^{19}F resonances resemble those of AlF_4 transition state analogue complexes with βG1P and G6P substrates [10, 12], but shifted slightly upfield (Fig. 6). ^{19}F 1D NMR spectra of βPGM_{WT} complexed with BeF_3^- show three protein bound peaks with a narrower linewidth than the MgF_3^- and AlF_4^- TSAs in βPGM but with a much broader linewidth than the corresponding complex in PSP (Fig. 6). The upfield shift and the broad linewidth of the four AlF_4^- resonances in the βPGM:AlF_4 complex indicates that the phospho-enzyme hydrolysis analogue is not coordinated as stably as either the ground state analogue complex (βPGM:BeF_3) or the βPGMMgF_3:G6P and βPGM:AlF_4:G6P transition state analogue complexes, which is a markedly different behaviour than the corresponding complexes in PSP. In βPGM it appears that water *as-substrate* is insufficient to form a stable complex in either GSA or TSA complexes and the ^{19}F NMR indicates that an exchange process may be present in βPGM (that causes an increased linewidth) that is not present in PSP (Fig. 6; Table S24).

2.5. Investigation of chemical exchange processes affecting βPGM but not PSP

In order to investigate the source of the linebroadening, ^{19}F-^{19}F EXSY NMR spectroscopy was used to determine the nature of any chemical exchange processes present. Namely, if the linebroadening in the ^{19}F spectra was due to rotation of the MF_x moiety in the active site, or to dissociation from the active site and exchange with fluoride containing species in solution, which may itself mimic the hydrolysis process. To address this question, ^{19}F-^{19}F EXSY spectra were recorded of the βPGM:AlF_4 complex with mixing delays of 50μs to 50 ms. Exchange peaks were observed on similar timescales

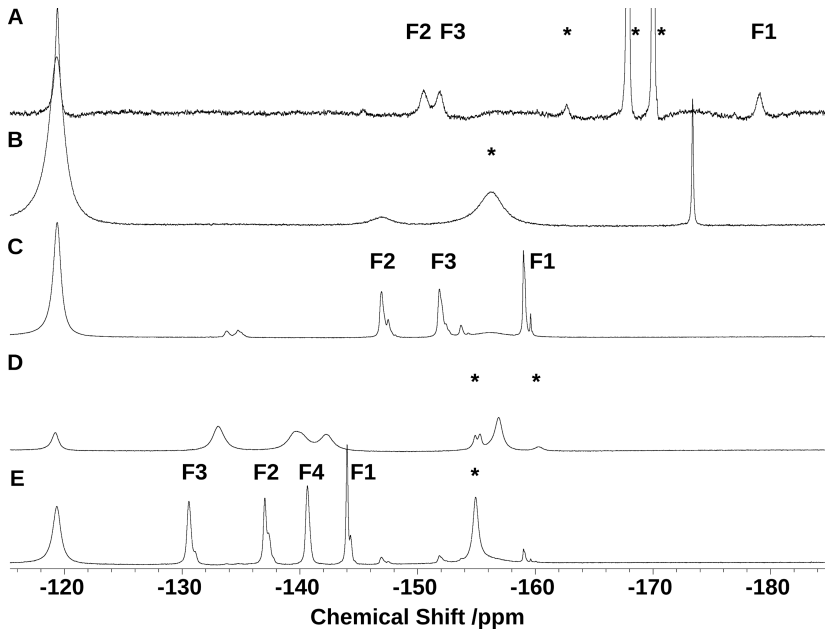


Figure 6: ^{19}F 1D NMR spectra of βPGM complexed with metal fluoride species, fluorine atoms (where known) are labeled according to Fig. **??** . **A)** βPGM:BeF_3 complex. **B)** βPGM:MgF_3, **C)** βPGM:MgF_3:G6P, **D)** βPGM:AlF_4, **E)** βPGM:AlF_4:G6P. Peaks denoted with an asterisk correspond to MFx species free is solution (see ref. [10, 11]), while the leftmost peak corresponds to free fluoride. In **D**, peaks at -154.9, -155.3, and -160.3 ppm correspond to AlFx species in solution. The chemical shifts of the metal fluoride resonances (in ppm) are: **A)** -178.7, -150.3, -151.7, **B)** -173.4, -147.0, **C)** -159.0, -147.0, -151.9, **D)** -156.9, -139.7, -133.1, -142.3, **E)** -144.0, -137.0, -130.6, -140.7, for fluorides F1, F2, F3, and F4 respectively (where applicable).

for all AlF_4 fluorides, with k_{ex} terms of 1000 s^{-1} fitted for the four AlF_4 fluorides (Fig. S23). Exchange with free fluoride is observed for one peak in the AlF_4 species on a timescale faster than the shortest mixing time of 50 μs. This peak is assigned as F1 due to the similarity in chemical shift to previously reported AlF_4 fluorides coordinated by the catalytic Mg^{2+} ion. No exchange peaks were observed between the AlF_4^- species and free AlF_x species in solution indicating that the whole AlF_4^- moiety does not dissociate from the active site on the timescale observed, instead the exchange with free fluoride is likely mediated by fluorine coordination of the catalytic Mg^{2+} and rotation into the F1 position.

To investigate any chemical exchange process present in the βPGM:BeF_3 complex ^{19}F-^{19}F EXSY NMR spectra were again recorded and exchange between each of the three protein-bound fluorine positions was observed with a k_{ex} of *ca.* 1000 s^{-1} (Fig. S22). However, exchange peaks were not observed between the protein bound BeF_3 group and either the free fluoride peak, or other BeF_x groups in solution, which demonstrates that the βPGM:BeF_3^- complex has a lifetime in excess of 50 ms, analogous to the phospho-enzyme. This demonstrates that in both βPGM:AlF_4 and βPGM:BeF_3 complexes a comparable rotational exchange process is present about the bond between the sidechain Oδ1 atom of D8 (O$\delta1_{D8}$) and the metal. In the βPGM:BeF_3 complex, this process is much faster than both the rate of phospho-enzyme hydrolysis (*ca.* 0.03-0.06 s^{-1}; [7, 9, 13]) and of catalysis (*ca.* 70 s^{-1} [7, 13]). The observation that this 1000 s^{-1} process is present in both BeF_3^- and AlF_4^- analogues suggests that there may be an underlying protein conformational cause, particularly as much of the active site is in intermediate exchange in the substrate-free form of βPGM [10].

To probe the active site dependence of the BeF_3^- group rotation, the K145A variant of βPGM (βPGM_{K145A}) was generated. The side chain amine of K145 directly coordinates F2 of the BeF_3^- group alongside the backbone NH group of A115 (Fig. 2). 1D ^{19}F NMR spectra of the βPGM_{K145A}:BeF_3 complex demonstrated an increased linewidth for all protein bound BeF_3^- peaks compared to the βPGM_{WT}:BeF_3 complex at 25 °C (Fig. S20, Table 6). This indicates that hydrogen bonding from the side chain amine of K145 restricts the rotation of the BeF_3^- group about the O$\delta1_{D8}$–Be^{3+} bond and consequently the exchange process observed. A non-linear and non-uniform linebroadening response to temperature was observed in the βPGM_{WT} and βPGM_{K145A} complexes with BeF_3^- (Fig. S24) which prohibited accurate activation energy calculations for rotation. However, the F1 fluoride displayed

a relatively linear response (within error) to temperature (Fig. S25) and given its proximity to the catalytic Mg^{2+} ion, it is tempting to speculate that dissociation of the Mg^{2+} ion may play a role in this exchange process. The observed effect on BeF_3^- group rotation is not as pronounced as could be expected given the loss of an ionic interaction, which suggests either that the interaction between the K145 sidechain amine and the BeF_3^- moiety is only a small component of the activation energy barrier for rotation, or that a cation from solution (eg. K^+) can substitute for the amine group.

3. Discussion

In this work, three central themes (translation, dislocation, and rotation) have been described that distinguish the specific phosphatase activity of PSP from the specific mutase activity of βPGM in a key enzyme superfamily [19]. Given the previous crystal structures of PSP using the GAB residue (D13) to align water for nucleophilic attack on the phospho-enzyme [15], there was little debate of the direct role that the GAB played. In βPGM where the same catalytic DXD motif could utilize the GAB residue (D10) for the same purpose, several crystal structures presented here indicate that it does not as was predicted previously [11, 13]. In these structures residue D10 not only adopts a rotamer that is both rotated *out* of the active site, but the residue is also translated away such that rotation from *out* to *in* is insufficient to align a water molecule for nucleophilic attack. In a structure where a partial *in* occupancy is observed, the GAB indirectly coordinates a nucleophilic water molecule (via a second water molecule), which may form the basis of a proton transfer network between nucleophilic water molecule to the GAB [18]. However, the minimal perturbation of the dephosphorylation rate when the GAB residue is mutated to a constitutively protonated mimic eliminates this possibility [13]. The comparison in PSP is that in both the PSP:BeF_3 and PSP:MgF_3 structures indicate that the GAB residue (D13) occupies an *in* rotamer that aligns water for nucleophilic attack on the phosphate group.

Key active site differences exist between PSP and βPGM beyond the GAB *in* – *out* transition. In PSP, the positioning of polar groups in the active site is near identical when either L-Ser or water are accommodated. This presents a mechanism whereby the active site in PSP acts to specifically orient a shell of water molecules around an activated water nucleophile in the same manner that a ligand is usually coordinated in the active site of an enzyme. This water *as-substrate* model is relatively robust, with no observed rotation of the

phosphate group mimic in the active site to cause a more global perturbation. By contrast, open-βPGM structures with either phosphate (or transferring phosphate) surrogates are coordinated by poorly defined water molecules either in solution, or crystallographically. This indicates that in βPGM, one of the protection mechanisms of the high energy phosphate group is to expose it to unstructured, bulk solvent, rather than preclude it from solvent.

Chemical exchange of the phosphate surrogate (BeF_3^-) is observed in βPGM but not in PSP, even when the number of coordinating groups is equalized between the two enzymes. This chemical exchange is the result of rotation of the BeF_3^- moiety around the apartyl Oδ1 – Be^{3+} bond and suggests a tightly controlled position of the BeF_3^- group in PSP which is not present in βPGM. The observation that removal of a coordinating positive charge (sidechain amine of K145) had only a moderate effect on this exchange process suggests that there are larger contributors to this exchange process present. One possibility is that there is an underlying conformational dynamic in the active site of βPGM that manifests as a rotation of the BeF_3^- moiety. A second possibility is that a more charged species is dominating the electrostatic environment surrounding the BeF_3^- group. Given the proximity of the $Mg^{2+}{}_{cat}$ ion coupled to the reportedly poor affinity, both catalytically [7] and structurally [13], it is tempting to speculate that dissociation of the $Mg^{2+}{}_{cat}$ ion may be responsible.

It has been asserted that as enzymes evolve towards a specific function, they rigidify (although no specific timescale was given, and a catalytically relevant one is assumed) [26–29]. By both X-ray crystallography and solution NMR, it is observed that PSP coordinates both L-Ser and water stably in the active site, with well defined and robust water network. Contrastingly, the active site of βPGM exposes the transferring phosphate group to unstructured solvent which is typically more plastic in nature than protein residues. Herein lies a potential distinction between the two enzymes, PSP binds water *as-substrate* in a stable manner conducive to specific phosphatase activity, whereas βPGM employs several conformational and geometric measures to prevent this from happening.

4. Acknowledgements

The authors would like to thank Sam Dix, Alicia Churchill-Angus, and Adli Aziz for technical support and helpful discussions. These studies were supported, in part, by BBSRC (Grants BB/E017541, BB/K016245 and BB/M021637

to J.P.W.). AJR was funded by a University of Sheffield department studentships and the authors would also like to thank the Universities of Sheffield and Manchester for support.

5. References

[1] N. Qian, G. Stanley, B. Hahn-Hagerdal, P. Rådström, Purification and characterization of two phosphoglucomutases from *Lactococcus lactis* subsp. lactis and their regulation in maltose- and glucose-utilizing cells, Journal of Bacteriology 176 (1994) 5304–5311.

[2] N. Qian, G. Stanley, A. Bunte, P. Rådström, Product formation and phosphoglucomutase activities in *Lactococcus lactis*: cloning and characterization of a novel phosphoglucomutase gene, Microbiology 143 (1997) 855–865.

[3] Levander, F and Andersson, U and Rådström, P , Physiological role of β-phosphoglucomutase in *Lactococcus lactis.*, Applied and environmental microbiology 67 (2001) 4546–4553.

[4] S. D. Lahiri, G. Zhang, J. Dai, D. Dunaway-Mariano, K. N. Allen, Analysis of the Substrate Specificity Loop of the HAD Superfamily Cap Domain, Biochemistry 43 (2004) 2812–2820.

[5] G. Zhang, J. Dai, L. Wang, D. Dunaway-Mariano, L. W. Tremblay, K. N. Allen, Catalytic cycling in β-phosphoglucomutase: A kinetic and structural analysis, Biochemistry 44 (2005) 9404–9416.

[6] J. Dai, L. Wang, K. N. Allen, P. Rådström, D. Dunaway-Mariano, Conformational cycling in β-phosphoglucomutase catalysis: Reorientation of the β-D-glucose 1,6-(bis)phosphate intermediate, Biochemistry 45 (2006) 7818–7824.

[7] M. Golicnik, L. F. Olguin, G. Feng, N. J. Baxter, J. P. Waltho, N. H. Williams, F. Hollfelder, Kinetic Analysis of β-Phosphoglucomutase and Its Inhibition by Magnesium Fluoride, Journal of the American Chemical Society 131 (2009) 1575–88.

[8] N. J. Baxter, L. F. Olguin, M. Golic, G. Feng, A. M. Hounslow, W. Bermel, G. M. Blackburn, F. Hollfelder, J. P. Waltho, N. H. Williams, A Trojan horse transition state analogue generated by MgF_3^- formation in an enzyme active site, Proceedings of the National Academy of Sciences (USA) 103 (2006) 14732–14737.

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

[10] N. J. Baxter, M. W. Bowler, T. Alizadeh, M. J. Cliff, A. M. Hounslow, B. Wu, D. B. Berkowitz, N. H. Williams, G. M. Blackburn, J. P. Waltho, Atomic details of near-transition state conformers for enzyme phosphoryl transfer revealed by MgF_3^- rather than by phosphoranes, Proceedings of the National Academy of Sciences (USA) 107 (2010) 4555–4560.

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

[12] Y. Jin, D. Bhattasali, E. Pellegrini, S. M. Forget, N. J. Baxter, M. J. Cliff, M. W. Bowler, D. L. Jakeman, G. M. Blackburn, J. P. Waltho, α-Fluorophosphonates reveal how a phosphomutase conserves transition state conformation over hexose recognition in its two-step reaction, Proceedings of the National Academy of Sciences (USA) 111 (2014) 12384–12389.

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

[14] W. Wang, R. Kim, J. Jancarik, H. Yokota, S.-H. Kim, Crystal Structure of Phosphoserine Phosphatase from *Methanococcus jannaschii*, a Hyperthermophile, at 1.8 Å Resolution, Structure 9 (2001) 65 – 71.

[15] W. Wang, H. S. Cho, R. Kim, J. Jancarik, H. Yokota, H. H. Nguyen, I. V. Grigoriev, D. E. Wemmer, S.-H. Kim, Structural Characterization of the Reaction Pathway in Phosphoserine Phosphatase: Crystallographic

snapshots of Intermediate States, Journal of Molecular Biology 319 (2002) 421 – 431.

[16] K. N. Allen, D. Dunaway-Mariano, Phosphoryl group transfer: evolution of a catalytic scaffold, Trends in Biochemical Sciences 29 (2004) 495–503.

[17] C. Pandya, J. D. Farelli, D. Dunaway-Mariano, K. N. Allen, Enzyme promiscuity: Engine of evolutionary innovation, Journal of Biological Chemistry 289 (2014) 30229–30236.

[18] K. N. Allen, D. Dunaway-Mariano, Catalytic scaffolds for phosphoryl group transfer, Current Opinion in Structural Biology 41 (2016) 172 – 179. Multi-protein assemblies in signaling Catalysis and regulation.

[19] H. Huang, C. Pandya, C. Liu, N. F. Al-Obaidi, M. Wang, L. Zheng, S. Toews Keating, M. Aono, J. D. Love, B. Evans, R. D. Seidel, B. S. Hillerich, S. J. Garforth, S. C. Almo, P. S. Mariano, D. Dunaway-Mariano, K. N. Allen, J. D. Farelli, Panoramic view of a superfamily of phosphatases through substrate profiling, Proceedings of the National Academy of Sciences 112 (2015) E1974–E1983.

[20] Y. Jin, N. G. Richards, J. P. Waltho, G. M. Blackburn, Metal Fluorides as Analogues for Studies on Phosphoryl Transfer Enzymes, Angewandte Chemie International Edition 56 (2017) 4110–4128.

[21] Y. Jin, R. W. Molt, G. M. Blackburn, Metal Fluorides: Tools for Structural and Computational Analysis of Phosphoryl Transfer Enzymes, Topics in Current Chemistry 375 (2017) 1–31.

[22] D. E. Wemmer, D. Kern, Beryllofluoride Binding Mimics Phosphorylation of Aspartate in Response Regulators, Journal of Bacteriology 187 (2005) 8229–8230.

[23] Y. Jin, R. W. Molt, J. P. Waltho, N. G. J. Richards, G. M. Blackburn, ^{19}F NMR and DFT Analysis Reveal Structural and Electronic Transition State Features for RhoA-Catalyzed GTP Hydrolysis, Angewandte Chemie International Edition 55 (2016) 3318–3322.

[24] Y. Jin, R. W. Molt, E. Pellegrini, M. J. Cliff, M. W. Bowler, N. G. J. Richards, G. M. Blackburn, J. P. Waltho, Assessing the Influence of

Mutation on GTPase Transition States by Using X-ray Crystallography, ^{19}F NMR, and DFT Approaches, Angewandte Chemie International Edition 56 (2017) 9732–9735.

[25] J. G. Sonicki, M. Langaard, P. E. Hansen, Long-range deuterium isotope effects on ^{13}C chemical shifts of intramolecularly hydrogen-bonded N-substituted 3-(cycloamine)thiopropionamides or amides: a case of electric field effects, The Journal of organic chemistry 72 (2007) 41084116.

[26] N. Tokuriki, D. S. Tawfik, Stability effects of mutations and protein evolvability, Current Opinion in Structural Biology 19 (2009) 596 – 604. Carbohydradtes and glycoconjugates / Biophysical methods.

[27] E. Dellus-Gur, M. Elias, E. Caselli, F. Prati, M. L. Salverda, J. A. G. de Visser, J. S. Fraser, D. S. Tawfik, Negative Epistasis and Evolvability in TEM-1 β-LactamaseThe Thin Line between an Enzyme's Conformational Freedom and Disorder, Journal of Molecular Biology 427 (2015) 2396 – 2409.

[28] A. Pabis, F. Duarte, S. C. L. Kamerlin, Promiscuity in the enzymatic catalysis of phosphate and sulfate transfer, Biochemistry 55 (2016) 3061–3081. PMID: 27187273.

[29] D. Petrović, V. A. Risso, S. C. L. Kamerlin, J. M. Sanchez-Ruiz, Conformational dynamics and enzyme evolution, Journal of The Royal Society Interface 15 (2018).

[30] S. D. Lahiri, G. Zhang, P. Rådström, D. Dunaway-Mariano, K. N. Allen, Crystallization and preliminary X-ray diffraction studies of β-phosphoglucomutase from *Lactococcus lactus*, Acta Crystallographica Section D 58 (2002) 324–326.

[31] S. D. Lahiri, G. Zhang, D. Dunaway-Mariano, K. N. Allen, Caught in the Act: The Structure of Phosphorylated β-Phosphoglucomutase from *Lactococcus lactis*, Biochemistry 41 (2002) 8351–8359. PMID: 12081483.

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

[33] A. Vagin, A. Teplyakov, *MOLREP*: an Automated Program for Molecular Replacement, Journal of Applied Crystallography 30 (1997) 1022–1025.

[34] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of *Coot*, Acta Crystallographica Section D 66 (2010) 486–501.

[35] G. N. Murshudov, A. A. Vagin, E. J. Dodson, Refinement of Macromolecular Structures by the Maximum-Likelihood Method, Acta Crystallographica Section D 53 (1997) 240–255.

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

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

[38] R. J. Read, A. J. Schierbeek, A phased translation function, Journal of Applied Crystallography 21 (1988) 490–495.

[39] S. Hayward, H. J. Berendsen, Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and t4 lysozyme, Proteins: Structure, Function, and Bioinformatics 30 (1998) 144–154.

[40] J. L. Griffin, Investigations of the Metal Fluoride Transition state and Ground State Analogue Complexes of HAD superfamily Proteins by Nuclear Magnetic Resonance Spectroscopy, Ph.D. thesis, University of Sheffield, 2011.

Supporting Information.

1. Methods

1.1. Protein expression and purification

The *pgmB* gene from *Lactococcus lactis* together with the *pgmB* gene containing the T16A mutation were cloned in pET22b+ expression vectors and used to express βPGM$_{WT}$ and βPGM$_{T16A}$ proteins in *E. coli* strain BL21(DE3). One liter cell cultures were grown to log phase in either LB media or M9 media (with ^{15}N isotopic enrichment), induced with 1 mM IPTG and grown for a further 16 h at 25 °C. Cells were harvested by centrifugation at 10,000 rpm for 10 min at 4 °C, decanted and frozen at –80 °C. Cell pellets were resuspended in ice-cold standard native buffer (50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3) supplemented with one tablet of cOmplete^{TM} protease inhibitor cocktail (Roche). The cell suspension was lysed on ice by sonication for 6 cycles of pulsation for 20 s with 60 s cooling intervals. The cell lysate was then separated by ultracentrifugation (Beckman Coulter Avanti centrifuge) at 24,000 rpm for 35 min at 4 °C to remove insoluble matter. The cleared cell lysate was filtered using a 0.2 μm syringe filter and loaded onto a DEAE-Sepharose fast flow ion exchange column connected to an ÄKTA purification system that had been washed previously with 1 column volume of 6 M guanidine hydrochloride (GuHCl), 1 column volume of 1 M NaOH and equilibrated with $>$ 2 column volumes of standard native buffer. Following extensive washing, proteins bound to the DEAE-Sepharose column were eluted with a gradient of 0 to 100% standard native buffer containing 0.5 M NaCl. Fractions containing βPGM were checked for purity using SDS-PAGE, were pooled together and concentrated by Vivaspin (10 kDa MWCO). The protein sample was filtered using a 0.2 μm syringe filter and loaded onto a prepacked Hiload 26/60 Superdex 75 size-exclusion column connected to an ÄKTA purification system that had been pre-equilibrated with filtered and degassed standard native buffer containing 1 M NaCl. βPGM eluted as a single peak and fractions containing βPGM were checked for purity using SDS-PAGE, were pooled together, buffer exchanged into standard native buffer and concentrated to 1 mM by Vivaspin (10 kDa MWCO) for storage as 1 mL aliquots at –20 °C. The overall yield for βPGM was *ca.* 60 mg protein from 1 L of bacterial culture.

The MJ1594 gene from *Methanococcus jannaschii* was cloned into pET-21a(+) vector using Nde1 and BamH1 restriction enzyme sites. To generate the E20A PSP variant (PSP$_{E20A}$), mutagenic primers were purchased from Eurofins Scientific, mutagenesis was performed using a QuikChange II site-

directed mutagenesis kit (Agilent Technologies), and the resulting plasmids were sequenced by GATC BioTech. *E. coli* strain BL21(DE3) was used to express PSP using largely the same protocol as described for βPGM, however cells were grown for a further 4-6 h at 30 °C following induction with 1mM IPTG. Cells were harvested by centrifugation at 10,000 rpm for 10 min at 4 °C, decanted and frozen at –80 °C. Cell pellets were resuspended ice-cold PSP lysis buffer (20 mM TRIS pH 8.0, 1mM EDTA, 2mM NaN_3) supplemented with one tablet of cOmplete^{TM} protease inhibitor cocktail (Roche). The cell suspension was lysed on ice by sonication for 3-5 cycles of pulsation for 20 s with 60 s cooling intervals and the cell lysate separated by ultracentrifugation (Beckman Coulter Avanti centrifuge) at 24,000 rpm for 35 min at 4 °C to remove insoluble matter. The supernatant following this initial centrifugation was then incubated in a water bath held at 70 °C for 20 min to denature native *E. coli* proteins (as PSP from *M. jannaschii* is highly thermostable). The sample was again centrifuged at 24,000 rpm for 35 min to remove insoluble matter. The supernatant was loaded onto a Q-sepharose ion exchange column (GE Healthcare) connected to an ÄKTA purification system that had been washed previously with 1 column volume of 6 M guanidine hydrochloride (GuHCl), 1 column volume of 1 M NaOH and equilibrated with $>$ 2 column volumes of PSP purification buffer (20mM TRIS pH 8.0, 2mM EDTA, 10mM DTT, 2mM NaN_3). PSP was eluted from the column without binding, while most of the contaminating protein bound to the column. Fractions containing PSP were checked for purity using SDS-PAGE and were then pooled together and concentrated by Vivaspin (10 kDa MWCO). The protein sample was loaded onto a prepacked Hiload 26/60 Superdex 75 size-exclusion column connected to an ÄKTA purification system that had been pre-equilibrated with filtered and degassed PSP purification buffer (20 mM TRIS pH 8.0, 2mM EDTA, 10mM DTT, 2mM NaN_3, 300mM NaCl). PSP eluted as a single peak and fractions containing PSP were checked for purity using SDS-PAGE, were pooled together, buffer exchanged (>8000 fold dilution) into standard PSP buffer (20 mM TRIS, 20mM BISTRIS, 10mM $MgCl_2$, 10mM DTT, 2mM NaN_3, at pH 7.5) and concentrated to 1.5 mM by Vivaspin (10 kDa MWCO) for storage as 0.5-1 mL aliquots at –20 °C.

Unless otherwise stated, reagents and purification equipment were purchased from Sigma-Aldrich, GE Healthcare, Melford Laboratories or Cotec-Net

1.2. Crystallization and soaking experiments

Crystallization of native βPGM_{WT} was achieved using the same conditions as described previously [10]. The βPGM_{WT} protein solution was mixed 1:1 with precipitants (26-30% (w/v) PEG 4000, 200 mM sodium acetate and 100 mM Tris-HCl (pH 7.5)) and crystals were grown at 290 K by hanging-drop vapor diffusion using a 2 μL drop suspended on a siliconized glass cover slip above a 700 μL well. Rod shaped crystals formed after several days which were cryo-protected in their original mother liquor containing an additional 25% (v/v) ethylene glycol prior to plunging into liquid nitrogen. For the acetylphosphate (AcP) soaking experiments, native βPGM_{WT} crystals were cryo-protected in their original mother liquor containing an additional 25% (v/v) ethylene glycol together with 30 mM AcP, and were incubated for a range of timescales (30 – 900 s) prior to plunging into liquid nitrogen. For the AlF_4 soaking experiments, native βPGM_{WT} crystals were cryo-protected in their original mother liquor containing an additional 25% (v/v) ethylene glycol with an further 5mM Al^{3+}, 20 mM NaF, and 10mM glucose, and were incubated for ca. 60s prior to plunging into liquid nitrogen.

Co-crystallization of βPGM_{WT} with MgF_3^- in the active site was achieved using the same conditions described above, but with the addition of 15mM NaF to the protein solution prior to mixing with precipitants. Rod shaped crystals grew and were cryo-protected in their original mother liquor containing an additional 25% (v/v) ethylene glycol prior to plunging into liquid nitrogen. Co-crystallization of βPGM_{T16A} with inorganic phosphate bound in the *distal* site was achieved serendipitously. Crystallization attempts were initially laid as βPGM_{T16A}:AlF_4:βG1P transition state analogue complexes [10] with the addition of 20 mM NaF, 5 mM $AlCl_3$ and 15 mM βG1P (synthesized enzymatically from maltose using maltose phosphorylase [13]) to 0.6 mM βPGM_{T16A} in standard native buffer. Protein solutions were mixed 1:1 with precipitants (24-30% (w/v) PEG 4000, 200 mM sodium acetate and 100 mM Tris-HCl (pH 7.5)) and crystals were grown at 290 K by hanging-drop vapor diffusion using a 2 μL drop suspended on a siliconized glass cover slip above a 700 μL well. Rod shaped crystals formed after several days which were cryo-protected in their mother liquor containing an additional 25% (v/v) ethylene glycol prior to plunging into liquid nitrogen. Refinement of structures from this complex yielded inorganic phosphate and Tris buffer coordinated in the active site.

1.3. Phospho-βPGM structures.

A structure of the native phospho-βPGM enzyme (βPGM^{P}, phosphorylated at residue D8) was also investigated for comparison with the metal fluoride complexes. The previously reported phospho-βPGM enzyme crystal (**PDB: 1LVH**, 2.3 Å resolution) [30, 31], was grown under condition where no phospho-βPGM enzyme was present [8, 10] and on the basis of 100 mM NH_4F in the crystallization conditions was postulated to contain an AlF_4 group in the active site [10]. On closer inspection of the difference Fourier maps calculated after refinement of structure PDB: 1LVH against the deposited structure factors there are significant discrepancies from the initial interpretation of the moieties present in the *proximal* site. In the difference map (Fig. **??**), negative peaks are observed at both the Mg^{2+} binding site (ca. 8–12 σ), and the phosphorylation site of D8 (ca. 6–9 σ) for chain B and A respectively. Replacement of the phosphate group with AlF_4 and refinement against the deposited structure factors eliminates peripheral peaks in the difference Fourier map, but a central negative peak remains at ca. 5.5–8.5 σ for chain B and A respectively. This central peak indicates that the true atomic species is likely to be heavier than aluminium, but that the overall moiety may be coordinated by additional waters or fluorides. Thus, the electron density in the catalytic site of the 1LVH structure is not satisfied by either a phosphate or an AlF_4^- group, and the question remains as to the identity of the species observed.

1.4. Data collection and refinement

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 [32] with resolution cut-offs applied using CC-half values and the structures were determined by molecular replacement with MolRep [33] using PDB: 2WHE of 2WFA as a search models. Model building was carried out in COOT [34] and a restrained refinement with isotropic temperature factors was performed using REFMAC5 [35] in the CCP4i suite [36]. Ligands and protein modifications were not included until the final stages of refinement to avoid biasing Fourier maps. Structure validation was carried out in COOT and MolProbity [37], superpositions were generated using PyMOL (The PyMOL Molecular Graphics System, version 1.8/2.0 Schrödinger, LLC), maps were generated using FFT [38] and domain movements were calculated using DynDom [39].

1.5. 1D ^{19}F NMR spectra

1D ^{19}F spectra were acquired at 298 K using a Bruker 500 MHz Avance III spectrometer equipped with a 5-mm QCI-F cryoprobe and z-axis gradients at the Manchester Institute of Biotechnology, The University of Manchester. 1-2 mM βPGM$_{\mathrm{WT}}$ samples were prepared in standard NMR buffer (50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3, 10% D_2O and 1 mM trimethylsilyl propionic acid (TSP). To form the metallofluoride complexes, either 15 mM NH_4F was added to the βPGM$_{\mathrm{WT}}$ solution to make the βPGM$_{\mathrm{WT}}$:MgF_3 complex or 15 mM NH_4F and 3 mM $AlCl_3$ were added to make the βPGM$_{\mathrm{WT}}$:AlF_4 complex. In some experiments, 10 mM glucose was also included in the solution, however glucose does not have a high affinity for βPGM$_{\mathrm{WT}}$ as demonstrated by the absence of resonance chemical shift changes in the ^{19}F NMR spectra (Fig. S21). 1-1.5mM PSP$_{\mathrm{WT}}$ samples were prepared in standard PSP NMR buffer (20 mM TRIS, 20mM BISTRIS, 10mM $MgCl_2$, 10mM DTT, 2mM NaN_3, 10% D_2O, at pH 7.5). To form BeF_3^- complexes a further 5mM $BeCl_2$ and 10mM NaF were added to these samples, while to form MgF_3^- compleses, a further 20mM NaF was added. 1D ^{19}F spectra were typically accumulations of 512 – 2048 transients incorporating a 1.5–2.5 s inter-scan delay over a spectral width of 120.77 ppm centered at -140 ppm, using Bruker internal referencing. The linewidth at half height of protein bound fluorine peaks was determined using the deconvolution tool in Topspin v3.5 (Bruker) and all errors are presented at one standard deviation.

2. Tables

Data Acquisition

Complex	βPGM$_{WT}$	βPGM$_{WT}$	βPGM$_{WT}$: AlF_4^-	βPGM$_{WT}$: MgF_3^-	βPGM$_{T16A}$: AlF_4^-	βPGM$_{T16A}$: MgF_3^-	βPGM$_{K145A}$: B
PDB code	6H8U	6H8V	6H8W	6H8X	6H8Y	6H8Z	6H90
Wavelength (Å)	0.97949	0.97949	0.97949	0.97950	0.97950	0.97950	0.97949
Beamline	i04	i04	i04	i04	i04	i04	i02
Facility	DLS	DLS	DLS	DLS	DLS	DLS	DLS
Space group	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions							
a, b, c, (Å)	53.15, 53.64, 81.07	53.10, 53.71, 81.09	53.44, 53.71, 80.88	38.16, 116.98, 53.11	52.55, 52.84, 78.32	47.27, 53.89, 82.27	52.32, 53.63, 81.5
α,β,γ (^{o})	90.00, 90.00, 90.00	90.00, 90.38, 90.00	90.00, 90.00, 90.00	90.00, 98.93, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.0
Resolution (Å) [1]	44.45 - 1.90 (1.95 - 1.90)	44.29 - 1.84 (1.89 - 1.84)	40.44 - 1.98 (2.03 - 1.98)	23.40 - 1.83 (1.86 - 1.83)	31.46 - 1.89 (1.92 - 1.89)	45.08 - 1.60 (1.63 - 1.60)	40.75 - 1.31 (1.34
R_{merge} [1,2]	0.118 (0.908)	0.073 (0.681)	0.116 (0.738)	0.100 (1.258)	0.096 (1.698)	0.067 (1.715)	0.068 (1.101)
R_{pim} [1,3]	0.073 (0.544)	0.063 (0.503)	0.075 (0.459)	0.040 (0.501)	0.034 (0.593)	0.027 (0.703)	0.032 (0.517)
CC-half	0.996 (0.604)	0.995 (0.556)	0.992 (0.576)	0.999 (0.568)	0.999 (0.580)	0.999 (0.579)	0.999 (0.539)
$< I/\sigma I >$ [1]	9.6 (1.6)	10.2 (1.6)	9.5 (1.9)	11.6 (1.2)	12.6 (1.3)	10.8 (1.0)	13.1 (1.5)
Completeness (%) [1]	99.7 (99.8)	85.5 (89.3)	99.8 (99.9)	98.5 (97.8)	100.0 (100.0)	100.0 (99.4)	99.9 (99.9)
Multiplicity [1]	4.3 (4.2)	2.6 (2.6)	4.3 (4.3)	7.0 (7.2)	8.8 (9.1)	6.9 (6.7)	6.5 (6.5)
Total reflections	80231	89124	71414	278152	159783	196086	360990
Unique reflections	18801	33973	16779	39841	18067	28426	55835
Molecular replacement model	2WHE	2WHE	2WHE	2WHE	2WHE	2WHE	2WHE

Data Refinement

Complex	βPGM$_{WT}$	βPGM$_{WT}$	βPGM$_{WT}$: AlF_4^-	βPGM$_{WT}$: MgF_3^-	βPGM$_{T16A}$: AlF_4^-	βPGM$_{T16A}$: MgF_3^-	βPGM$_{K145A}$: B
PDB code	6H8U	6H8V	6H8W	6H8X	6H8Y	6H8Z	6H90
R(%) [4] / R_{free} (%) [5]	21.07 / 23.93	18.60 / 23.56	20.46 / 24.81	18.27 / 22.58	20.67 / 25.95	23.64 / 28.37	13.53 / 16.23
Protein [6]	1680	1688, 1689	1697	1689, 1697	1686	1696	1741
Ligands [7]	0	0	5	8	5	4	4
Metal ions	1	2	1	2	1	1	2
Water	110	403	136	214	44	273	291
Protein residues [6]	218	218	219	219	218	219	219
RMS deviations:							
Length	0.013	0.013	0.013	0.012	0.013	0.012	0.012
Angles	1.48	1.50	1.48	1.54	1.51	1.53	1.49
Average B factors(A^2)							
Main chain [6]	25.71	18.70, 19.97	22.78	28.59, 29.82	38.68	31.73	13.42
Side chains [6]	29.49	21.65, 23.55	26.56	33.76, 34.79	44.53	36.45	17.95
Ligands	–	–	29.31	35.28	35.72	28.85	11.18
Metal Ions [8]	27.52	21.26	33.27	28.13	33.68	21.74	10.80, 46.19
Water	29.25	28.30	26.70	35.32	39.45	34.06	27.32
Ramachandran analysis							
Favored/allowed (%)	97.69	97.22	97.70	98.39	98.15	96.31	98.22
Disallowed (%)	0.00	0.00	0.00	0.00	0.46	0.00	0.00
Molprobity score (percentile)	1.00 (100^{th})	0.94 (100^{th})	0.91 (100^{th})	0.61 (100^{th})	0.77 (100^{th})	1.49 (89^{th})	1.18 (96^{th})

Table S1:

[1]Values for the higher resolution shell are in parenthesis.

[2] $R_{merge} = \Sigma_{hkl}\Sigma_i \mid I_i - I_m \mid / \Sigma_{hkl}\Sigma_i I_i$.

[3] $R_{pim} = \Sigma_{hkl} \sqrt{1/n\text{-}1}\Sigma_{i=1} \mid I_i - I_m \mid / \Sigma_{hkl}\Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively.

[4]$R=\sum_{hkl} || F_{obs} | - k| F_{calc} || I \sum_{hkl} | F_{obs} |$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.

[5]$R_{free}=\sum_{hklIT} || F_{obs}| - k| F_{calc}|| I \sum_{hklIT} | F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes and T is the test set of data omitted from refinement (5% in this case).

[6] For structures where there are two proteins in the asymmetric unit, the value for chain A will be given first, then the value for chain B.

[7] Only the ligands that are the subject of investigation are presented. Other ligands such as ethylene glycol and acetate, etc. that are part of the crystallization solution are not presented.

[8] Frequently the only observable metal ion was Mg^{2+}, however in some cases Na^+ ions were also observed. Where this was the case, the B-factor for Mg^{2+} ions were given first, followed by Na^+ ions.

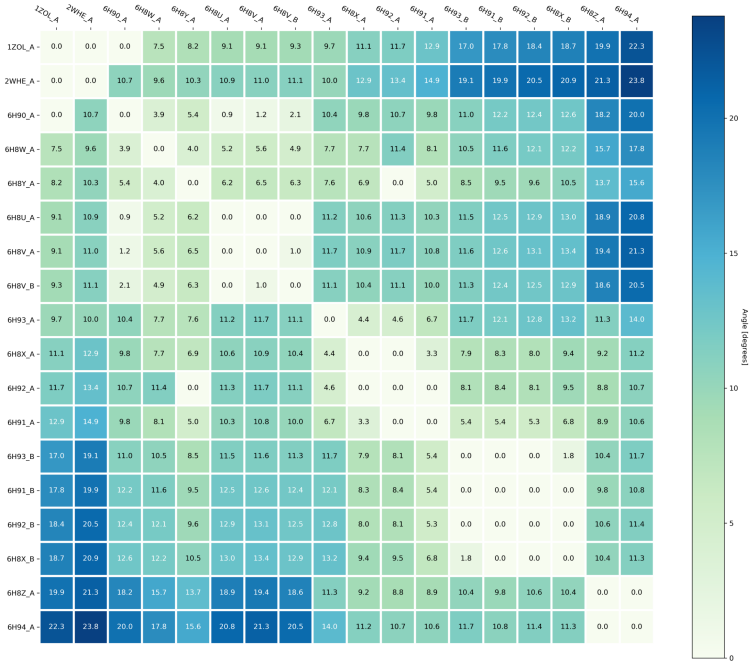


Table S2: A structural comparison of all complexes discussed in the text using and DynDom [39] to generate angles (°) of rotation to map cap domains between different complexes in βPGM. Comparisons where no dynamic domains were found are denoted with an angle of 0.0.

Table S3: A structural comparison of all complexes discussed in the text using and DynDom [39] to generate non-H atom RMSD values for aligned cap and core domain between different complexes in βPGM. The top right side of the matrix indicates the RMSD value for the cap domain (Å), while the bottom left side of the matrix indicates the RMSD for the core domains (Å). The diagonal is necessarily indicates no difference. Comparisons where no dynamic domains were found are denoted with an angle of 0.0.

3. Phosphorylation of pre-formed substrate-free βPGM crystals by AcP

3.1. The previously reported phospho-enzyme in βPGM

While BeF_3^- species are reportedly good mimics of native phospho-enzyme states [11], only few direct comparisons have been made structurally [22], and a direct comparison has not yet been made in βPGM. This phospho-enzyme state in βPGM (βPGM$^{\mathrm{P}}$, phosphorylated at residue D8) is only transient in solution (ca. 30s lifetime [7, 13]) which precludes standard crystallographic approaches of co-crystallization. The structure of βPGM$^{\mathrm{P}}$ was previously reported (**PDB: 1LVH**, 2.3 Å resolution) [30, 31], however, the chemical implausibility of this complex (especially in the presence of 100 mM fluoride) was highlighted previously [10]. In the follow up study, substrate-free βPGM crystals were grown under conditions where no βPGM$^{\mathrm{P}}$ was present as validated by solution NMR [8, 10], on the basis of 100 mM NH_4F in the crystallization conditions, it was postulated that the previously ascribed phosphate group may in fact be an AlF_4 group in the active site [10].

On closer inspection of the difference Fourier maps calculated after refinement of structure PDB: 1LVH against the deposited structure factors, there are significant discrepancies from the initial interpretation of the moieties present in the *proximal* site (Fig. S1). In the difference map, negative peaks are observed at both the Mg^{2+} binding site (*ca.* 8–12 σ), and the phosphorylation site of D8 (*ca.* 6–9 σ) for chain B and A respectively. Replacement of the phosphate group with AlF_4 and refinement against the deposited structure factors eliminates peripheral peaks in the difference Fourier map, but a central negative peak remains at *ca.* 5.5–8.5 σ for chain B and A respectively. This central peak indicates that the true atomic species is likely to be heavier than aluminum, but that the overall moiety may be coordinated by additional waters or fluorides.

Inspection of the hydrogen bonding pattern around the phosphate group reveals that in PDB: 1LVH, a charge-balancing hydrogen bond is not made between the sidechain amine of K145 and the phosphate group that is made in all phosphate and phosphate-analogue structures subsequently [10, 11, 13]. Furthermore, hydrogen bonds from the backbone amides of residues L9, D10, and A115 to the phosphate group are also missing, together indicating that moiety present in the crystal displays a different charge and/or geometry to that of phosphate. Thus, the electron density in the catalytic site of the 1LVH structure is not satisfied by either a phosphate or an AlF_4 group, and

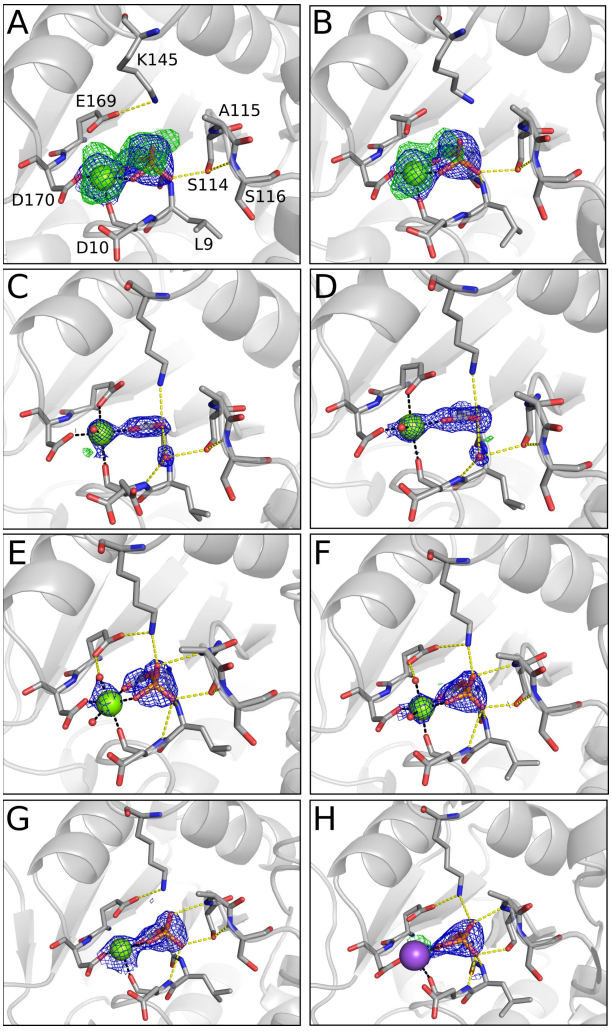


Figure S1: Electron density (2Fo-Fc, blue) and omit map density (Fo-Fc, green) are drawn around the catalytic Mg^{2+} ion and D8 residue for PDB: 1LVH (**A, B**), 6H8V (**C, D**), 6H91 (**E, F**), 6H92 (**G, H**). For **A** and **B**, 2Fo-Fc density is contoured at 3.0 σ and Fo-Fc density contoured at 5.0 σ. For **C** and **D**, 2Fo-Fc density is contoured at 2.0 σ and Fo-Fc density contoured at 3.0 σ. For **E** and **F**, 2Fo-Fc density is contoured at 3.0 σ and Fo-Fc density contoured at 3.0 σ. For **G** and **H**, 2Fo-Fc density is contoured at 2.0 σ and Fo-Fc density contoured at 3.0 σ. Hydrogen bonds ($\leq$ 3.2 Å) around the phosphate group are drawn as dashed yellow lines and metal coordination is illustrated with dashed black lines. Atoms are drawn using standard CPK colors, but with carbon colored in grey.

the identity of the species observed is not readily solvable. Given that the crystal was a Seleno-Met preparation for initial phase determination, if heavy metals were also included for phasing, then a cation would likely display the repulsion of positively charged moieties seen here.

3.2. Crystallization of the $\beta PGM_{WT}{}^{P}$:AcP_{90} complex

In order to attempt to trap the short-lived $\beta\mathrm{PGM_{WT}}^{\mathrm{P}}$ species under analogous conditions (and timeframe) to those used to directly observe the $\beta\mathrm{PGM_{WT}}^{\mathrm{P}}$ complex in solution [13], a flash freezing approach was adopted. Substrate-free $\beta\mathrm{PGM_{WT}}$ crystals (PDB: 6H8V) were transferred to cryoprotectant containing 30 mM AcP and were incubated for a range of timescales prior to flash freezing in liquid nitrogen. Crystals that were soaked with AcP for 90 s still belonged to the $P2_1$ spacegroup and were refined to a resolution of 2.4 Å (**PDB: 6H91**, Table S1). The two monomers present in the asymmetric unit closely resemble each other and the cap and core domains have an open arrangement as observed substrate-free βPGM structures (Table S3). Both monomers in this structure show clear electron density for a phosphate group covalently bonded to the carboxylate Oδ1 atom of the catalytic aspartate residue (D8), identifying the complex as $\beta\mathrm{PGM_{WT}}^{\mathrm{P}}$, with Mg^{2+} acting to charge balance the *proximal* binding site (Fig. S2). This complex is termed the $\beta\mathrm{PGM_{WT}}^{\mathrm{P}}$:$\mathrm{AcP_{90}}$ complex for clarity. Furthermore, both monomers demonstrate a close resemblance to the structure of the $\beta\mathrm{PGM_{WT}}$:$\mathrm{BeF_3}$ complex (PDB: 2WFA, [11]), with rotations of the cap domain relative to the core domain of 15°$ – 20°$ (Table S4, S5, S6). These rotations indicate a subtle cap opening (relative to a closed transition state analogue) compared to previously reported substrate free structures (PDB: 2WHE, 1Z0L) as a result of the change in spacegroup.

The catalytic Mg^{2+} ion is coordinated with a regular octahedral geometry by the sidechain carboxylate groups of D8 and D170, the backbone carbonyl group of D10, two structural water molecules and a phosphate oxygen atom from the $PO_3{}^-$ group, with Mg–O bond lengths in the range 1.9 - 2.2 Å. An equivalent Mg^{2+} ion coordination is present in the βPGM:$\mathrm{BeF_3}$ complex (PDB: 2WFA, [11]) except that the phosphate oxygen atom is replaced by a fluorine atom from the $\mathrm{BeF_3}{}^-$ moiety. Furthermore, near-identical hydrogen bond organization in the *proximal* binding site for the $\mathrm{BeF_3}{}^-$ moiety in the $\beta\mathrm{PGM_{WT}}$:$\mathrm{BeF_3}$ complex and for the $PO_3{}^-$ group in $\beta\mathrm{PGM_{WT}}^{\mathrm{P}}$:$\mathrm{AcP_{90}}$, indicate that the $\beta\mathrm{PGM_{WT}}$:$\mathrm{BeF_3}$ complex is validated as an excellent structural model for $\beta\mathrm{PGM_{WT}}^{\mathrm{P}}$ (Fig. S3, S4), as suggested previously [11, 13].

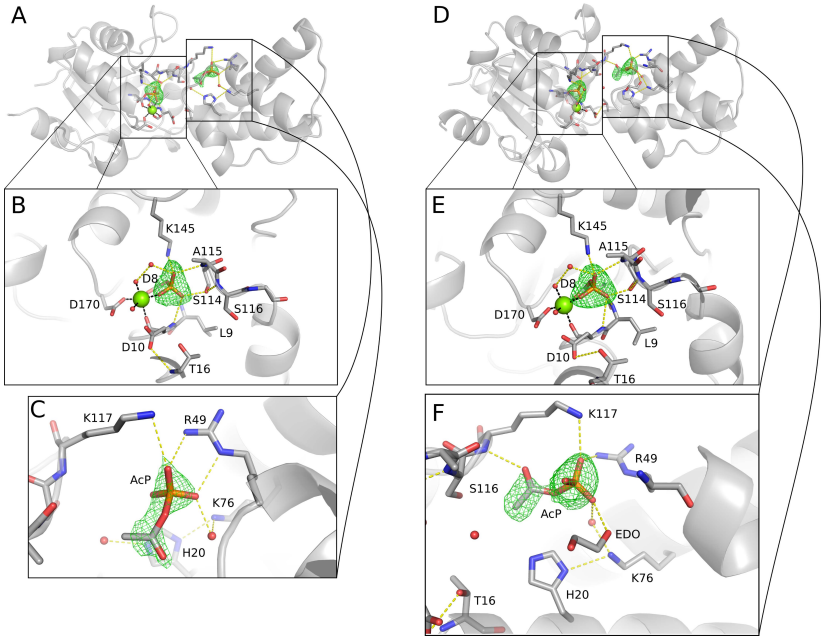


Figure S2: The active site of the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{90}$ complex (PDB: 6H91) with chain A (A), and B (B), shown with Fo-Fc difference density contoured at 3.0 σ. Hydrogen bonds to the phosphate group are drawn as dashed yellow lines and ionic interaction with the Mg^{2+} ion is illustrated with dashed black lines. Atoms are drawn using standard CPK colors, but with carbon colored in grey.

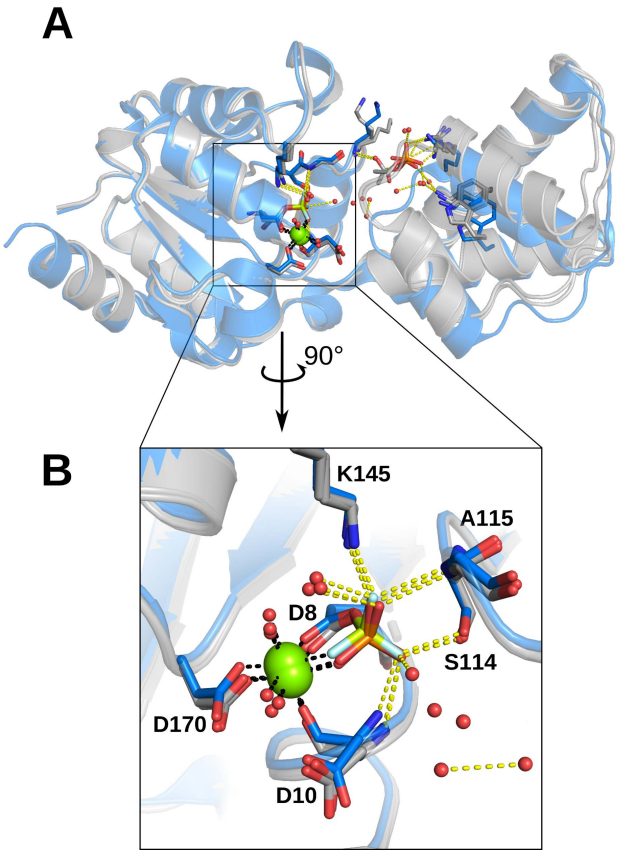


Figure S3: Overlay of the acetylphosphate soak after 90s with the BeF_3 phospho-enzyme analogue structure 2WFA. Chains A and B of the βPGM^P complex (PDB: 6H91) are colored in gray, while the βPGM:BeF_3 complex (PDB: 2WFA) is colored in blue for clarity. Water molecules are displayed as red spheres and atoms are colored using CPK coloring with O=red, N=blue, P=orange, Mg=green, Be=green, and F=gray. Metal coordination are shown with black dashed lines, and hydrogen bonds are shown with yellow dashed lines. **A** shows a subtle cap domain rotation between 2WFA and the new structures, but a near identical core domain. **B** A zoomed and rotated view of the the proximal catalytic site, highlighting almost identical coordination patterns of the phosphate group in the native phospho-enzyme, and the BeF_3 phospho-enzyme analog.

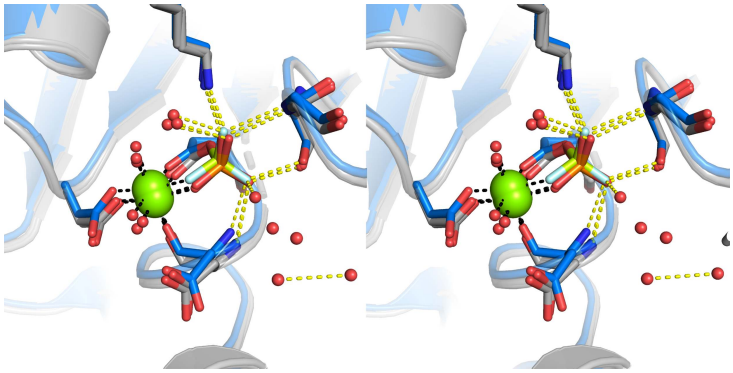


Figure S4: Stereo view of the phosphate coordination in the active site of PDB: 6H91 for chains A and B, overlaid with the BeF_3 phospho-enzyme analogue structure 2WFA. Chains A and B are colored in grey, while 2WFA is colored in blue for clarity. Water molecules are displayed as red spheres and atoms are colored using CPK coloring with O=red, N=blue, P=orange, Mg=green, Be=green, and F=gray. Metal coordination are shown with black dashed lines, and hydrogen bonds are shown with yellow dashed lines.

3.3. Crystallization of the $\beta PGM_{WT}{}^{P}$:AcP_{180} complex

Apo-βPGM$_{\mathrm{WT}}$ crystals that had been soaked with AcP for 180 s again belonged to the $P2_1$ spacegroup and were refined to a resolution of 2.6 Å (**PDB: 6H92**, Table S1), whereas βPGM$_{\mathrm{WT}}$ crystals incubated for longer times disintegrated prior to X-ray analysis. The two monomers present in the asymmetric unit closely resemble each other and their respective monomer in the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{90}$ complex (Table S4, S5, S6). The cap and the core domains are again in an open arrangement and both monomers are phosphorylated at residue D8 (Fig. S5). This complex is termed the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{180}$ complex. In chain A, the catalytic Mg^{2+} ion has incomplete coordination, being liganded by the sidechain carboxylate groups of D8 and D170, the backbone carbonyl group of D10 and a phosphate oxygen atom from the PO_3^- group, with Mg–O bond lengths in the range 2.0 - 2.4 Å. In chain B, the catalytic Mg^{2+} ion has been replaced by a Na^+ ion with identical coordinating partners, but with Na–O bond lengths in the range 2.3 - 3.0 Å.

In both βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{90}$ and βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{180}$ complexes, there is electron density present that indicates the presence of un-hydrolyzed AcP bound in the *distal* phosphate binding sites (FIG as evidence). In chain A of the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{90}$ complex, two phosphate oxygen atoms of AcP are hydrogen bonded by the guanidinium group of residue R49 (in the cap domain)

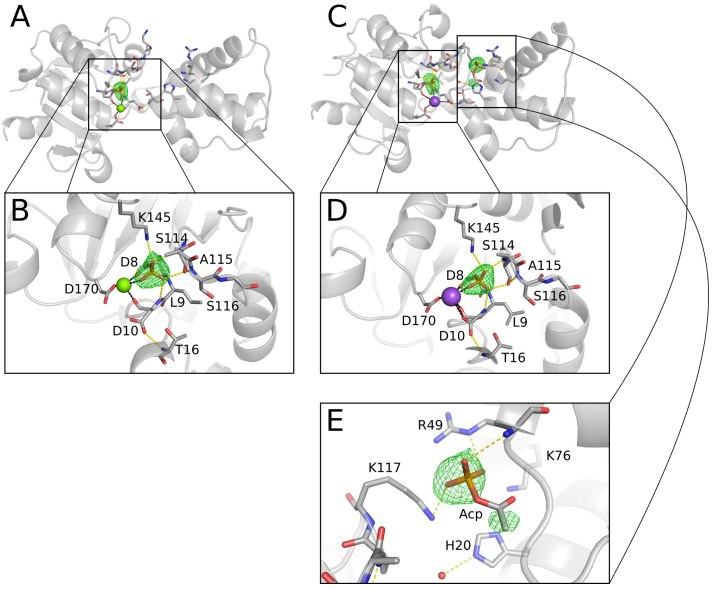


Figure S5: The active site of the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{180}$ complex (PDB: 6H92) with chain A (A) and B (B) shown with Fo-Fc difference density contoured at 3 σ. Hydrogen bonds to the phosphate group are drawn as dashed yellow lines and ionic interaction with the Mg^{2+} (A) or Na^{+} (B) ion is illustrated with dashed black lines.

and the carboxylate oxygen atom of AcP is coordinated by structural waters. In chain B, the phosphate group of AcP is hydrogen bonded as for chain A, but the carboxylate oxygen atom is hydrogen bonded by the backbone amide group of K117 (in the core domain), with the larger domain rotation observed for this monomer. In the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{180}$ complex, two phosphate oxygen atoms of AcP are hydrogen bonded by both the guanidinium group and the backbone amide group of R49 (in the cap domain) and the final phosphate oxygen atom is coordinated by the sidechain amine group of K117 (in the core domain), which reflects a different AcP hydrogen bond arrangement to that described for the monomers of the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{90}$ complex.

3.4. Commentary

Together, the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{90}$ and the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{180}$ complexes demonstrate that AcP is able to phosphorylate βPGM selectively at the catalytic aspartate residue (D8) yielding a crystal form of βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$ that is stable over a similar time-frame to solution forms of this complex [6, 7, 13]. The active sites of the resulting structures corroborate the structural homology between βPGM$^{\mathrm{P}}$ and the βPGM:BeF_3 GSA complex, further validating the use of $BeF_3{}^-$ as a phospho-enzyme surrogate. However, while $BeF_3{}^-$ was only observed to bind to the *proximal* site, AcP was also observed to bind to the *distal* phosphate binding site. In both βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{90}$ and βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{180}$ complexes, un-hydrolyzed AcP makes several hydrogen bonds to key conserved residues in the *distal* site, which may serve as a structural model of an AcP-dependent inhibition of catalysis reported previously [4].

Given that the resolution achievable by the βPGM:BeF_3 complex (1.3 Å) is significantly higher than for the βPGM$^{\mathrm{P}}$ structures generated here (2.4 Å), when combined with the higher longevity of the βPGM:BeF_3complex, it appears that $BeF_3{}^-$ is still a better alternative for structural investigation. However, if $BeF_3{}^-$ complexes do not form, or the phospho-enzyme state is of significant importance, then the inclusion of AcP in the cryoprotectant prior to flash freezing may serve as a suitable approach to generate native phospho-enzyme *in the crystal*.

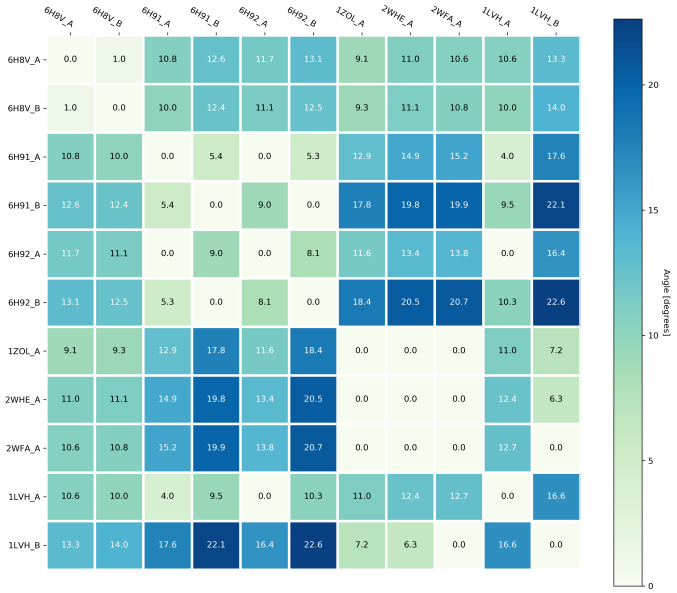


Table S4: A structural comparison of all complexes discussed in the text using and DynDom [39] to generate angles (°) of rotation to map cap domains between different complexes in βPGM. Comparisons where no dynamic domains were found are denoted with an angle of 0.0 and are approximated to be $< 1°$.

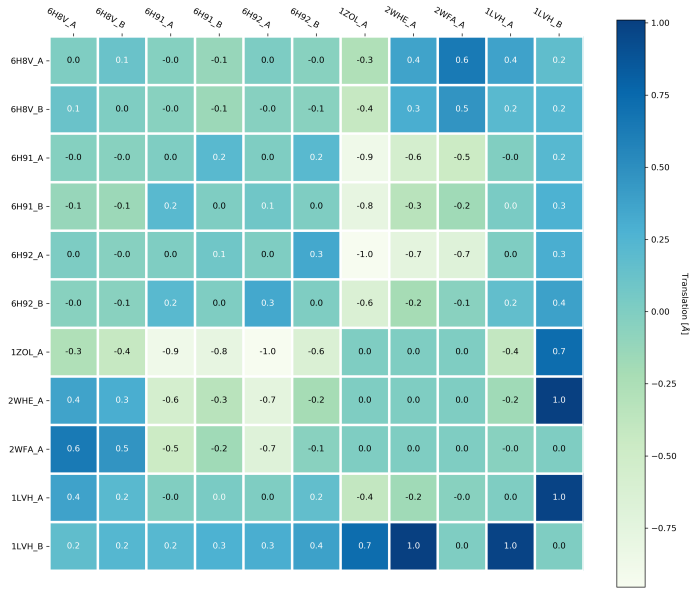


Table S5: A structural comparison of all complexes discussed in the text using and DynDom [39] to generate a translation (Å) necessary to map the cap domains between different complexes in βPGM following the rotation presented in Table S2. Comparisons where no dynamic domains were found are denoted with an angle of 0.0 and are approximated to be < 0.1 Å.

Table S6: A structural comparison of all complexes discussed in the text using and DynDom [39] to generate non-H atom RMSD values for aligned cap and core domain between different complexes in βPGM. The top right side of the matrix indicates the RMSD value for the cap domain (Å), while the bottom left side of the matrix indicates the RMSD for the core domains (Å). The diagonal necessarily indicates no difference. Comparisons where no dynamic domains were found are denoted with an angle of 0.0 and are approximated to be < 0.1 Å.

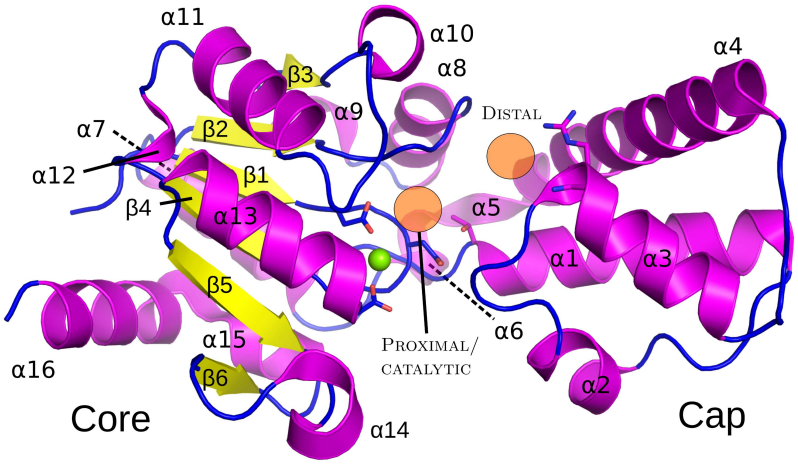


Figure S6: **A)** A ribbon representation of βPGM showing the core (left) and cap (right) domains with α-helices colored purple, β-sheets yellow, with standard CPK colors for atoms. The proximal and distal phosphate binding sites are indicated by orange circles.

4. Supplementary Figures

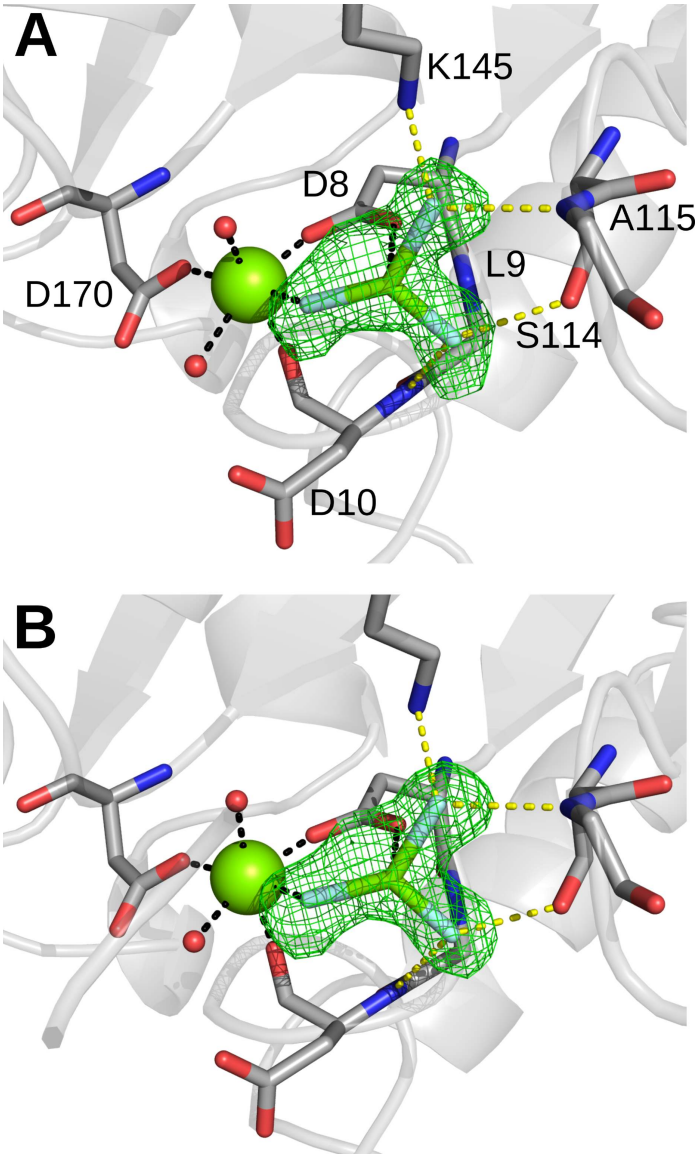


Figure S7: The active site of the βPGM$_{\mathrm{WT}}{}^{\mathrm{P}}$:AcP$_{90}$ complex with chain A (A) and B (B) shown with Fo-Fc difference density contoured at 3 σ. Hydrogen bonds to the phosphate group are drawn as dashed yellow lines and ionic interaction with Mg^{2+} is illustrated with dashed black lines.

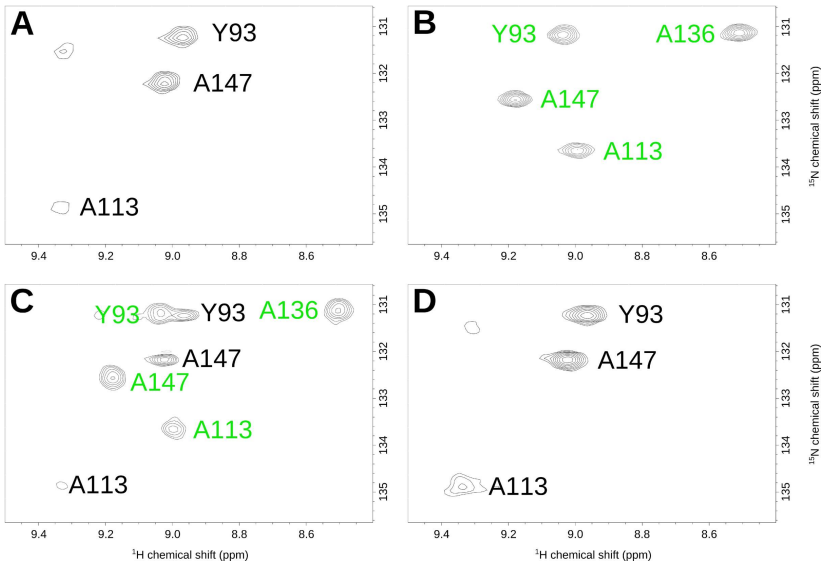


Figure S8: 2D ^{1}H^{15}N-TROSY NMR spectra of 1mM βPGM$_{\mathrm{T16A}}$ (50mM Tris pH 7.2, 5mM $MgCl_2$ 2mM NaN_3) in **A)** the open conformer and βPGM$_{\mathrm{T16A}}$ incubated with *ca.* 30mM acetylphosphate for **B)** 15 minutes, **C)** 21 minutes, **D)** 27 minutes. Transferred assignments from βPGM$_{\mathrm{WT}}$ open and βPGM$_{\mathrm{WT}}$:BeF_3^- complexes are shown in black and green respectively.

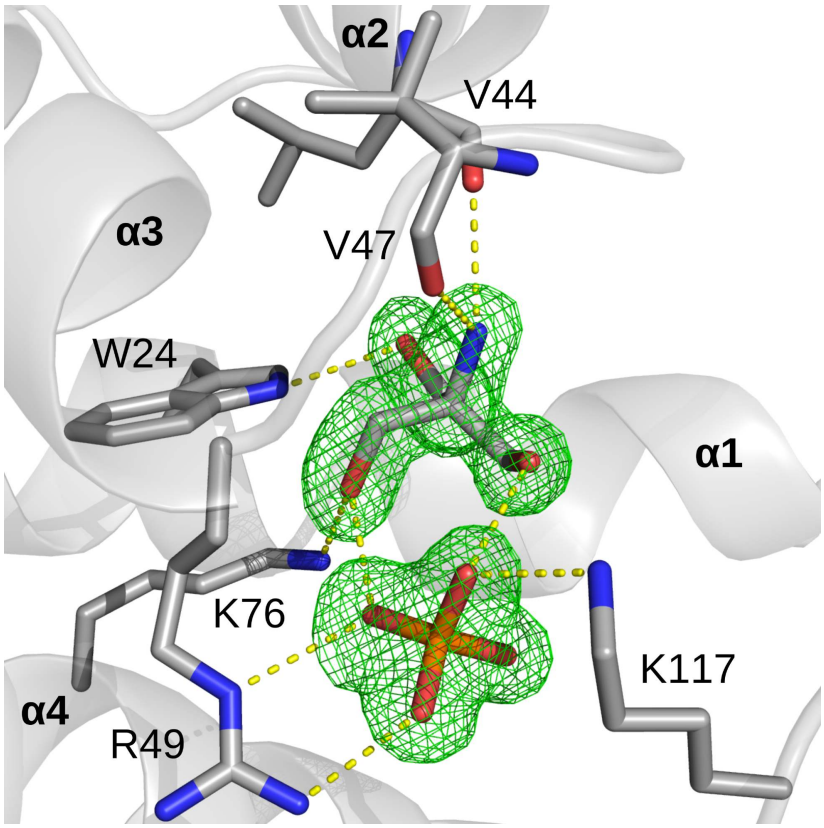


Figure S9: The *distal* phosphate binding site is shown with Fo-Fc difference density contoured at 3 σ. Hydrogen bonds to the phosphate group and TRIS molecule are drawn as dashed yellow lines.

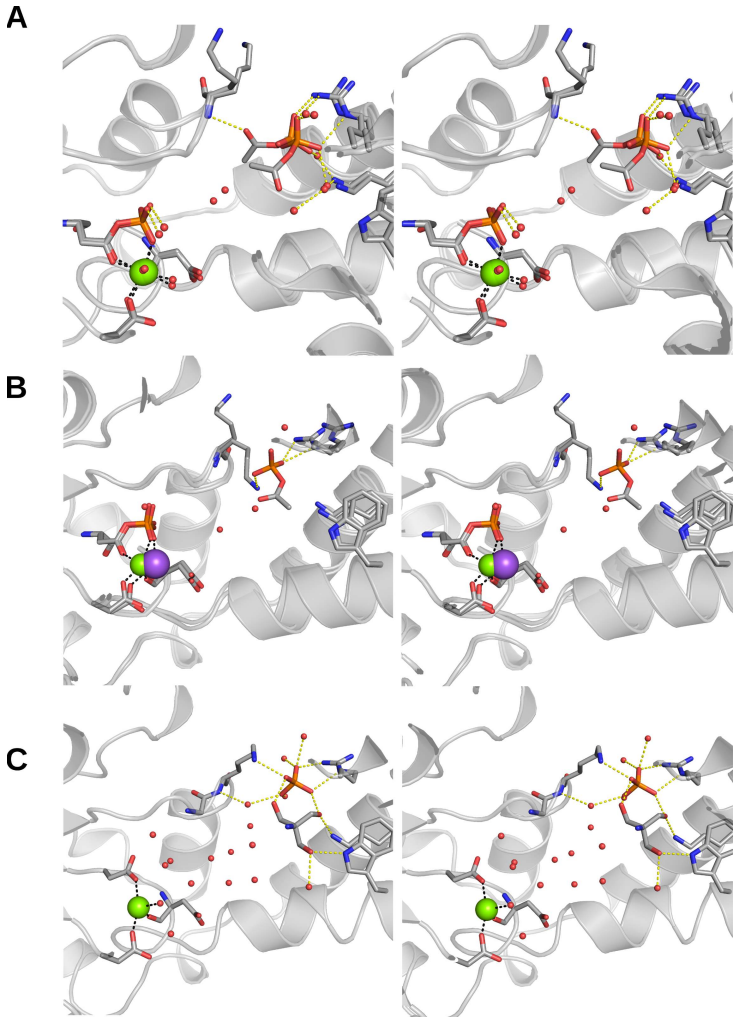


Figure S10: Stereo view of the active site of **A)** PDB: 6H91, 90s incubation with AcP, **B)** PDB: 6H92, 180s incubation with AcP, and **C)** the active site of the βPGM$_{\mathrm{T16A}}$:Pi complex. All structures were either P12$_1$1 or P2$_1$2$_1$2$_1$ spacegroup symmetry, with both monomers are overlaid in the stereo figure where applicable. Acetyl phosphate is shown in the *distal* site of **A** and **B** with phosphate and TRIS molecules shown in the distal site of **C**. Active site waters are shown as red spheres and atoms are colored using CPK coloring with O=red, N=blue, P=orange, Mg=green, Na=purple. Metal coordination are shown with black dashed lines, and hydrogen bonds are shown with yellow dashed lines. Unfortunately titration of Pi into substrate-free βPGM did not cause active site residues to come out of intermediate exchange (Fig. S12)

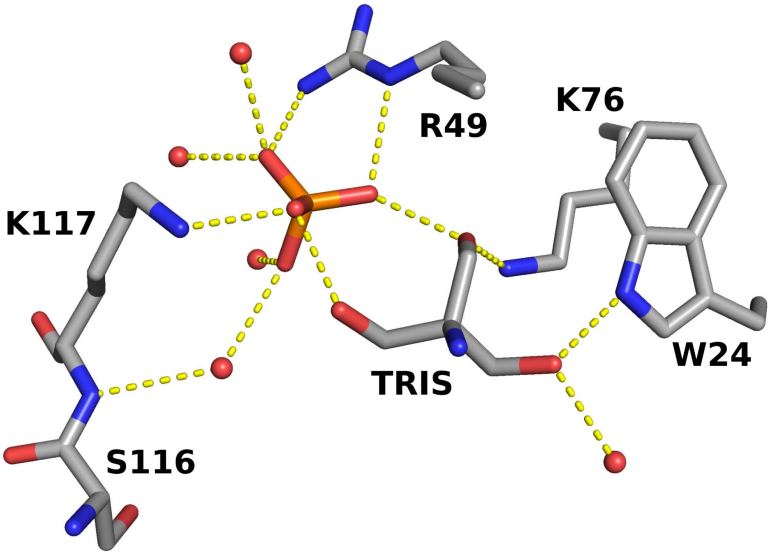


Figure S11: The *distal* phosphate binding site of the βPGM$_{\mathrm{T16A}}$:Pi complex, with Pi and TRIS bound in the the active site of the enzyme. Selected active site residues and ligands are shown as sticks using CPK coloring with O=red, N=blue, P=orange, Mg=green, while water molecules are displayed as red spheres. Hydrogen bonds are shown with yellow dashed lines.

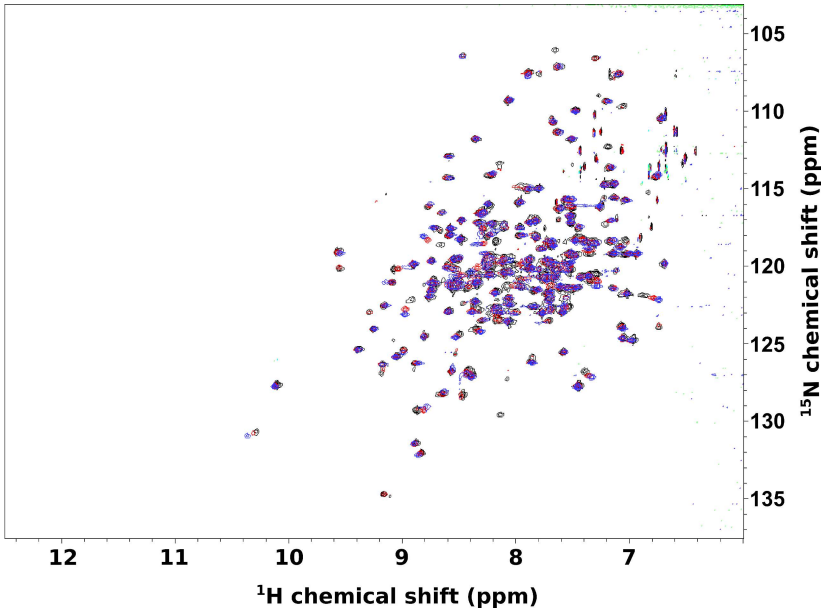


Figure S12: 2D $^{1}H^{15}N$-TROSY NMR spectra of 1mM apo-βPGM$_{WT}$ (50mM Tris pH 7.2, 5mM $MgCl_2$ 2mM NaN_3) in the presence of 0mM (black), 25mM (red), and 100mM (blue) phosphate.

5. SIIS value determination

Complex	F1 (ppm)	F2 (ppm)	F3 (ppm)	F4 (ppm)
PSP:MgF_3	-175.5	-140.7	-144.3	–
PSP:MgF_3:L-Ser	-160.6	-140.9	-147.1	–
PSP:AlF_4	-148.7	-134.1	-133.0	-140.8
PSP:AlF_4:L-Ser	-142.2	-136.3	-133.1	-141.3
PSP:BeF_3	-180.5	-147.5	-147.8	–

Table S7: ^{19}F chemical shifts of metal fluoride TSA complexes in PSP. Values are presented for the chemical shift in H_2O.

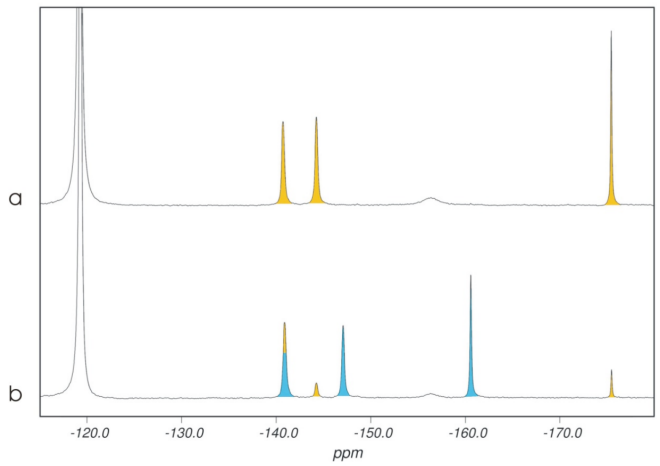


Figure S13: 1D ^{19}F NMR spectra of **(a)** PSP:$MgF_3$and **(b)** PSP:MgF_3:L-Ser. The PSP:MgF_3 complex is highlighted in yellow, while the PSP:MgF_3:L-Ser complex is highlighted in blue. The peak on the left is free fluoride in solution, whereas the broad peak at -156 ppm is MgF^+ free in solution. Chemical shifts associated with the two complexes are presented in Table S7. Figure adapted from [40].

6. Linewidth analysis and chemical exchange

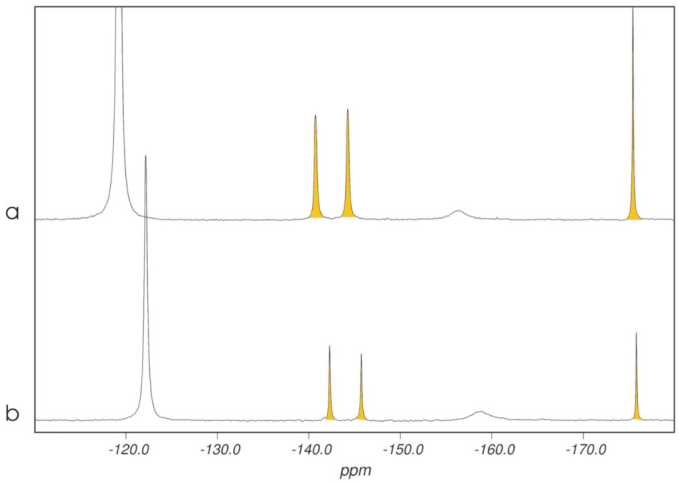


Figure S14: 1D ^{19}F NMR spectra of the PSP:MgF_3 complex (highlighted in yellow) in **(a)** 100% H_2O and **(b)** 100% D_2O. The peak on the left is free fluoride in solution, whereas the broad peak at -156 ppm is MgF^+ free in solution. Chemical shifts associated with the two complexes are presented in Table S8 as well as the resulting SIIS value. Figure adapted from [40].

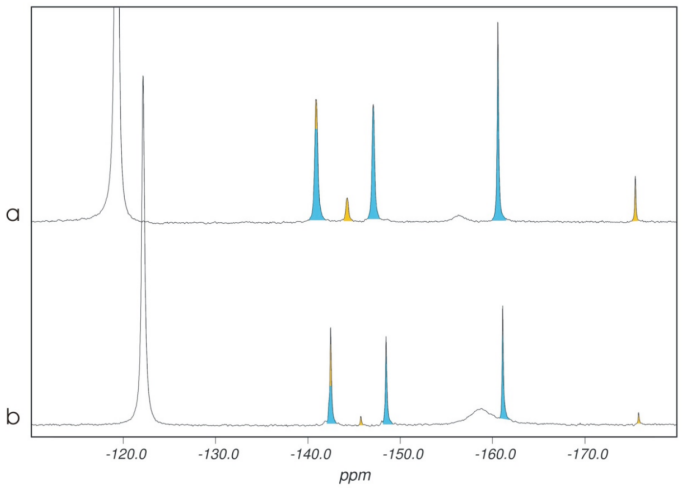


Figure S15: 1D ^{19}F NMR spectra of the PSP:MgF_3:L-Ser complex (highlighted in blue) in **(a)** 100% H_2O and **(b)** 100% D_2O. The peak on the left is free fluoride in solution, whereas the broad peak at -156 ppm is MgF^+ free in solution. Chemical shifts associated with the two complexes are presented in Table S8 as well as the resulting SIIS value. Figure adapted from [40].

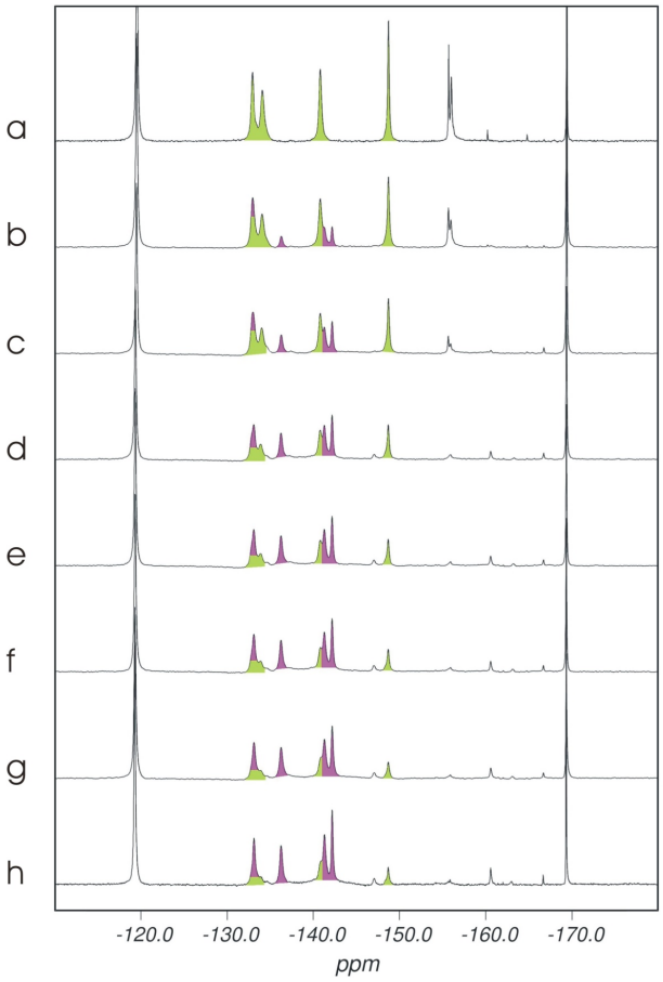


Figure S16: 1D ^{19}F NMR spectra used to determine the saturation of the PSP:AlF_4:L-Ser complex (highlighted in purple). The concentration of L-Ser was **(a)** 0 mM, **(b)** 5 mM, **(c)** 10 mM, **(d)** 20 mM, **(e)** 30 mM, **(f)** 40 mM, **(g)** 50 mM, **(h)** 70 mM. Chemical shifts associated with the two complexes are presented in Table S7. The peak at -119 ppm corresponds to free fluoride in solution and the peak at -169 is of unknown origin, but due to the narrow linewidth observed, likely corresponds to a small molecule. Figure adapted from [40].

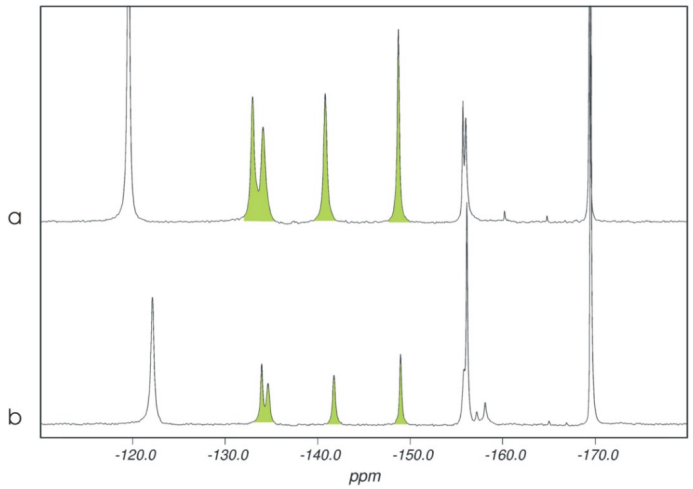


Figure S17: 1D ^{19}F NMR spectra of the PSP:AlF_4 complex (highlighted in green) in **(a)** 100% H_2O and **(b)** 100% D_2O. Chemical shifts associated with the two complexes are presented in Table S8 as well as the resulting SIIS value. Figure adapted from [40].

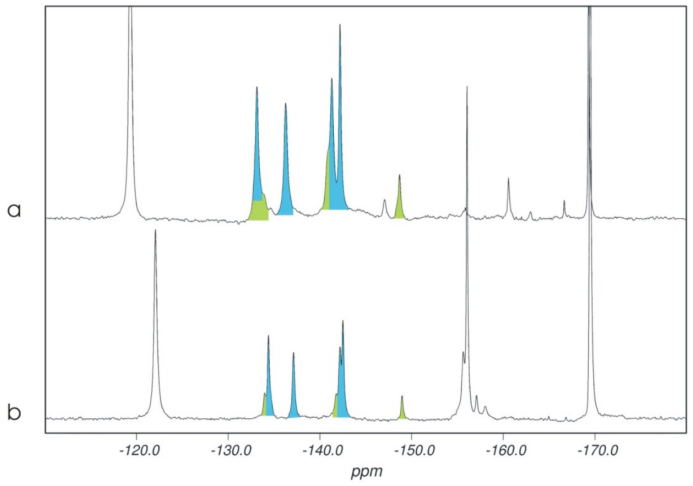


Figure S18: 1D ^{19}F NMR spectra of the PSP:AlF_4:L-Ser complex (highlighted in blue) in **(a)** 100% H_2O and **(b)** 100% D_2O. Chemical shifts associated with the two complexes are presented in Table S8 as well as the resulting SIIS value. Figure adapted from [40].

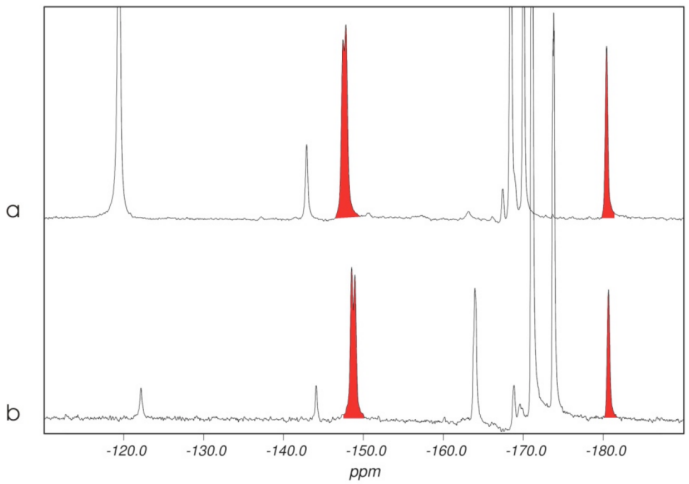


Figure S19: 1D ^{19}F NMR spectra of the PSP:BeF_3 complex (highlighted in red) in **(a)** 100% H_2O and **(b)** 100% D_2O. Chemical shifts associated with the two complexes are presented in Table S8 as well as the resulting SIIS value. Figure adapted from [40].

Fluorine nuclei	$\delta(H_2O)$, ppm	$\delta(D_2O)$, ppm	SIIS, ppm
PSP:MgF_3			
F1	-175.45	-175.83	0.38
F2	-140.71	-142.28	1.57
F3	-144.26	-145.74	1.48
PSP:MgF_3:L-Ser			
F1	-160.59	-161.11	0.52
F2	-140.89	-142.46	1.57
F3	-147.70	-148.48	1.38
PSP:AlF_4			
F1	-148.72	-148.97	0.25
F2	-134.10	-134.65	0.55
F3	-132.97	-133.96	0.99
F4	-140.82	-141.77	0.95
PSP:AlF_4:L-Ser			
F1	-142.20	-142.52	0.32
F2	-136.26	-137.12	0.86
F3	-133.14	-134.39	1.25
F4	-141.31	-142.19	0.88
PSP:BeF_3			
F1	-180.45	-180.63	0.18
F2	-147.45	-148.55	1.10
F3	-147.82	-148.94	1.12

Table S8: ^{19}F chemical shifts of metal fluoride TSA complexes in PSP. Chemical shift (δ) for the resonance in 100% H_2O and 100% D_2O are presented as well as the sum SIIS value.

Complex	F1	F2	F3
βPGM$_{WT}$:BeF_3^-	296 ± 2.1	404 ± 2.7	426 ± 3.0
βPGM$_{K145A}$:BeF_3^-	397 ± 3.3	530 ± 3.1	592 ± 2.5
βPGM$_{WT}$:MgF_3^-:G6P	102.7 ± 1.5	187 ± 3.5	199 ± 3.3
PSP$_{WT}$:BeF_3^-	115 ± 2.8	188 ± 3.8	196 ± 4.9
PSP$_{WT}$:MgF_3^-	54 ± 0.2	149 ± 1.5	213 ± 2.0
PSP$_{E20A}$:BeF_3^-	127 ± 1.6	218 ± 3.7	189 ± 2.3
PSP$_{E20A}$:MgF_3^-	80 ± 0.3	148 ± 1.6	182 ± 5.1
PSP$_{N170A}$:MgF_3^-	108 ± 0.7	184 ± 1.7	400 ± 3.5

Table S9: Linewidths of fluorine resonances in βPGM and PSP complexes with metal fluoride ground and transition state analogues. Linewidths (in Hz) of each resonance were fitted using the *dcon* tool in Topspin v.3.5 and error given to one standard deviation.

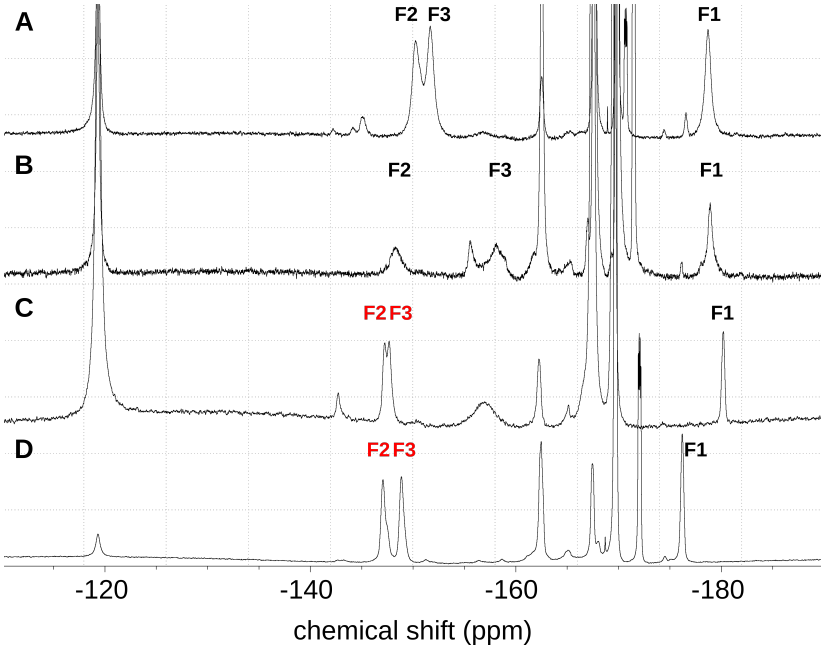


Figure S20: ^{19}F 1D NMR spectra of βPGM and PSP complexed with metal fluoride species, fluorine atoms (where known) are labeled in black while inferred assignments are labeled in red according to Fig. 2 . **A)** βPGM_{WT}:BeF_3^-, **B)** βPGM_{K145A}:BeF_3^-, **C)** PSP_{WT}:BeF_3^-, **D)** PSP_{E20A}:BeF_3^-. Peaks denoted with an asterisk correspond to MFx species free is solution (see ref. [10, 11]), while the leftmost peak corresponds to free fluoride. The chemical shifts of the fluorides (in ppm) are **A)** -178.7, -150.3, -151.7, **B)** -178.6, -148.3, -158.0, **C)** -180.4, -147.5, -147.9, and **D)** -176.4, -147.3, -149.1, for fluorides F1, F2 and F3 respectively in each of the complexes.

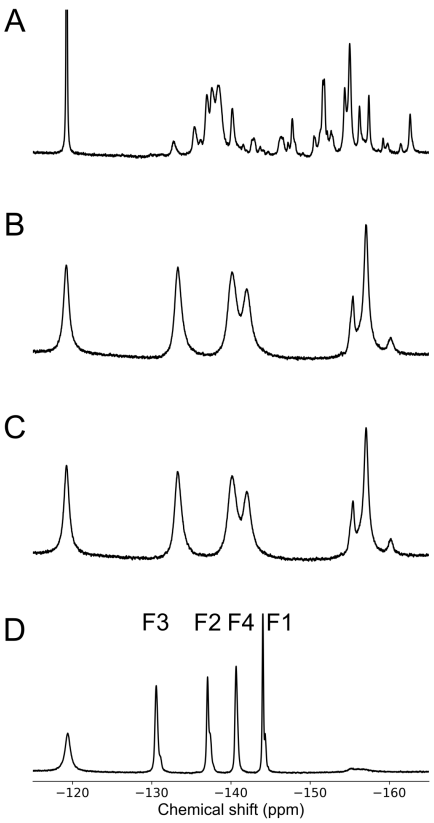


Figure S21: ^{19}F 1D NMR spectra of βPGM in standard NMR buffer, complexed with 5 mM AlF_4, **A)** in the absence of Mg^{2+}, **B)** in the presence of 5mM $MgCl_2$, **C)** in the presence of 5mM $MgCl_2$ and 10mM glucose, and **D)** in the presence of 5mM $MgCl_2$ and 10mM G6P. The addition of Mg^{2+} to an βPGM:AlF_4 complex greatly simplifies the ^{19}F spectra, indicating that the formation of the βPGM:AlF_4 complex is Mg^{2+} dependent and that the AlF_4 species is interacting with the catalytic Mg^{2+} ion in the *proximal* site. Addition of glucose does not change the chemical shift of AlF_4 bound species, while addition of G6P forms a stable complex with a much narrower ^{19}F linewidth.

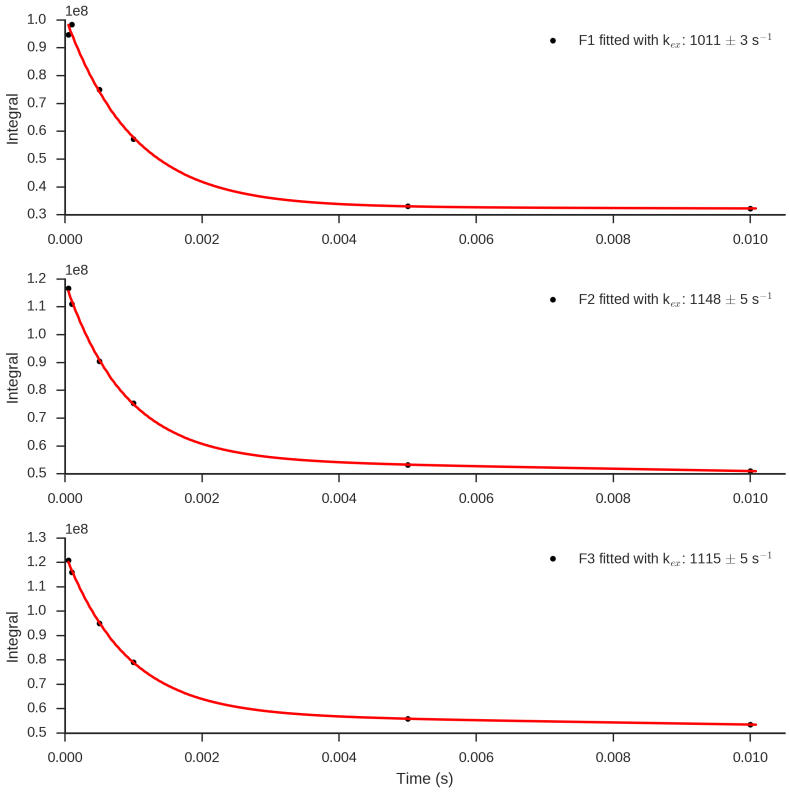


Figure S22: Integral vs. mixing time for the diagonal peaks that correspond to each fluoride position in 2D ^{19}F-EXSY spectra of the βPGM$_{WT}$:BeF_3^- complex. Exchange was fitted using a python least squares minimisation algorithm, and the error of the fit is presented at one standard deviation.

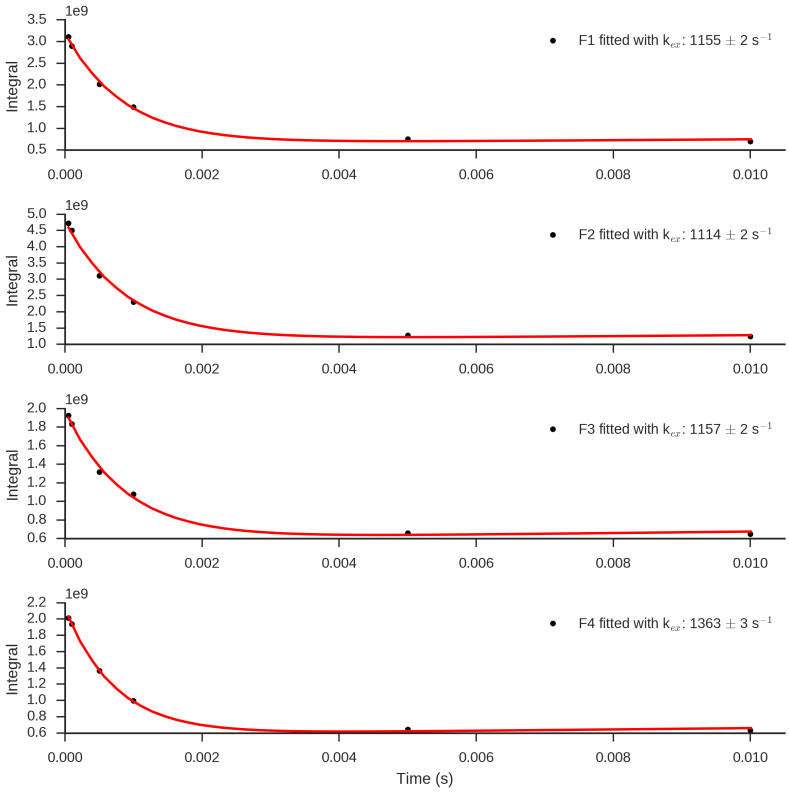


Figure S23: Integral vs. mixing time for the diagonal peaks that correspond to each fluoride position in 2D ^{19}F-EXSY spectra of the βPGM$_{WT}$:AlF$_4^-$ complex. Exchange was fitted using a python least squares minimisation algorithm, and the error of the fit is presented at one standard deviation.

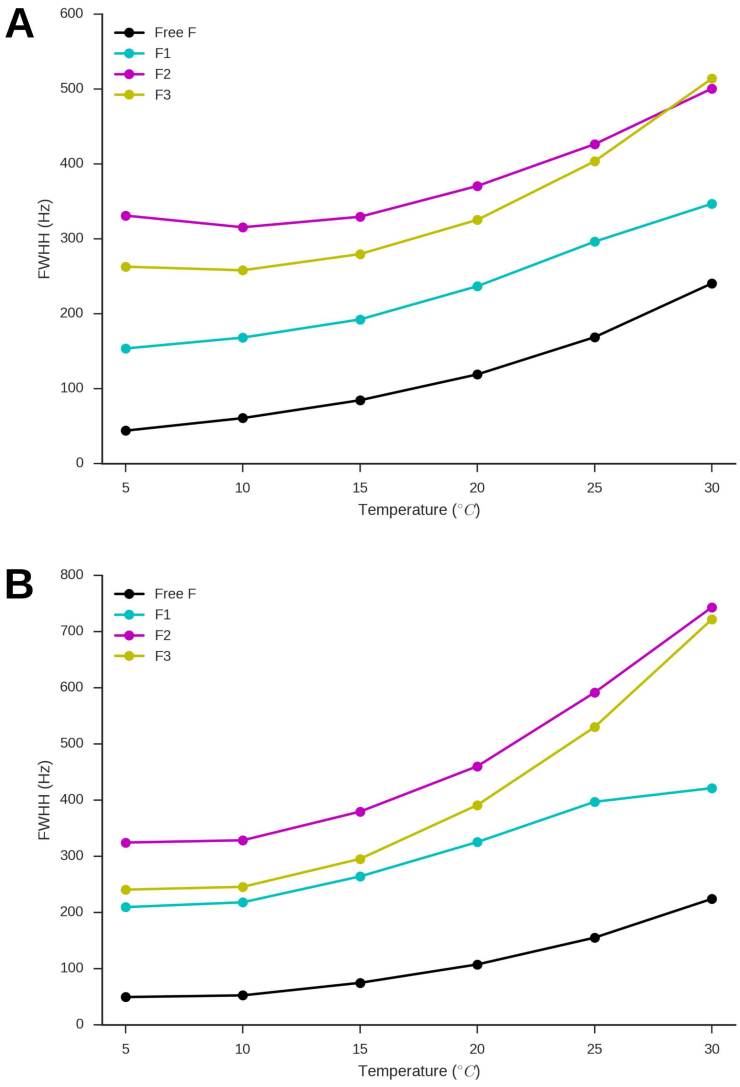


Figure S24: The temperature dependence of fluorine linewidth at half height (FWHH, in Hz) for BeF_3^- fluorides in **A)** βPGM_{WT}:BeF_3^- and **B** βPGM_{K145A}:BeF_3^-complexes.

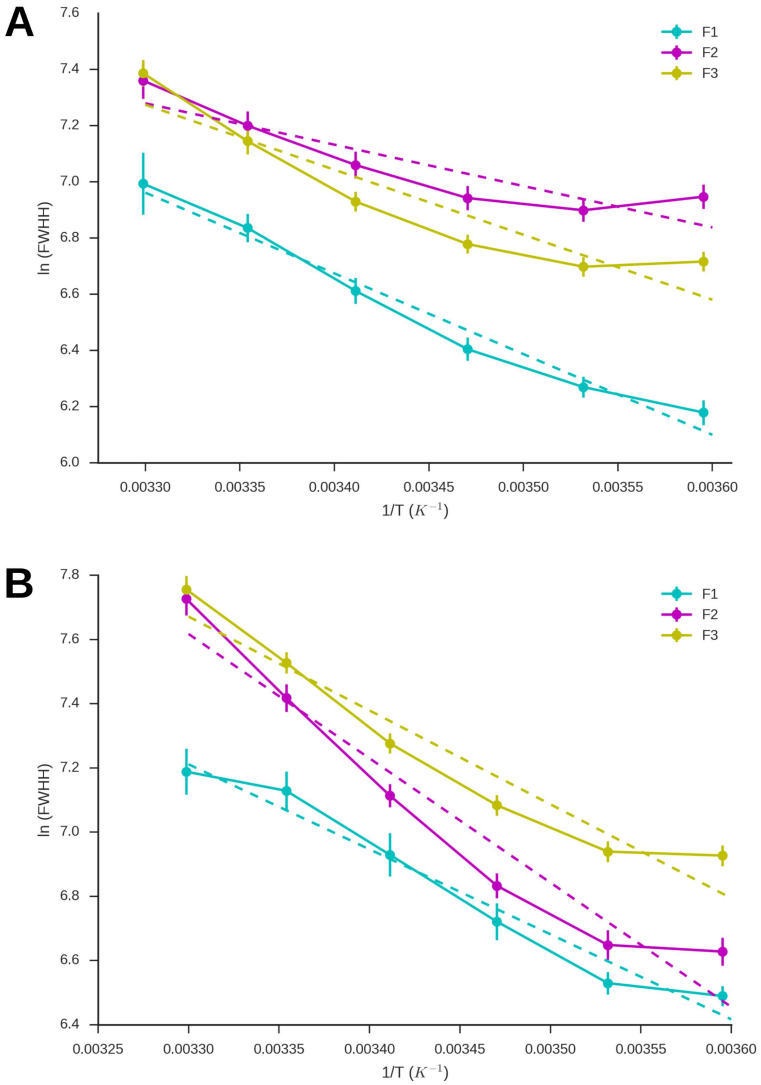


Figure S25: The inverse temperature dependence of the natural logarithm of fluorine linewidth at half height (FWHH, in Hz) for BeF_3^- fluorides in **A)** βPGM_{WT}:BeF_3^- and **B** βPGM_{K145A}:BeF_3^-complexes. Data were fit to a linearized Arrhenius equation $ln(k) = \frac{-E_a}{R} \cdot \frac{1}{T} + ln(A)$, where R is the gas constant, T is the temperature in Kelvin, A is the pre-exponential factor, and E_a is the activation energy.

Appendix B

Appendix

B.1 Protein purification

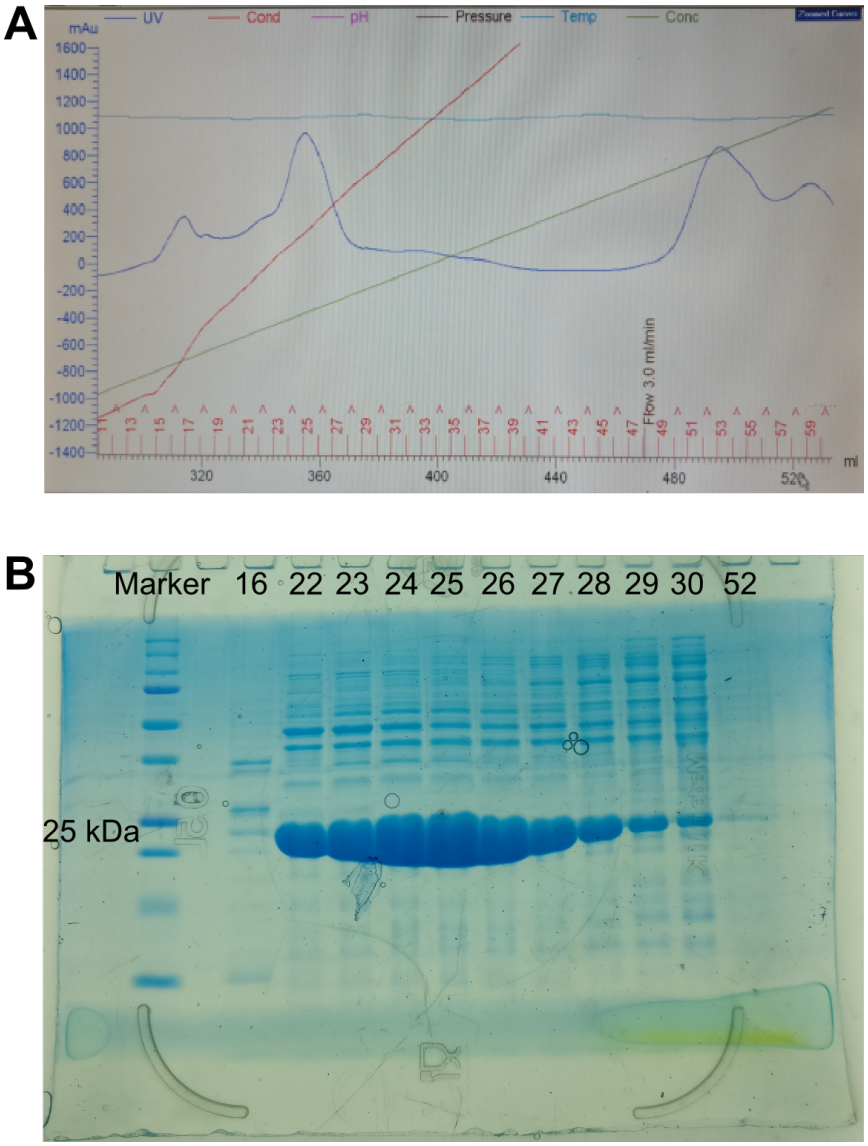


FIGURE B.1: DEAE sephadex ion exchange chromatography step of βPGM purification. **A)** UV readout from AKTAprime with fraction numbers illustrated at the bottom. **B)** PAGE-gel, 10 μl loaded of fraction diluted by a factor of 0.75 (addition of 4x SDS loading buffer) and run at 50 V for 10 min, then 180 V for 50 min. The gel was stained with Coomasie InstantBlue and incubated overnight, numbers denote fractions loaded.

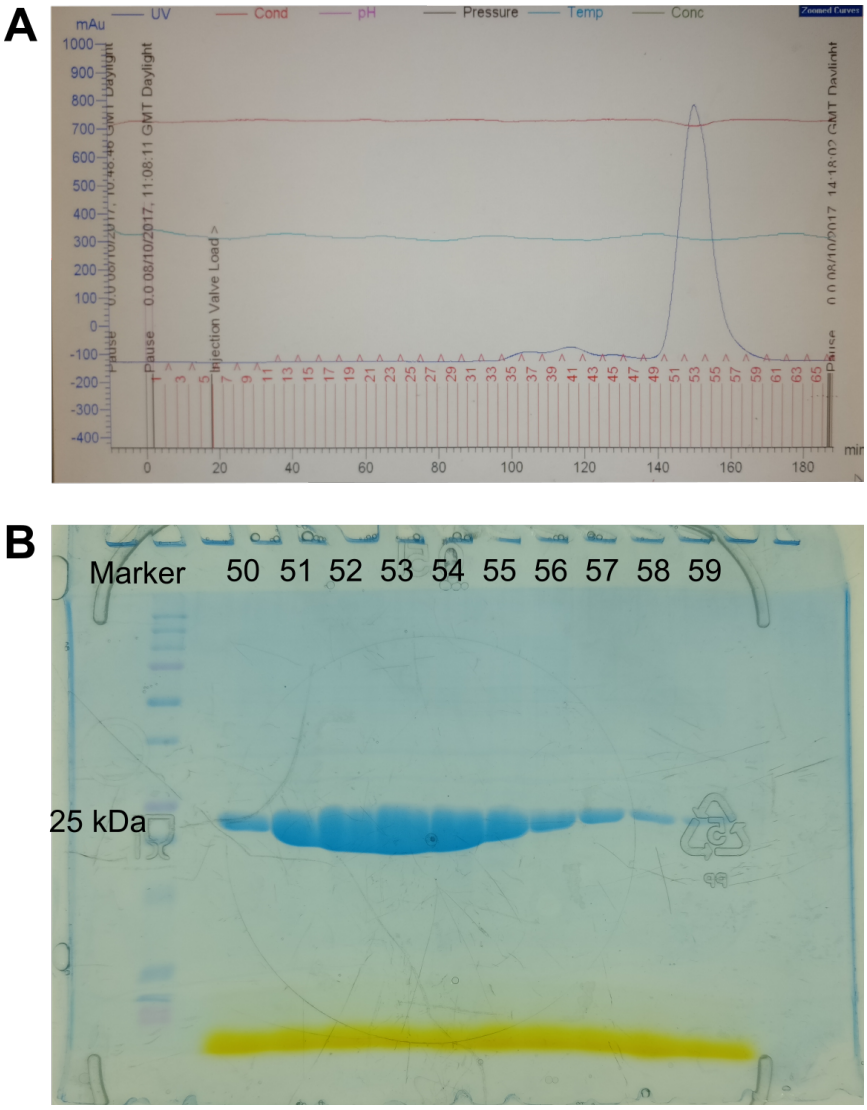


FIGURE B.2: Sephadex G75 chromatography step of βPGM purification. **A)** UV readout from AKTAprime with fraction numbers illustrated at the bottom. **B)** PAGE-gel, 10 μl loaded of fraction diluted by a factor of 0.75 (addition of 4x SDS loading buffer) and run at 50 V for 10 min, then 180 V for 50 min. The gel was stained with Coomasie InstantBlue and incubated overnight, numbers denote fractions loaded.

B.2 Backbone relaxation macros

Many of these scripts are for data extraction, plotting, and comparison. Code can be made available upon request. This master shell script runs each of the analyses in order and was designed to be a semi-automated pipeline where HydroNMR is automatically incorporated.

B.2.1 ajr-wonder-macro-for-relaxation-analysis.sh

```
#!/bin/csh

# wonder macro for relaxation analysis

#-------------------------------------------------
echo ""
echo "-------------------------------------------------"
echo "Making SS file ... "
echo "-------------------------------------------------"
# Make ss file  (soon to be redundant)
ajr-ssFromPDB.py 2wf5.pdb



#-------------------------------------------------
echo ""
echo "-------------------------------------------------"
echo "Plotting T1, T2, NOE values ... "
echo "-------------------------------------------------"
# Collect T1, T2 and NOE values together and plot.
# Outputs PDB files with b-factors
# Outputs scatter plots with secondary structures
# Outputs relax formatted results file
# Outputs text file for hammock plotting


BBRel_plot2.py expfit.txt R2.txt noe.txt 2wf5.pdb ajr_ss_file.txt



#-------------------------------------------------
echo ""
echo "-------------------------------------------------"
echo "Plotting Hammock plots from acquired data ..."
echo "-------------------------------------------------"
# Plots hammock plot of T1 vs T2
# Plots hammock plot of T1 vs NOE
ajr-hammock3.py ajr_T1_T2_HetNOE.txt -s "WT AlF4 G6P TSA" 18.8 3 27 3 "T1vT2_Hammock" "T1vHetNOE_Hammock"



#-------------------------------------------------
```

```
echo ""
echo "-------------------------------------------------"
echo "Running HydroNMR, Have you edited the hydronmr.dat file? ... "
echo "-------------------------------------------------"
# NOTE!!!!
# Go and edit the hydronmr.dat file with correct field strength in T.

# Runs hydroNMR taking the hydronmr.dat as direction.
# Outputs pgmutase.XXXX.res
#------------------------------ uncomment as appropriate ---------------------
# cp hydronmr800.dat hydronmr.dat
# cp hydronmr600.dat hydronmr.dat
# cp hydronmr500.dat hydronmr.dat

hydronmr7c2lnx.exe


#-------------------------------------------------
echo ""
echo "-------------------------------------------------"
echo "Plotting T1, T2, NOE values generated from HydroNMR"
echo "-------------------------------------------------"
# Extracts T1, T2, NOE data from XXXX.res file generated by HydroNMR
# Outputs PDB files with b-factors
# Outputs scatter plots with secondary structures
# Outputs text file for hammock plotting

ajr-HydroNMR_XP.py pgmutase.800.res 2wf5.pdb


#-------------------------------------------------
echo ""
echo "-------------------------------------------------"
echo "Plotting Hammock plots from HydroNMR run ..."
echo "-------------------------------------------------"
# Plots hammock plot of T1 vs T2
# Plots hammock plot of T1 vs NOE

ajr-hammock3.py ajr_HydroNMR_T1_T2_HetNOE.txt -s "WT AlF4 G6P TSA" 18.8 3 27 3
 "HydroNMR_T1vT2_Hammock" "HydroNMR_T1vHetNOE_Hammock"


#-------------------------------------------------
echo ""
echo "-------------------------------------------------"
echo "Plotting T1, T2, NOE values compared to HydroNMR ... "
```

```
echo "-------------------------------------------------"
# Collect T1, T2 and NOE values together and compares to HydroNMR.
# Outputs PDB files with b-factors
# Outputs scatter plots with secondary structures

ajr-BBR_HNMR_comparison.py ajr_T1_T2_HetNOE.txt ajr_HydroNMR_T1_T2_HetNOE.txt 2wf5.pdb


#-------------------------------------------------
echo ""
echo "-------------------------------------------------"
echo "Moving png files to /png_output "
echo "-------------------------------------------------"
mkdir png_output
mv *.png png_output


#-------------------------------------------------
echo ""
echo "-------------------------------------------------"
echo "Moving PDB files to /PDB_output "
echo "-------------------------------------------------"
mkdir PDB_output
mv *_*.pdb PDB_output

#-------------------------------------------------
echo ""
echo "-------------------------------------------------"
echo "Making relax input pdb file"
echo "MAKE SURE THAT THE PDB FILE HAS HYDROGENS ADDED"
echo "-------------------------------------------------"

# eg.
# pymol 2wfa.pdb
# action - add hydrogens - save molecule as 2wfa_H.pdb

nawk '$3=="N" || $3=="H02" {print $0}' 2wfa_H.pdb | nawk '$1=="ATOM" {print $0}' |\
nawk '!($3=="H02" && $4=="PRO") {print $0}' |\
nawk '!($4=="HOH") {print $0}' > pgmutase_N_H.pdb
```

B.2.2 ajr-hammock3.py

This program uses the original LS MF equation to predict order parameters and correlation times.

```
#!/usr/bin/python

# Incorporating Jeremy's relvals script
# Will plot T1 and T2 values



import math
import matplotlib.pyplot as plt
import sys
import os

#----------------------------Usage information at cmd line----------------------------------
if len(sys.argv) < 3:
        print "Program to plot hammock plot in matplotlib"
        print "Written by Angus Robertson Jun 2017 \n"
        print "Usage:"
        print "ajr_hammock.py t1_t2_fileajr_T1_T2_HetNOE.txt _s\n"
        print "\nUsage options:\n_c for command line "
        print "_s for scripting mode (most common)"
        print "Scripting mode Takes files with format \nAss T1 T1err T2 T2err NOE NOEerr\n"
        sys.exit()

#-------------------------------------------------------------------------------------------



if sys.argv[2]=="_c":
        ##################################
        #Set exp param


        Title="WT AlF4 G6P complex"
        B0=18.8
        tm_low=3
        tm_hi=27
        tm_increment=3
        s2range = [0.5, 0.6, 0.7, 0.8, 0.9, 1.0]
        file1_title="T1vT2_Hammock"
        file2_title="T1vHetNOE_Hammock"

elif sys.argv[2]=="_s":
        Title=sys.argv[3]
        B0=float(sys.argv[4])
        tm_low=int(sys.argv[5])
```

```
        tm_hi=int(sys.argv[6])
        tm_increment=int(sys.argv[7])
        s2range = [0.5, 0.6, 0.7, 0.8, 0.9, 1.0]
        file1_title=sys.argv[8]
        file2_title=sys.argv[9]


# TAKE CARE: Make sure that the B0 (T) is set appropriately
#
#
# NB: Increasing tm extends x axis
# increasing s2 extends the y axis


###################################


#

pi=3.1416
h=6.63E-34
u0=1.26E-6
rnh=1.02E-10
yn=-4.31e+6*2.0*pi
yh=4.26e+7*2.0*pi
r2a=0.0
d=((u0*h)/(8*(pi*pi)))*yn*yh*(1/(rnh*rnh*rnh))


def srj(tm, te, s, w):
        rj=((s*s)*tm)/(1+(w*w*tm*tm))
        t=1.0/((1.0/tm)+(1.0/te))
        rj=rj+((1-(s*s))*t)/(1+(w*w*t*t))
        rj=rj*0.4
        return rj


def sr1(tm, te, s, d, wh, wn, c):
        r1=srj(tm,te,s,wh-wn)
        r1=r1+(3.0*srj(tm,te,s,wn))+(6.0*srj(tm,te,s,wh+wn))
        r1=r1*((d*d)/4.0)
        r1=r1+((c*c)*srj(tm,te,s,wn))
        return r1


def sr2(tm, te, s, d, wh, wn, c, rr2):
```

```
        r2=(4.0*srj(tm,te,s,0))+srj(tm,te,s,wh-wn)
        r2=r2+(3.0*srj(tm,te,s,wn))+(6.0*srj(tm,te,s,wh))
        r2=r2+(6.0*srj(tm,te,s,wh+wn))
        r2=r2*((d*d)/8.0)
        r2=r2+((((c*c)/6.0)*(3.0*srj(tm,te,s,wn)+(4.0*srj(tm,te,s,0))))+rr2
        return r2


def snoe(tm, te, s, d, rr1, yh, yn, wh, wn):
        noe=(6.0*srj(tm,te,s,wh+wn))-srj(tm,te,s,wh-wn)
        noe=1.0+(noe*(((d*d)/(4.0*rr1))*(yh/yn)))
        return noe


def relvals(tm, s2, te, b0):
        if (b0==""):
                b0=11.744
        wn=yn*b0
        wh=yh*b0
        c=160
        c=wn*((c*1e-6)/(math.sqrt(3.0)))
        te=te*1e-12
        tm=tm*1e-9
        s=math.sqrt(s2)
        # Calculating R1, R2 and NOE
        r1=sr1(tm, te, s, d, wh, wn, c)
        r2=sr2(tm, te, s, d, wh, wn, c, r2a)
        noe=snoe(tm, te, s, d, r1, yh, yn, wh, wn)

        return tm*1e+9,te*1e+12,s2,1000/r1,1000/r2,noe

#-------------------------------------------------------------------------
# plotting scripts

def draw_hammock_s2(s2, tml, tmh, tmi, b0):
        holder=[]
        for i in range(tml, tmh, tmi):
                # relvals(tm, s2, te, b0)
                a=relvals(i, s2, 40, b0)
                holder.append([a[3], a[4]])
        x_val = [x[0] for x in holder]
        y_val = [x[1] for x in holder]
        plt.plot(x_val, y_val, 'k-')
        label="S2="+str(s2)
        plt.annotate(label,xy=holder[-0], xytext=(0, 0), textcoords='offset points', ha='right', va='bottom')
```

```
def draw_hammock_tm(tm, s2_range_lis, b0):
        holder=[]
        for i in s2_range_lis:
                # relvals(tm, s2, te, b0)
                a=relvals(tm, i, 40, b0)
                holder.append([a[3], a[4]])
        x_val = [x[0] for x in holder]
        y_val = [x[1] for x in holder]
        plt.plot(x_val, y_val, 'k-')
        label="tm="+str(tm)
        plt.annotate(label,xy=holder[-0], xytext=(0, 0), textcoords='offset points', ha='left', va='bottom', rotation
```

 # plotting scripts for HetNOE hammock

```
def draw_hammock_s2_HetNOE(s2, tml, tmh, tmi, b0):
        holder=[]
        for i in range(tml, tmh, tmi):
                # relvals(tm, s2, te, b0)
                a=relvals(i, s2, 40, b0)
                holder.append([a[3], a[5]])
        x_val = [x[0] for x in holder]
        y_val = [x[1] for x in holder]
        plt.plot(x_val, y_val, 'k-')
        label="S2="+str(s2)
        plt.annotate(label,xy=holder[-1], xytext=(0, 0), textcoords='offset points', ha='left', va='bottom')

def draw_hammock_tm_HetNOE(tm, s2_range_lis, b0):
        holder=[]
        for i in s2_range_lis:
                # relvals(tm, s2, te, b0)
                a=relvals(tm, i, 40, b0)
                holder.append([a[3], a[5]])
        x_val = [x[0] for x in holder]
        y_val = [x[1] for x in holder]
        plt.plot(x_val, y_val, 'k-')
        label="tm="+str(tm)
        plt.annotate(label,xy=holder[0], xytext=(0, -15), textcoords='offset points', ha='right', va='bottom')
```

```
################################################################################
```

```
if sys.argv[2]=="-c":
        # read in file
```

```
t1t2_file=sys.argv[1]


with open(t1t2_file, 'r') as fh:
        f=fh.readlines()

# check that assignment is identical in col 1 and 3:
for line in f:
        a=line.split()[0]
        b=line.split()[2]
        if not a==b:
                print "\nThere appears so be an assignment discrepancy"
                print "\nMismatched line"
                print line
                print "System exiting"
                sys.exit()

# Set up hammock plot
fig = plt.figure()
ax = fig.add_subplot(111)

# plotting hammock
for val in s2range:
        draw_hammock_s2(val, tm_low, tm_hi, tm_increment, B0)


for val in range(tm_low, tm_hi, tm_increment):
        draw_hammock_tm(val, s2range, B0)



# read T1 and T2 into lists


t1t2list=[]


for line in f:
        a=line.split()[1]
        b=line.split()[3]
        t1t2list.append([a, b])


x_val = [x[0] for x in t1t2list]
y_val = [x[1] for x in t1t2list]

# plot t1 t2 scatter


plt.scatter(x_val, y_val, marker='o', c='b', s=2)


plt.title(Title)
plt.xlabel('T1 (ms)')
```

```
        plt.ylabel('T2 (ms)')

        savefile="T1_T2_Hammock_"+str(B0)+"T.png"
        plt.savefig(savefile)
        plt.show()




elif sys.argv[2]=="-s":
        # read in file
        parameter_file=sys.argv[1]

        with open(parameter_file, 'r') as fh:
                f=fh.readlines()


        # Set up hammock plot
        fig = plt.figure()
        ax = fig.add_subplot(111)

        # plotting hammock
        for val in s2range:
                draw_hammock_s2(val, tm_low, tm_hi, tm_increment, B0)

        for val in range(tm_low, tm_hi, tm_increment):
                draw_hammock_tm(val, s2range, B0)

        # read T1 and T2 into lists

        t1t2list=[]
        noelist=[]

        for line in f:
                t1=line.split()[1]
                t2=line.split()[3]
                noe_val=line.split()[5]
                t1t2list.append([t1, t2])
                noelist.append([t1, noe_val])

        x_val = [x[0] for x in t1t2list]
        y_val = [x[1] for x in t1t2list]



        # plot t1 t2 scatter
```

FIGURE B.3: T1 vs T2 hammock generated using the original LS MF equation to predict correlation times (ns) and order parameters.

```
plt.scatter(x_val, y_val, marker='o', c='b', s=2)


plt.title(Title)
plt.xlabel('T1 (ms)')
plt.ylabel('T2 (ms)')


savefile=file1_title+"_"+str(B0)+"T.png"
plt.savefig(savefile)
#plt.show()



# Set up hammock plot for HetNOE vs T1
fig = plt.figure()
ax = fig.add_subplot(111)

# plotting hammock
for val in s2range:
        draw_hammock_s2_HetNOE(val, tm_low, tm_hi, tm_increment, B0)


for val in range(tm_low, tm_hi, tm_increment):
        draw_hammock_tm_HetNOE(val, s2range, B0)


x_val_noe = [x[0] for x in noelist]
y_val_noe = [x[1] for x in noelist]


plt.scatter(x_val_noe, y_val_noe, marker='o', c='b', s=2)


plt.title(Title)
plt.xlabel('T1 (ms)')
plt.ylabel('HetNOE')


savefile=file2_title+"_"+str(B0)+"T.png"
plt.savefig(savefile)
#plt.show()
```

Example output

FIGURE B.4: T1 vs HetNOE hammock generated using the original LS MF equation to predict correlation times (ns) and order parameters.

B.3 Kinetic characterization

B.3.1 PANDAlyze: from spreadsheet to text files

This program uses the python package *pandas* to do zeroing, path length correction, and conversion to G6P concentration, on a spreadsheet, reducing the risk of manual error and saving significant time. Thanks to Henry Wood in the development for some very useful proofreading.

```
# -*- coding: utf-8 -*-
"""
Created on 13 Jun 18 by AJR

This script will read a CSV with coupled assay output and output dynafit readable
files

@coauthor Angus Robertson
@coauthor: Chris Sharratt
"""
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from matplotlib.pyplot import cm

current_dir="/home/gus/pgmutase/kcat_determination/R49K_catalysis/PANDAlyse"
data_dir=current_dir
CSV_file="290618_r49k_g1p_edited.csv"

Blank_ref='B02'      # Change this to a string of the id used.
columnnames = ['well', 'times', 'B02', 'C02', 'C03', 'C04', 'C05', 'C06', 'C07', 'C08', 'C09', 'C10', \
    'D02', 'D03', 'D04', 'D05', 'D06', 'D07', 'D08', 'D09', 'D10', \
    'E02', 'E03', 'E04', 'E05', 'E06', 'E07', 'E08', 'E09', 'E10']

concentrations = [0,5,15,35,50,70,100,160,230,330,5,15,35,50,70,100,160,230,330,5,15,35,50,70,100,160,230,330]
csv_out = "no"      # Change to "yes" if you want full CSV output.


#%%
#-------------------------------------------------------------------------------
# This generates a (file-name, concentration) list of tuples for curve display and
# naming in the Dynafit data script.

nc_coup=[]
counter=0
for i in columnnames[2:]:
    nc_coup.append((concentrations[counter], i))
    counter+=1
```

```
nc_sorted = sorted(nc_coup, key=lambda x: x[0])


#=========================================================================
# Here we make the dataframe and do multiplicaitons
#=========================================================================
#This time we try reading it, but NOT using the well in the index.
# Should now be able to match on the index field.

df1 = pd.read_csv(CSV_file, header=7, index_col = 'times', names = columnnames )

#Note the Syntax, [RowFrom:RowTo, ColumnFrom, ColumnTo],
# So specifying a specific unrecognised column in the range means make me a new column.

print('Columns to calc are: ')
print(columnnames[2:])

for colname in columnnames[2:]:

    #df1.loc[df1['well'] == 'Raw Data (340 1)', (colname + 'Corr')] = df1.loc[df1['well']
    #== 'Raw Data (340 1)'][colname] - df1.loc[df1['well'] == 'Raw Data (340 1)']['C03']

    # Eq: c = A / l e    so.
    # Eq. used:    (OD_340_sample - OD_340_blank) / ((OD_977 - OD_900) / 0.183) * 6220

    df1.loc[df1['well'] == 'Raw Data (340 1)', (colname + 'Corr')] = df1.loc[df1['well']
    == 'Raw Data (340 1)'][colname] - df1.loc[df1['well'] == 'Raw Data (340 1)'][Blank_ref]

    df1.loc[df1['well'] == 'Raw Data (340 1)', (colname + 'GusCalc')] = df1.loc[df1['well']
    == 'Raw Data (340 1)', (colname + 'Corr')] / (((df1.loc[df1['well'] == 'Raw Data (977 3)'][colname] -
    df1.loc[df1['well'] == 'Raw Data (900 2)'][colname])/0.183)*6220)

    #******** Added after I finished the script to write all of the files independently. **********

    df2 =   df1[['well', colname, colname + 'Corr', colname + 'GusCalc']]

    if csv_out=="yes":
        df2.loc[df2['well'] == 'Raw Data (340 1)'].to_csv(colname+'_verbose.csv')


#%%
#=========================================================================
# This outputs just the corrected value vs. time for each column
#=========================================================================
```

```
if csv_out=="yes":
    for colname in columnnames[2:]:
        df2 = df1[[colname + 'GusCalc']]
        df2 = df2.dropna()
        df2.to_csv(colname+'.csv')

for colname in columnnames[2:]:
    df2 = df1[[colname + 'GusCalc']]
    df2 = df2.dropna()
    df2.to_csv(colname+'.txt', sep=' ')


#%%

#============================================================================
# This plots each of the reactions with color order
#============================================================================

color=iter(cm.rainbow(np.linspace(0,1,len(columnnames)-2)))
fig=plt.figure(figsize=(16,16))
plt.title("Turnover series:%s \nPath to file:%s" % (CSV_file, current_dir))

for i in nc_sorted:
    c=next(color)
    a = df1[[i[1]+'GusCalc']]
    plt.plot(a[i[1]+'GusCalc'], label=str(i[0])+" - "+i[1], c=c)
    plt.xlabel("Time /s")
    plt.ylabel("Concentration of G6P / M")
plt.legend(bbox_to_anchor=(1,1), loc="upper left")
fig.subplots_adjust(right=0.85)
plt.savefig("All_data_plotted.png", dpi=300)


#%% This section plots all turnovers separately on the same figure.

#============================================================================
# tot = len(columnnames[2:])
# matsizex=4
# matsizey = math.ceil(tot/4.0)
# counter=1
#
# fig=plt.figure(figsize=(32,32))
# plt.title("Turnover series:%s \nPath to file:%s" % (CSV_file, current_dir))
# for colname in columnnames[2:]:
#     ax = fig.add_subplot(matsizey,matsizex,counter)
#     a = df1[[colname+'GusCalc']]
```

```
#       ax.plot(a[colname+'GusCalc'], label=colname)
#       plt.legend(loc='upper left')
#       plt.xlabel("Time /s")
#       plt.ylabel("Absorbance")
#       counter+=1
# plt.savefig("plotted_separately.png", dpi=300)
#=============================================================

#%%

color=iter(cm.rainbow(np.linspace(0,1,len(columnnames)-2)))

fig=plt.figure(figsize=(16,16))
plt.title("Turnover series:%s \nPath to file:%s" % (CSV_file, current_dir))
for colname in columnnames[2:]:
    c=next(color)
    ax = fig.add_subplot(111)
    a = df1[[colname+'GusCalc']]
    ax.plot(a[colname+'GusCalc'], label=colname, c=c)
    plt.xlabel("Time /s")
    plt.ylabel("Concentration of G6P / M")
plt.legend(bbox_to_anchor=(1,1), loc="upper left")
fig.subplots_adjust(right=0.85)
plt.savefig("All_data_plotted_bywell.png", dpi=300)

#%%

#=============================================================
# Output dynafit text for script correlating concentration and filename.
#=============================================================

DF_text=open("DynaFit_text.txt", 'w')
DF_text.write("[data]\n")
DF_text.write("directory %s \n" % (data_dir))
DF_text.write("extension txt\n\n")

for i in nc_sorted:
    DF_text.write("file %s | concentration G1P = %s \n" % (i[1], i[0]))
```

Example output

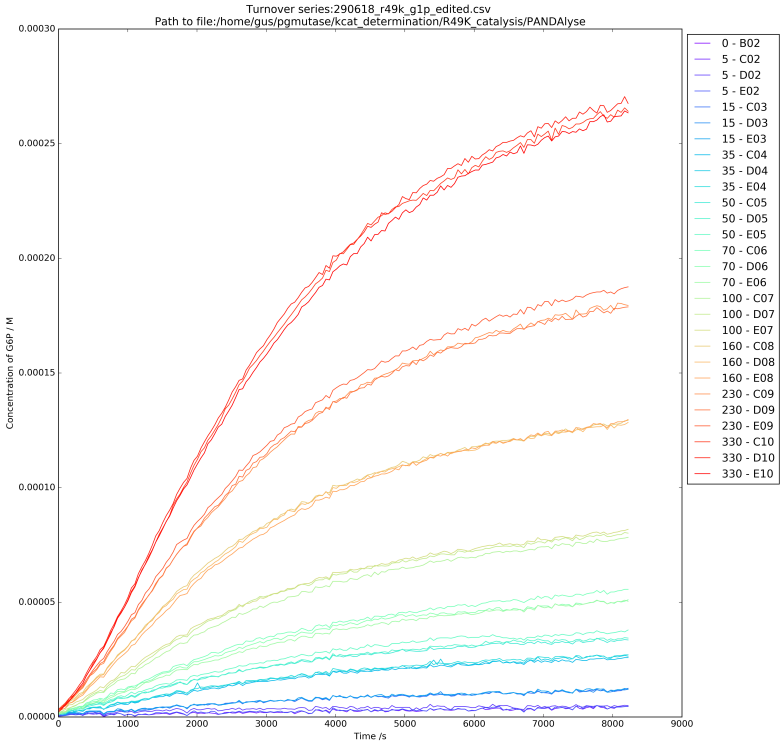


FIGURE B.5: The concentration of G6P (M) vs. time (s) for each well in the plate. Rainbow spectrum used to illustrate concentraion, with well IDs in the key.

B.3.2 FITalyze: from text files to k_{cat} and Km values

This program takes the output from PANDAlyze and allows you to iterate through each well, defining the linear portion of the reaction curve (with linear regression analysis), and ultimately using all of the rates determined in a Michaelis-Menten analysis of k_{cat} and Km. Errors were estimated using a bootstrap resampling protocol (see code) and again, thanks go to Henry Wood for some very useful proofreading and discourse.

```
#!/usr/bin/python

"""
Created on 25th Jun 2018

This script will read input files from PANDAlyze and output:
Plot of whole region.
Plot and fit of selected region.
Fit statistics for region.
Kcat determined given starting concentrations of reactants.

IMPORTANT NOTES:
    The G6P concentrations from the input files are formatted with time (s) and conc (M)
```

The Enzyme concentration must be in molar to get an accurate kcat!

 # Updates
28/06/18 – Fixed memory bug on generation of lots of figures, and added silent mode using plt.ioff().

@author: gus
"""

```
import sys, os, shutil, time, math
from scipy import stats
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
from scipy.optimize import curve_fit
from scipy import optimize
import gc

file_version_no=1.0

#=========================================================================
# Dataseries specific parameters
#=========================================================================

enz_conc=60E-9          # Conc in M!!!!

# Data time range to fit.
lo_val=0       #10000
hi_val=50000

# Initial estimates for Bootstrap fitting.
pstart = [20.0, 1.0]          # [ Km , Vmax ]
err_stdev = 0.2

# Experiment numbers to integrate, by populting the expno_to_exclude list with valid integers,
# you can exclude certain experiments.
Blank_ref='E05'      # Change this to a string of the well that contains the blank.
columnnames = ['well', 'times', 'C06', 'C07', 'C08', 'C09', 'C10', 'D06', 'D07',
'D08', 'D09', 'E05', 'E06','E07', 'E08', 'E09', 'F05', 'F06', 'F07', 'F08', 'F09']

concentrations = [5,35,70,160,330,15,50,100,230,0,15,50,100,230,5,35,70,160,330]

# Plotting and output parameters
savefig="true"  # this shoud be either "true", "show", "false"
log_file="linfit_output.txt"
```

```
#%%
#============================================================================
#============================================================================
#============================================================================
# Defining funcitons
#============================================================================
def abline(slope, intercept, pltcolor='blue'):
    """Plot a line from slope and intercept"""
    axes = plt.gca()
    x_vals = np.array(axes.get_xlim())
    y_vals = intercept + slope * x_vals
    plt.plot(x_vals, y_vals, '--', color=pltcolor, label="LOBF")


def eqn(m,x,c):
    temp=[]
    for i in x:
        y = m*i + c
        temp.append(y)
    return temp

#============================================================================
# Opening Lists etc.
#============================================================================

def calc_kcat(text_file, lo_val, hi_val, enz_conc, substrate_conc, kcat_list, show=False):
    """ This function returns appends a (conc,kcat) to the input kcat_list.
    It also plots a figure of total reaction profile, liear portion plotted, and residuals.
    The function takes an enzyme concentration (M) and substrate concentration (uM) and determines
    kobs.
    """
    if show==False:
        plt.ioff()


    G6P_values_x=list()
    G6P_values_y=list()
    All_G6P_values_x=list()
    All_G6P_values_y=list()

    #============================================================================
    # Opening files
    #============================================================================


    f = open(text_file, 'r')
    fh = f.readlines()
    f.close()
```

```
for i in fh[1:]:
    line = i.split()
    All_G6P_values_x.append(int(line[0]))
    All_G6P_values_y.append(float(line[1]))
    if int(line[0]) > lo_val and int(line[0]) < hi_val:
        G6P_values_x.append(int(line[0]))
        G6P_values_y.append(float(line[1]))

#=========================================================================
# Fitting slopes to specific regions
#=========================================================================

G6P_slope, G6P_intercept, G6P_r_value, G6P_p_value, \
G6P_std_err = stats.linregress(G6P_values_x,G6P_values_y)

#=========================================================================
# Calculating regions
#=========================================================================

G6P_kcat=(G6P_slope)/enz_conc

#=========================================================================
# Plot all of turnover reaction
#=========================================================================
fig = plt.figure(figsize=(16,16))

ax = fig.add_subplot(313)

#Plot
plt.errorbar(All_G6P_values_x, All_G6P_values_y, yerr=G6P_std_err,
             color='black', fmt='o—', markersize=4, label="All data")
plt.errorbar(G6P_values_x, G6P_values_y, yerr=G6P_std_err,
             color='red', fmt='o—', markersize=4, label="Region fitted")
plt.grid(True)
plt.legend(loc="best")
#Axes labels
plt.xlabel('Time (s)',fontsize=18)
plt.ylabel("Value",fontsize=18)

#=========================================================================
# Plot selected region
#=========================================================================
ax = fig.add_subplot(312)

#Plot
plt.errorbar(G6P_values_x, G6P_values_y, yerr=G6P_std_err,
```

```
            color='red', fmt='o—', markersize=4, label="Region fitted")
abline(G6P_slope, G6P_intercept, pltcolor='blue')
plt.legend(loc="best")
# Axes labels
plt.ylabel("Value",fontsize=18)
plt.grid(True)

#Text
plt.annotate('$y=%0.2E x + %0.2f ..with..  R^2 = %0.2f$'\
 % (G6P_slope, G6P_intercept, math.pow(G6P_r_value,2)), xy=(0.65, 0.15), \
 xycoords='axes fraction')
plt.annotate('$k_{cat} = %.3f  s^{—1}$' % (abs(G6P_kcat)), xy=(0.65, 0.05), \
 xycoords='axes fraction')

#=============================================================================
# Plot residuals
#=============================================================================
ax = fig.add_subplot(311)
tempy=[]
for i in G6P_values_x:
    tempy.append(0.)

plt.plot(G6P_values_x, tempy, 'b——', label="LOBF")

predicted = eqn(G6P_slope, G6P_values_x, G6P_intercept)

difference = []
for i in predicted:
    #print i
    ind = predicted.index(i)
    dat1 = predicted[ind]
    dat2 = G6P_values_y[ind]
    diff = dat1 — dat2
    difference.append(diff)

plt.errorbar(G6P_values_x, difference, yerr=G6P_std_err, color='red', \
fmt="o", label="Difference")
plt.legend(loc="best")
plt.grid(True)
title_text=text_file+" , [E] = %0.1E M ,  [G1P] = %0.2f $\mu$M" % (enz_conc, substrate_conc)
plt.title(title_text)
text = "kcat_"+text_file+".png"
plt.savefig(text, dpi=300)

if show==False:
    plt.clf()
```

```
        fig.clf()
        plt.close()
    elif show==True:
        plt.show()




    #==========================================================================
    # Appending values to kcat list
    #==========================================================================
    kcat_list.append((substrate_conc, G6P_kcat))

def jcalc_kcat(text_file, lo_val, hi_val, enz_conc, substrate_conc, kcat_list):
    G6P_values_x=list()
    G6P_values_y=list()
    All_G6P_values_x=list()
    All_G6P_values_y=list()

    #==========================================================================
    # Opening files
    #==========================================================================
    f = open(text_file, 'r')
    fh = f.readlines()
    f.close()

    for i in fh[1:]:
        line = i.split()
        All_G6P_values_x.append(int(line[0]))
        All_G6P_values_y.append(float(line[1]))
        if int(line[0]) > lo_val and int(line[0]) < hi_val:
            G6P_values_x.append(int(line[0]))
            G6P_values_y.append(float(line[1]))
    #==========================================================================
    #  Fitting slopes to specific regions
    #==========================================================================
    G6P_slope, G6P_intercept, G6P_r_value, G6P_p_value, \
    G6P_std_err = stats.linregress(G6P_values_x,G6P_values_y)

    #==========================================================================
    #  Calculating regions
    #==========================================================================
    G6P_kcat=(G6P_slope)/enz_conc

    #==========================================================================
    # Appending values to kcat list
```

```
    #=========================================================================
    kcat_list.append((substrate_conc, G6P_kcat))



def Km_eqn(substrate, Vmax, Km):
    return Vmax * (substrate / (substrate + Km))

def Km_eqn_old(substrate, Vo, Vmax, Km):
    return Vmax * (substrate / (substrate + Km))

def Kf(x, p):
    return Km_eqn(x, *p)


#%%

def fit_bootstrap(p0, datax, datay, function, yerr_systematic=0.0):
    """ Function taken from: https://stackoverflow.com/questions/
    14581358/getting-standard-errors-on-fitted-parameters-using-the-optimize-leastsq-method-i

    Bootstrap error estimation using the Scipy optimize package. This function uses
    optimize.leastsq and minimizes differences.

    Returns:
        parameters_fit, parameter_error
    """

    errfunc = lambda p, x, y: function(x,p) - y

    # Fit first time
    pfit, perr = optimize.leastsq(errfunc, p0, args=(datax, datay), full_output=0)



    # Get the stdev of the residuals
    residuals = errfunc(pfit, datax, datay)
    sigma_res = np.std(residuals)


    sigma_err_total = np.sqrt(sigma_res**2 + yerr_systematic**2)

    # 100 random data sets are generated and fitted
    ps = []
    for i in range(100):

        randomDelta = np.random.normal(0., sigma_err_total, len(datay))
        randomdataY = datay + randomDelta
```

```
        randomfit, randomcov = \
            optimize.leastsq(errfunc, p0, args=(datax, randomdataY),\
                             full_output=0)

        ps.append(randomfit)

    ps = np.array(ps)
    mean_pfit = np.mean(ps,0)

    # You can choose the confidence interval that you want for your
    # parameter estimates:
    Nsigma = 1. # 1sigma gets approximately the same as methods above
                # 1sigma corresponds to 68.3% confidence interval
                # 2sigma corresponds to 95.44% confidence interval
    err_pfit = Nsigma * np.std(ps,0)

    pfit_bootstrap = mean_pfit
    perr_bootstrap = err_pfit
    return pfit_bootstrap, perr_bootstrap


#=============================================================================
#=============================================================================
#=============================================================================

#%%
#=============================================================================
# Functional part of code
#=============================================================================
# Cleaning up
gc.collect()
remake_figures="no"
just_calc_kcats="yes"

if remake_figures=="yes":
    kcat_list=[]
    # Sheet 2
    calc_kcat("C02.txt", 0, 600, enz_conc, 5, kcat_list)
    calc_kcat("C03.txt", 0, 1700, enz_conc, 15, kcat_list)
    calc_kcat("C04.txt", 0, 1600, enz_conc, 35, kcat_list)
    calc_kcat("C05.txt", 0, 1900, enz_conc, 50, kcat_list)
    calc_kcat("C06.txt", 0, 2100, enz_conc, 70, kcat_list)
    calc_kcat("C07.txt", 600, 1700, enz_conc, 100, kcat_list)
    calc_kcat("C08.txt", 700, 2100, enz_conc, 160, kcat_list)
    calc_kcat("C09.txt", 800, 2200, enz_conc, 230, kcat_list)
    calc_kcat("C10.txt", 900, 2300, enz_conc, 330, kcat_list)
```

```
    calc_kcat("D02.txt", 0, 600, enz_conc, 5, kcat_list)
    calc_kcat("D03.txt", 0, 1700, enz_conc, 15, kcat_list)
    calc_kcat("D04.txt", 0, 1600, enz_conc, 35, kcat_list)
    calc_kcat("D05.txt", 0, 1900, enz_conc, 50, kcat_list)
    calc_kcat("D06.txt", 0, 2100, enz_conc, 70, kcat_list)
    calc_kcat("D07.txt", 600, 1700, enz_conc, 100, kcat_list)
    calc_kcat("D08.txt", 700, 2100, enz_conc, 160, kcat_list)
    calc_kcat("D09.txt", 800, 2200, enz_conc, 230, kcat_list)
    calc_kcat("D10.txt", 900, 2300, enz_conc, 330, kcat_list)

    calc_kcat("E02.txt", 0, 600, enz_conc, 5, kcat_list)
    calc_kcat("E03.txt", 0, 1700, enz_conc, 15, kcat_list)
    calc_kcat("E04.txt", 0, 1600, enz_conc, 35, kcat_list)
    calc_kcat("E05.txt", 0, 1900, enz_conc, 50, kcat_list)
    calc_kcat("E06.txt", 0, 2100, enz_conc, 70, kcat_list)
    calc_kcat("E07.txt", 600, 1700, enz_conc, 100, kcat_list)
    calc_kcat("E08.txt", 700, 2100, enz_conc, 160, kcat_list)
    calc_kcat("E09.txt", 800, 2200, enz_conc, 230, kcat_list)
    calc_kcat("E10.txt", 900, 2300, enz_conc, 330, kcat_list)



if just_calc_kcats=="yes":
    kcat_list=[]
    # Sheet 2
    jcalc_kcat("C02.txt", 100, 500, enz_conc, 5, kcat_list)
#NOTE: This is bad. calc_kcat("C03.txt", 1200, 1700, enz_conc, 15, kcat_list)
    jcalc_kcat("C04.txt", 600, 1000, enz_conc, 35, kcat_list)
    jcalc_kcat("C05.txt", 600, 1200, enz_conc, 50, kcat_list)
    jcalc_kcat("C06.txt", 800, 1400, enz_conc, 70, kcat_list)
    jcalc_kcat("C07.txt", 900, 1600, enz_conc, 100, kcat_list)
    jcalc_kcat("C08.txt", 1300, 2300, enz_conc, 160, kcat_list)
    jcalc_kcat("C09.txt", 1700, 2700, enz_conc, 230, kcat_list)
    jcalc_kcat("C10.txt", 2400, 3400, enz_conc, 330, kcat_list)

    jcalc_kcat("D02.txt", 100, 500, enz_conc, 5, kcat_list)
    jcalc_kcat("D03.txt", 300, 700, enz_conc, 15, kcat_list)
    jcalc_kcat("D04.txt", 600, 1000, enz_conc, 35, kcat_list)
    jcalc_kcat("D05.txt", 600, 1200, enz_conc, 50, kcat_list)
    jcalc_kcat("D06.txt", 800, 1400, enz_conc, 70, kcat_list)
    jcalc_kcat("D07.txt", 900, 1600, enz_conc, 100, kcat_list)
    jcalc_kcat("D08.txt", 1300, 2300, enz_conc, 160, kcat_list)
    jcalc_kcat("D09.txt", 1700, 2700, enz_conc, 230, kcat_list)
    jcalc_kcat("D10.txt", 2400, 3400, enz_conc, 330, kcat_list)
```

```
    jcalc_kcat("C02.txt", 100, 500, enz_conc, 5, kcat_list)
    jcalc_kcat("D03.txt", 300, 700, enz_conc, 15, kcat_list)
    jcalc_kcat("C04.txt", 600, 1000, enz_conc, 35, kcat_list)
    jcalc_kcat("C05.txt", 600, 1200, enz_conc, 50, kcat_list)
    jcalc_kcat("C06.txt", 800, 1400, enz_conc, 70, kcat_list)
    jcalc_kcat("C07.txt", 900, 1600, enz_conc, 100, kcat_list)
    jcalc_kcat("C08.txt", 1300, 2300, enz_conc, 160, kcat_list)
    jcalc_kcat("C09.txt", 1700, 2700, enz_conc, 230, kcat_list)
    jcalc_kcat("C10.txt", 2400, 3400, enz_conc, 330, kcat_list)


#%%
bootstrap="yes"

plt.cla()

if bootstrap=="yes":

    # Sorting kcat list by concentration order
    kcat_list.sort(key=lambda tup: tup[0])
    xlist =[]
    ylist =[]
    for i in kcat_list:
        xlist.append(i[0])
        ylist.append(i[1])
    xdata=xlist
    ydata=ylist

    pfit, perr = fit_bootstrap(pstart, xdata, ydata, Kf)

    print("\n# Fit parameters and parameter errors from bootstrap method :")
    print("pfit = ", pfit)
    print("perr = ", perr)

    kcat = pfit[0]/enz_conc
    kcaterr = perr[0]/enz_conc

    fig = plt.figure()
    plt.plot(xdata, Km_eqn(xdata, *pfit), 'r-', label='Fit: $V_{max}=%5.2f \pm %5.2f s^{-1}, \
    K_{m}=%5.1f \pm %5.2f \mu M $' % (pfit[0], perr[0], pfit[1], perr[1]))
    plt.plot(xdata, ydata, '.k', label="Data")
    plt.xlabel("G1P concentration / $\mu M$")
    plt.ylabel("$Rate / M^{-1} s^{-1}$")
    #plt.plot(xdata, ydata, '.k', mfc='None')
    #plt.title("Errors estimated using Bootstrap")
    #plt.legend(loc="best")
```

```
    plt.savefig("R49K_Rate_vs_conc.png", dpi=300)
    plt.show()

#%%

#popt, pcov = curve_fit(Km_eqn_old, xlist, ylist, p0=[1 , 18 , 52])
#std = np.sqrt(np.diag(pcov))
#
#fig = plt.figure()
#ax = fig.add_subplot(111)
#plt.plot(xlist, Km_eqn(xlist, *popt), 'r-', label='Fit: $V_{max}=%5.2f s^{-1}, \
#K_{m}=%5.1f \muM$' % (popt[1], popt[2]))
#plt.plot(xdata, ydata, '.b', label="Data")
#plt.xlabel("beta-G1P concentration / $\muM$")
#plt.ylabel("$Rate / M^{-1} s^{-1}$")
#
#plt.legend(loc="best")
#plt.savefig("Rate_vs_conc.png", dpi=300)


# ============================================================
# Writing fitted output
# ============================================================

#%%

outfile="Rates_and_fit.txt"
out=open(outfile, "w")
out.write("# LABELS Concentration  Rate \n")

for i in kcat_list:
    line = "%s %s \n" % (i[0], i[1])
    out.write(line)

out.write("\n# LABELS  Vmax   VmaxErr   Km   KmErr \n")

line = "# PARAMETERS %s %s %s %s \n" % (pfit[0], perr[0], pfit[1], perr[1])
out.write(line)

out.write("\n# LABELS   kcat  kcatErr \n")
line = "# KCAT %s %s \n" % (kcat, kcaterr)
out.write(line)
```

Example output

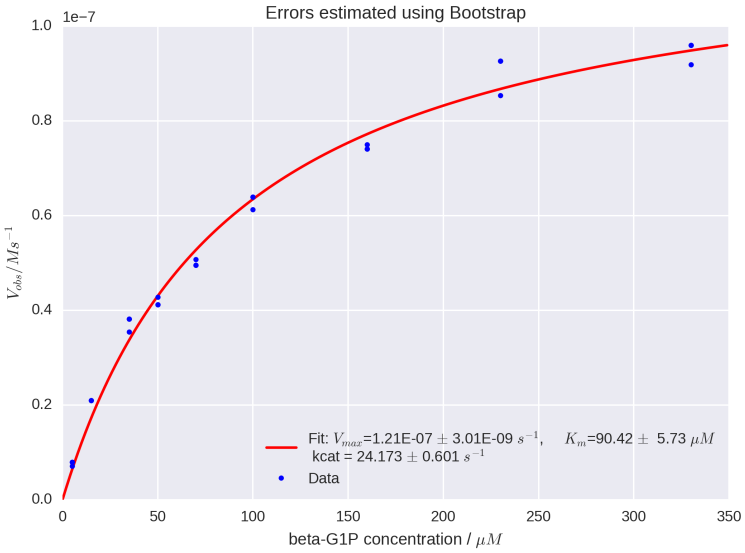


FIGURE B.6: Michaelis-Menten analysis for the βG1P dependence of k_{obs}. Bootstrap resampling is used to generate errors (see code).

B.3.3 DynDom heat map: when you have to compare an extensive number of pdb files

This program uses *dyndom* which is distributed in the CCP4i package to align a large number of PDB files. The code below and associated figures were generated for Paper III.

```
#!/usr/bin/python
"""
Created on 20 Jun 18 by AJR

This script will take a list of PDB files and run dyndom on each of the monomers.

This script needs to make comaprison directories from the input PDB file list.
This script needs to auto generate dyndom runfiles from the input pdb files.
This script needs to run dyndom on each of the runfiles in their respective dir.
It then needs to extract the rotation and translation value, as well as RMSD value.
DynDom values need to be read into a comparison matrix in pandas.

comparison column make to index.


@author Angus Robertson
"""
import os, sys, shutil, itertools
import numpy as np
import pandas as pd
import matplotlib
import matplotlib.pyplot as plt
import gc


plotting='yes'
plt.ioff()

#f=open("ensemble_models.pdb", 'r')
#fh=f.readlines()
#f.close()


mol_counter=0
text=str()


replace_contents=True


current_dir=os.getcwd()+"/"



#PDB_list = [("xtal6.pdb", "A"), ("xtal68.pdb", "A"), ("xtal68.pdb", "B"),
#("xtal59.pdb", "A"), ("xtal16.pdb", "A"), ("xtal16.pdb", "B"), ("xtal14.pdb", "A"),
```

```
#("xtal14.pdb", "B"), ("xtal13.pdb", "A"), ("xtal13.pdb", "B"), ("xtal70.pdb", "A"),
#("xtal70.pdb", "B"), ("xtal62.pdb", "A"), ("xtal55.pdb", "A"), ("xtal75.pdb", "A") ]

# Ordered list by angle of rotation away from 1ZOL

PDB_list = [("1ZOL", "A"),("2WHE", "A"),("6HDH", "A"),("6HDI", "A"),("6H93", "A"),
            ("6HDF", "B"),("6HDF", "A"),("6H93", "B"),("6HDH", "B"),("6HDI", "B"),
            ("6HDM", "A"),("6HDK", "A"),("2WF5", "A"),("2WF6", "A"),("6HDJ", "A"),
            ("6HDG", "A"),("6HDL", "A")]


DD_dict={}

#=========================================================================
# Functions
#=========================================================================

def write_command_file(directory, pdb1, chain1, pdb2, chain2):
    writefile=directory+"/run.command"
    f=open(writefile, "w")
    pdb_name=pdb1+pdb2+".w5"
    f.write("title=%s \n"%(pdb_name))
    f.write("filename1=%s \n" % (pdb1))
    f.write("chain1id=%s \n" % (chain1))
    f.write("filename2=%s \n" % (pdb2))
    f.write("chain2id=%s \n" % (chain2))
    f.write("clusters=20 \n" )
    f.write("iterations=100 \n" )
    f.write("window=5 \n" )
    f.write("domain=21 \n" )
    f.write("ratio=1 \n" )
    f.close()

def copy_pdb_files(directory, pdb1, pdb2):
    current_dir=os.getcwd()+"/"
    pdb1_curr=current_dir+pdb1
    pdb2_curr=current_dir+pdb2
    pdb1_dest=directory+"/"+pdb1
    pdb2_dest=directory+"/"+pdb2
    shutil.copyfile(pdb1_curr, pdb1_dest)
    shutil.copyfile(pdb2_curr, pdb2_dest)

def run_dyndom(directory):
    os.chdir(directory)
    os.system("dyndom run.command")
    os.chdir("..")
```

```
def DD_value_scrape(directory, pdb1, pdb2, chain1, chain2, DD_dict):
    #os.chdir(directory)
    pdb_name=directory+"/"+pdb1+pdb2+".w5"
    info_file=pdb_name+"_info"
    f=open(info_file, "r")
    fh=f.readlines()
    angle = np.nan
    trans = np.nan
    rmsd1 = np.nan
    rmsd2 = np.nan
    if len(fh) > 11:
        counter=1
        for i in fh:
            line = i.split()
            if "ANGLE OF ROTATION:" in i:
                angle = float(line[-2])
            if "TRANSLATION ALONG AXIS:" in i:
                trans = float(line[-2])
            if "BACKBONE RMSD ON THIS DOMAIN:" in i and counter==1:
                rmsd1 = float(line[-1][:-1])
                counter+=1
            if "BACKBONE RMSD ON THIS DOMAIN:" in i and counter==2:
                rmsd2 = float(line[-1][:-1])

    PDB = pdb1+"_"+chain1+"_"+pdb2+"_"+chain2
    DD_dict[PDB] = [angle, trans, rmsd1, rmsd2]
    values = [angle, trans, rmsd1, rmsd2]
    return values

def _color_sheet(val):
    color = 'white' if val == np.nan else 'green'
    return 'color: %s' % color


#==========================================================================
# Script
#==========================================================================

#%%
shortlist=[]
for i in PDB_list:
    shortlist.append(str(i[0]+"_"+i[1]))

df_angle = pd.DataFrame(index=shortlist, columns=shortlist)
df_trans = pd.DataFrame(index=shortlist, columns=shortlist)
```

```
df_rmsd1 = pd.DataFrame(index=shortlist, columns=shortlist)
df_rmsd2 = pd.DataFrame(index=shortlist, columns=shortlist)



#%%

#df_angle.loc["xtal6.pdb_A","xtal6.pdb_A"] = 1.5


#%%


for pair in itertools.combinations(PDB_list, 2):
    directory=pair[0][0]+"_"+pair[1][0]
    dest=os.getcwd()+'/'+directory
    #print directory
    try:
        os.mkdir(dest)
    except:
        pass

    write_command_file(dest, pair[0][0], pair[0][1], pair[1][0], pair[1][1])

    copy_pdb_files(dest, pair[0][0], pair[1][0])

    run_dyndom(dest)

    values = DD_value_scrape(dest, pair[0][0], pair[1][0], pair[0][1], pair[1][1], DD_dict)

    df_angle.loc[str(pair[0][0]+"_"+pair[0][1]), str(pair[1][0]+"_"+pair[1][1])] = values[0]
    df_trans.loc[str(pair[0][0]+"_"+pair[0][1]), str(pair[1][0]+"_"+pair[1][1])] = values[1]
    df_rmsd1.loc[str(pair[0][0]+"_"+pair[0][1]), str(pair[1][0]+"_"+pair[1][1])] = values[2]
    df_rmsd2.loc[str(pair[0][0]+"_"+pair[0][1]), str(pair[1][0]+"_"+pair[1][1])] = values[3]

    #input1 = open('file_%s' % pair[0], 'r')
    #input2 = open('file_%s' % pair[1], 'r')



#%%


"""
This section taken from:
```

```
https://matplotlib.org/gallery/images_contours_and_fields/image_annotated_heatmap.html
"""



def heatmap(data, row_labels, col_labels, ax=None,
            cbar_kw={}, cbarlabel="", **kwargs):
    """
    Create a heatmap from a numpy array and two lists of labels.

    Arguments:
        data       : A 2D numpy array of shape (N,M)
        row_labels : A list or array of length N with the labels
                     for the rows
        col_labels : A list or array of length M with the labels
                     for the columns
    Optional arguments:
        ax         : A matplotlib.axes.Axes instance to which the heatmap
                     is plotted. If not provided, use current axes or
                     create a new one.
        cbar_kw    : A dictionary with arguments to
                     :meth:`matplotlib.Figure.colorbar`.
        cbarlabel  : The label for the colorbar
    All other arguments are directly passed on to the imshow call.

    This function taken from:
    https://matplotlib.org/gallery/images_contours_and_fields/image_annotated_heatmap.html
    """


    if not ax:
        ax = plt.gca()

    # Plot the heatmap
    im = ax.imshow(data, **kwargs)

    # Create colorbar
    cbar = ax.figure.colorbar(im, ax=ax, **cbar_kw)


    cbar.ax.set_ylabel(cbarlabel, rotation=-90, va="bottom")



    # We want to show all ticks...
    ax.set_xticks(np.arange(data.shape[1]))
    ax.set_yticks(np.arange(data.shape[0]))
    # ... and label them with the respective list entries.
```

```
    ax.set_xticklabels(col_labels)
    ax.set_yticklabels(row_labels)

    # Let the horizontal axes labeling appear on top.
    ax.tick_params(top=True, bottom=False,
                   labeltop=True, labelbottom=False)

    # Rotate the tick labels and set their alignment.
    plt.setp(ax.get_xticklabels(), rotation=-30, ha="right",
             rotation_mode="anchor")

    # Turn spines off and create white grid.
    for edge, spine in ax.spines.items():
        spine.set_visible(False)

    ax.set_xticks(np.arange(data.shape[1]+1)-.5, minor=True)
    ax.set_yticks(np.arange(data.shape[0]+1)-.5, minor=True)
    ax.grid(which="minor", color="w", linestyle='-', linewidth=3)
    ax.tick_params(which="minor", bottom=False, left=False)

    return im, cbar


def annotate_heatmap(im, data=None, valfmt="{x:.2f}",
                     textcolors=["black", "white"],
                     threshold=None, **textkw):
    """
    A function to annotate a heatmap.

    Arguments:
        im          : The AxesImage to be labeled.
    Optional arguments:
        data        : Data used to annotate. If None, the image's data is used.
        valfmt      : The format of the annotations inside the heatmap.
                      This should either use the string format method, e.g.
                      "$ {x:.2f}", or be a :class:`matplotlib.ticker.Formatter`.
        textcolors  : A list or array of two color specifications. The first is
                      used for values below a threshold, the second for those
                      above.
        threshold   : Value in data units according to which the colors from
                      textcolors are applied. If None (the default) uses the
                      middle of the colormap as separation.

    Further arguments are passed on to the created text labels.
    """
```

```
    if not isinstance(data, (list, np.ndarray)):
        data = im.get_array()

    # Normalize the threshold to the images color range.
    if threshold is not None:
        threshold = im.norm(threshold)
    else:
        threshold = im.norm(data.max())/2.

    # Set default alignment to center, but allow it to be
    # overwritten by textkw.
    kw = dict(horizontalalignment="center",
              verticalalignment="center")
    kw.update(textkw)

    # Get the formatter in case a string is supplied
    if isinstance(valfmt, str):
        valfmt = matplotlib.ticker.StrMethodFormatter(valfmt)

    # Loop over the data and create a `Text` for each "pixel".
    # Change the text's color depending on the data.
    texts = []
    for i in range(data.shape[0]):
        for j in range(data.shape[1]):
            kw.update(color=textcolors[im.norm(data[i, j]) > threshold])
            text = im.axes.text(j, i, valfmt(data[i, j], None), **kw)
            texts.append(text)

    return texts

#%%


if plotting == 'yes':
#====================================================================
# Making Angles plot.
#====================================================================
    fig, ax = plt.subplots(figsize=(12,12))

    df_angle1 = df_angle.fillna(value=0.0)

    angles = df_angle1.values
    anglesTP = np.matrix.transpose(angles)
    all_angles = np.add(angles, anglesTP)

    im, cbar = heatmap(all_angles, shortlist, shortlist, ax=ax,
```

```
                    cmap="GnBu", cbarlabel="Angle [degrees]", cbar_kw={"shrink": .8})

    texts = annotate_heatmap(im, valfmt="{x:.1f}")

    plt.tight_layout()
    plt.savefig("angles.png", dpi=300)

    #%%
#============================================================================
# Making Translation plot.
#============================================================================

    fig, ax = plt.subplots(figsize=(12,12))

    df_trans1 = df_trans.fillna(value=0.0)

    trans = df_trans1.values
    transTP = np.matrix.transpose(trans)
    all_trans = np.add(trans, transTP)

    im, cbar = heatmap(all_trans, shortlist, shortlist, ax=ax,
                    cmap="GnBu", cbarlabel="Translation [$\AA$]", cbar_kw={"shrink": .8})

    texts = annotate_heatmap(im, valfmt="{x:.1f}")

    plt.tight_layout()
    plt.savefig("translation.png", dpi=300)

#============================================================================
# Making rmsd domain 1 plot.
#============================================================================
    #%%
    fig, ax = plt.subplots(figsize=(12,12))

    df_rmsd1 = df_rmsd1.fillna(value=0.0)

    rmsd1 = df_rmsd1.values
    rmsd1TP = np.matrix.transpose(rmsd1)
    all_rmsd1 = np.add(rmsd1, rmsd1TP)

    im, cbar = heatmap(all_rmsd1, shortlist, shortlist, ax=ax,
                    cmap="YlOrRd", cbarlabel="RMSD domain 1", cbar_kw={"shrink": .8})

    texts = annotate_heatmap(im, valfmt="{x:.3f}")

    plt.tight_layout()
```

```
    plt.savefig("RMSD1.png", dpi=300)

#=============================================================================
# Making rmsd domain 2 plot.
#=============================================================================

    fig, ax = plt.subplots(figsize=(12,12))

    df_rmsd2 = df_rmsd2.fillna(value=0.0)

    rmsd2 = df_rmsd2.values
    rmsd2TP = np.matrix.transpose(rmsd2)
    all_rmsd2 = np.add(rmsd2, rmsd2TP)

    im, cbar = heatmap(all_rmsd2, shortlist, shortlist, ax=ax,
                       cmap="PuRd", cbarlabel="RMSD domain 2", cbar_kw={"shrink": .8})

    texts = annotate_heatmap(im, valfmt="{x:.3f}")

    plt.tight_layout()
    plt.savefig("RMSD2.png", dpi=300)


#=============================================================================
#       Making combined RMSD plot.
#=============================================================================

    fig, ax = plt.subplots(figsize=(12,12))

    rmsd_combined = np.add(rmsd1, rmsd2TP)

    im, cbar = heatmap(rmsd_combined, shortlist, shortlist, ax=ax,
                       cmap="PuRd", cbarlabel="RMSD both domains", cbar_kw={"shrink": .8})

    texts = annotate_heatmap(im, valfmt="{x:.3f}")

    plt.tight_layout()
    plt.savefig("Both_RMSD.png", dpi=300)

plt.clf()
plt.close()
gc.collect()
```

Example output

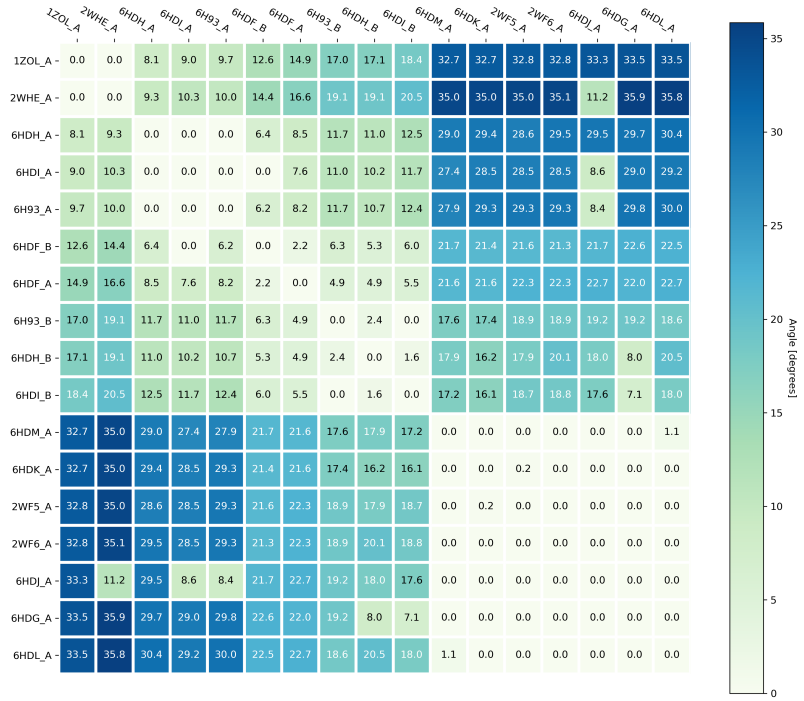


FIGURE B.7: The DynDom derived angle of rotation for each pairwise comparison

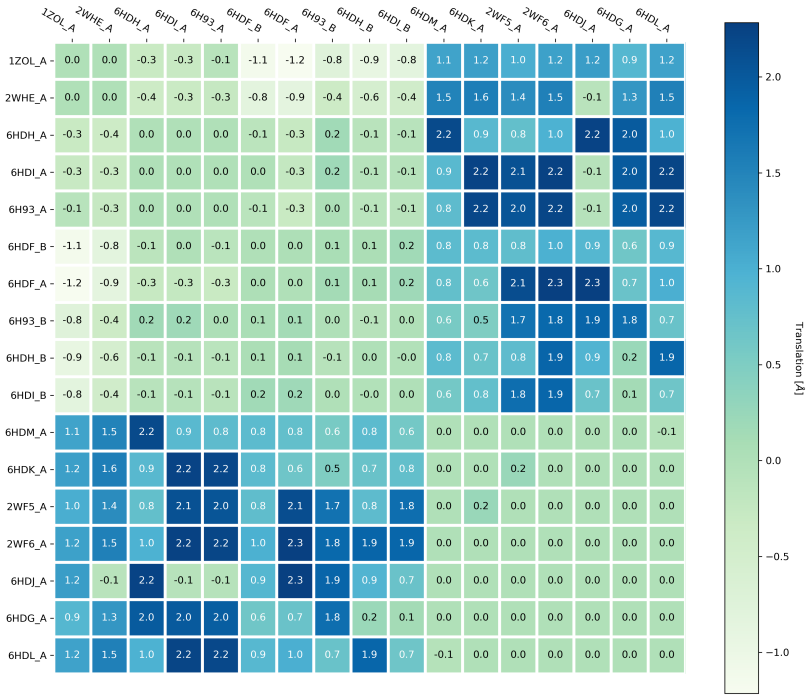


FIGURE B.8: The DynDom derived translation for each pairwise comparison

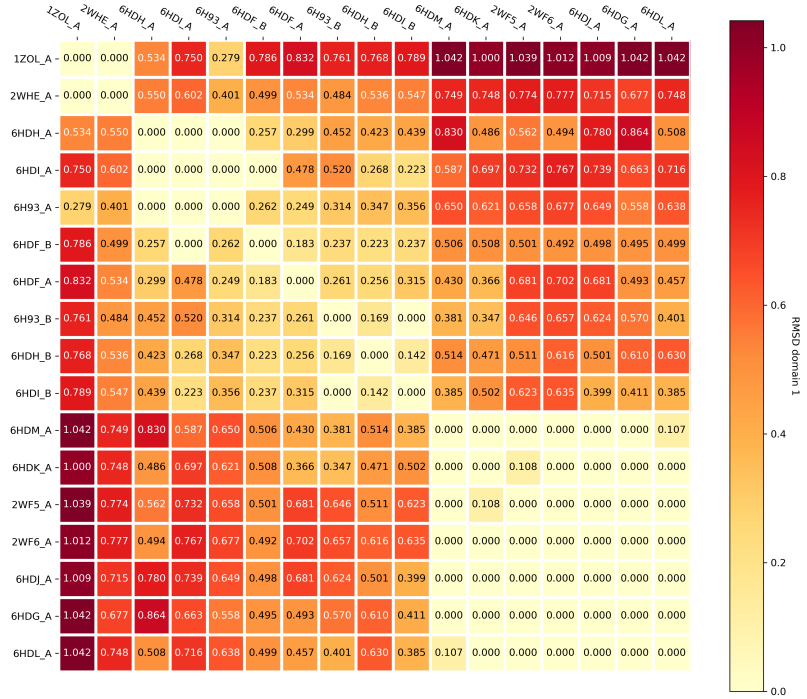


FIGURE B.9: The DynDom derived RMSD comparison for domain 1 (core domain of βPGM) for each pairwise comparison.

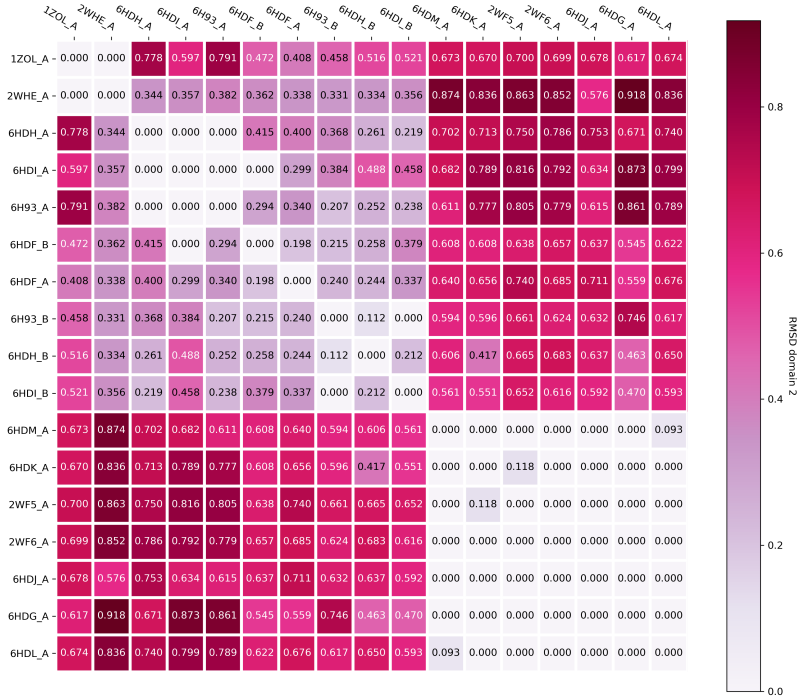


FIGURE B.10: The DynDom derived RMSD comparison for domain 2 (cap domain of βPGM) for each pairwise comparison.

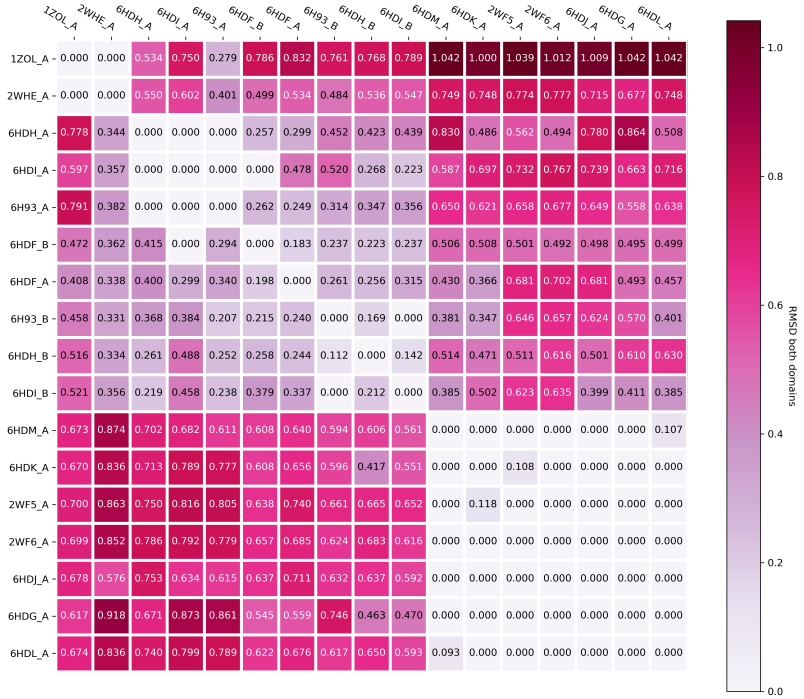


FIGURE B.11: The DynDom derived RMSD comparison for domain 1 (top right) and 2 (bottom left) for each pairwise comparison.