Thermodynamics of hinge bending in β-Phosphoglucomutase

Adam J. Flinders

Supervisors:

Professor Jonathan P. Waltho

Professor Jane A. Grasby

Department of Chemistry

Thesis submitted to the University of Sheffield

for the degree of Doctor of Philosophy

June 2023

Table of Contents

List of Figures v
List of Tables vi
Abbreviations vii
Symbols viii
Acknowledgements ix
Abstract x
1 General Introduction 1
  1.1 Enzyme catalysis 1
    1.1.1 Transition state theory 1
    1.1.2 Electrostatic pre-organisation 2
    1.1.3 The Spatiotemporal Hypothesis 2
    1.1.4 Near Attack Conformers 3
    1.1.5 Ground State Destabilisation 5
    1.1.6 Desolvation 7
    1.1.7 Conformational Dynamics 7
  1.2 β-Phosphoglucomutase 8
    1.2.1 Introduction 8
    1.2.2 Metal Fluoride Complexes 10
    1.2.3 Step-2 TSA complexes 10
    1.2.4 Step-1 TSA complexes 12
    1.2.5 Trifluoroberyllate complexes 13
    1.2.6 βPGM_{D10N} complexes 14
    1.2.7 Computational studies of proton transfer 15
    1.2.8 Allomorphy 15
2 Characterisation of Hinge Dynamics in βPGM 17
  2.1 Introduction 17
    2.1.1 Conformational dynamics of globular domains 17
    2.1.2 MD Simulations – Theory and Developments 22
  2.2 Results: MD Simulations 28
    2.2.1 βPGM Crystal Contacts 28
    2.2.2 Molecular Dynamics 30
    2.2.3 Conformer c_2 33
    2.2.4 Conformer c_1 36
    2.2.5 Conformer c_3 39
    2.2.6 Conformer c_4 41
    2.2.7 P146 isomerisation effects 42

2.2.8 A partially closed, substrate-free conformer 45
2.3 Results: Experimental Validation.. 47
2.3.1 Introduction 47
2.3.2 Small Angle X-ray Scattering 47
2.3.3 Analytical ultracentrifugation 53
2.3.4 NMR Relaxation 58
2.4 Discussion 65
2.4.1 Describing Domain Orientation in βPGM.. 65
2.4.2 Experimental Validation 65
2.4.3 D10 artefacts 66
2.4.4 Salt-bridge artefacts... 67
2.4.5 A hydrophobic cavity facilitates product release. 68
2.5 Conclusions 72
3 Describing thermodynamic coupling between phosphodianion binding and hinge bending 73
3.1 Introduction 73
3.1.1 The phosphodianion-driven activation framework 73
3.1.2 Distal site mutants in βPGM 74
3.1.3 How is phosphodianion binding coupled to domain closure?......................... 77
3.2 Results 78
3.2.1 A communication pathway between the hinge and the distal site.................. 78
3.2.2 Solution properties of the $\beta PGM_{N77A}:MgF_3:G6P$ TSA complex.................... 82
3.2.3 G1P turnover kinetics.. 85
3.2.4 BeF3, G6P titration 88
3.2.5 α-Galactose-1-phosphate titration 91
3.2.6 N77A TSA crystal structures 92
3.3 Discussion 95
3.3.1 Energy landscape of βPGM_{N77A} 95
3.3.2 The coupling mechanism 96
3.4 Conclusions 98
4 Future perspectives.. 99
5 Materials & Methods.. 101
5.1 Buffers Solutions & Media 102
5.1.1 Luria Bertani (LB) Agar.. 102
5.1.2 LB Liquid Media.. 102
5.1.3 M9 Minimal media... 102
5.1.4 SOB media 102

5.1.5 SOC media 102
5.1.6 Trace elements 103
5.1.7 TFB1 buffer 103
5.1.8 TFB2 buffer 103
5.1.9 Standard HEPES Buffer 103
5.1.10 βPGM K-HEPES High Salt Buffer 104
5.1.11 Chromatography Buffer A 104
5.1.12 Chromatography Buffer B 104
5.1.13 Lysis Buffer 104
5.1.14 4X SDS-PAGE Stacking Gel Buffer 104
5.1.15 4X SDS-PAGE Resolving Gel Buffer 104
5.1.16 SDS-PAGE Stacking Gel (4.5% BisAcrylamide) 105
5.1.17 SDS-PAGE Resolving Gel (18% BisAcrylamide) 105
5.1.18 SDS-PAGE Running Buffer 105
5.1.19 2X SDS-PAGE Loading Buffer 105
5.1.20 Gel stain and destain 105
5.2 Methods 106
5.2.1 Crystal structure processing 106
5.2.2 Crystal contact analysis 107
5.2.3 Molecular Dynamics acquisition 107
5.2.4 MD Analysis 109
5.2.5 Molecular Biology 118
5.2.6 Small Angle X-ray Scattering 123
5.2.7 Analytical Ultracentrifugation 126
5.2.8 NMR Relaxation 127
5.2.9 X-ray Crystallography 128
5.2.10 G6P titration in the presence of MgF_3 TSA 130
5.2.11 αGal1P titration 130
5.2.12 Steady State Kinetics 131
5.2.13 G6P titration in the presence of BeF_3 GSA 134
5.2.14 Multiple Sequence Alignment 134
6 Appendix 135
6.1 Thermodynamics 135
6.2 The Michaelis-Menton model 137
6.3 Two-state 1:1 binding model 139
6.4 The King-Altman Model: Bi Bi Ping-Pong Reactions 140
6.5 Linear Discriminant Analysis 141

7 Supplementary Data 142
  7.1 Euler Angles by simulation 142
  7.2 Y19 sidechain trajectory 143
  7.3 P146 loop convergence 144
  7.4 Hierarchical clustering of βPGM crystal structures 145
  7.5 Crystal Structure Euler Angles 146
  7.6 Hydrogen bond lengths in βPGM TSA crystal structures 147
  7.7 Structure of the 70s helix 148
  7.8 Crystal Structure Processing and Refinement Statistics 149
  7.9 βPGM NMR resonance assignments 150
  7.10 ShiftX2 predictions 151
  7.11 Conserved Residues 152
  7.12 List of pgmB sequences 153
  7.13 Y19 sidechain morph animation 178
8 Bibliography 179

List of Figures

Figure 1-1 – Near Attack Conformer Model 4
Figure 1-2 – Models of enzyme catalysis 6
Figure 1-3 – Ground State Destabilisation 6
Figure 1-4 – β-Phosphoglucomutase isomerisation mechanism 9
Figure 1-5 – βPGM$_{WT}$ active site in a Step-2 TSA complex 11
Figure 1-6 – Sugar coordination in βPGM TSA complexes 12
Figure 2-1 – βPGM crystal packing contacts 29
Figure 2-2 – βPGM crystal packing contact examples 29
Figure 2-3 – Cap Domain Euler Angles 32
Figure 2-4 – A hydrophobic cavity in substrate-free βPGM 33
Figure 2-5 – c_2 Ramachandran distributions 34
Figure 2-6 – Hinge conformations in βPGM 37
Figure 2-7 – c_1 Hinge conformation 37
Figure 2-8 – c_1 Ramachandran distributions 38
Figure 2-9 – c_3 domain orientation and hydrogen bonding 40
Figure 2-10 – c_3 hinge conformations 40
Figure 2-11 – A salt bridge biases the c_4 conformer 41
Figure 2-12 – Forcefield electrostatics bias K145 recruitment 42
Figure 2-13 – Metadynamics sampling of K145/P146 cis/trans isomerism 44
Figure 2-14 – Forcefield electrostatics bias domain orientation 44
Figure 2-15 – Sampling of a substrate-free closed conformation 46
Figure 2-16 – SAXS curves for open and closed βPGM complexes 48
Figure 2-17 – Model optimisation for SAXS data analysis 50
Figure 2-18 - χ^2 vs. R_g profiles from an optimised implicit solvation model 51
Figure 2-19 - R_g Distributions 52
Figure 2-20 – Linearly extrapolated Sedimentation coefficients 55
Figure 2-21 – Radius of Hydration distributions 57
Figure 2-22 – Correlation time (τ_m) distributions 60
Figure 2-23 – local τ_m distributions 62
Figure 2-24 – Collapse of a hydrophobic cavity in the hinge of βPGM 71
Figure 3-1 – Phosphodianion-driven activation framework 76
Figure 3-2 – Distal site coordination in open and closed conformations 77
Figure 3-3 – βPGM Chemical Shift Library 80
Figure 3-4 – Communication between N77 and I84 via the Y80/V81 peptide bond 81
Figure 3-5 – Conformation of backbone from M83-N77 is sensitive to domain closure 81
Figure 3-6 – βPGM$_{WT}$ MgF$_3$ G6P titration 83
Figure 3-7 – Chemical shift changes between βPGM$_{WT}$ and βPGM$_{N77A}$ TSA complexes 84
Figure 3-8 – King-Altman diagrams 86

Figure 3-9 – Steady-state kinetics data 87
Figure 3-10 – G6P titration into a βPGM:BeF_3 complex 89
Figure 3-11 – αGal1P titration into substrate-free βPGM 91
Figure 3-12 – Conformation of mutation site within βPGM_{N77A}:MgF_3:G6P TSA complex 93
Figure 3-13 – RMSF analysis of the βPGM_{N77A}:MgF_3:G6P TSA complex 94
Figure 5-1 – The pET-22b(+) vector 119
Figure 7-1 – Cap domain Euler angles from βPGM MD simulations 142
Figure 7-2 – Y19 occupancy of hinge cavity 143
Figure 7-3 – Conformational convergence in the 140s loop under a metadynamics bias 144
Figure 7-4 – Hierarchical clustering of βPGM crystal structures 145
Figure 7-5 – βPGM_{N77A} TSA crystal structures have distorted cap domain Euler Angles 146
Figure 7-6 – βPGM_{N77A} TSA hydrogen bonding in the 70s Helix 147
Figure 7-7 – Residues in α4 which are sensitive to closure 148
Figure 7-8 – ShiftX2 chemical shift predictions 151
Figure 7-9 – βPGM Residue Conservation Scores. 152

List of Tables

Table 2-1 – Summary of molecular dynamics simulations 30
Table 2-2 – Cap Domain Euler Angles. 32
Table 2-3 – βPGM R_g values for crystal structures and MD conformers 51
Table 2-4 – βPGM R_H values for crystal structures and MD conformers 54
Table 2-5 – Hydrodynamic variables for R_H calculations 56
Table 2-6 – Diffusion tensor eigenvalue distributions 61
Table 2-7 – Prolate diffusion models of relaxation data 64
Table 3-1 – Parameters derived from titrations of G6P into βPGM:BeF_3 complexes 90
Table 3-2 – Free energy changes for different βPGM_{N77A} and βPGM_{WT} complexes 95
Table 5-1 – The βPGM PDB dataset 106
Table 5-2 – Electronic Continuum Correction Parameters for monoatomic ions 108
Table 7-1 – Summary of βPGM NMR resonance assignments 150

Abbreviations

AcP	Acetyl Phosphate
ADP	Adenosine Diphosphate
ATP	Adenosine Triphosphate
αGal1P	α-galactose 1-phosphate
α4	α-Helix 4 (70s helix)
AUC	Analytical Ultracentrifugation
COMT	Catechol O-Methyltransferase
CoA	Coenzyme A
CV	Collective Variable
ECC	Electronic Continuum Correction
EVB	Empirical Valence Bond
FAD	Flavine Adenine Dinucleotide
FRET	Fourier Resonance Energy Transfer
F16BP	Fructose 1,6-Bisphosphate
βG1P	β-Glucose 1-Phosphate
βG16BP	β-Glucose 1,6-Bisphosphate
G6PDH	Glucose 6-Phosphate Dehydrogenase
G6P	Glucose 6-phosphate
GDH	Glutamate Dehydrogenase
GPU	Graphics Processing Unit
GSD	Ground State Destabilisation
HAD	Haloalkanoic Acid Dehalogenase
IPTG	Isopropyl ß-D-1-Thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
KSI	Ketosteroid Isomerase
LJ	Lennard Jones
LDA	Linear Discriminant Analysis
LB	Luria Bertani / Luria Broth / Lysogeny Broth
MD	Molecular Dynamics
NAC	Near Attack Conformer
NSE	Neutron Spin Echo
NAD	Nicotinamide Adenine Dinucleotide
NMR	Nuclear Magnetic Resonance
NSR	Nuclear Spin Relaxation
OMPDC	Orotidine 5'-Monophosphate
PME	Particle Mesh Ewald
PBC	Periodic Boundary Conditions
PBP	Periplasmic Binding Proteins
PGM	Phosphoglucomutase
PGK	Phosphoglycerate Kinase
PDB	Protein Data Bank
PCS	Pseudo Contact Shift
QM	Quantum Mechanical
RDC	Residual Dipolar Couplings
RMSD	Root-Mean-Square-Deviation
SAXS	Small Angle X-ray Scattering
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
T4L	T4 Lysozyme

TSA	Transition State Analogue
TROSY	Transverse Relaxation Optimised Spectroscopy
TSP	Trimethylsilyl propionate
TIM	Triose Phosphate Isomerase

Symbols

$\mathbf{k_B}$	Boltzmann Constant
$\mathbf{k_{cat}}$	Catalytic Rate Constant
$\boldsymbol{\kappa}$	transmission coefficient
‡	Transition State
$\mathbf{R_g}$	Radius of Gyration
$\mathbf{K_m}$	Michaelis Constant

Acknowledgements

I would like to thank Jon for taking me on as a student, for your support and patience. I really appreciate having been given the freedom to explore ideas and develop new skills. Thank you also for providing clear guidance when I needed it and for sharing your insights and experience into research and academic life. Nicky, thank you for everything you had done to help me throughout my studies. You have been a great source of knowledge and enthusiasm, keeping me on track and motivated.

To all members of the lab, past and present, thanks for your uplifting and entertaining conversation and for creating a cheerful working environment. Thank you to Nick F for taking an interest in my work and for all the advice and encouragement! Thank you to Mike and Rosie for keeping the lab running smoothly, for asking helpful questions at lab meetings and for sharing your experience and expertise!

Thanks to Andy for making the time to help running and troubleshooting simulations, fixing scripts, and running analyses. Thank you to Michelle and Matt for keeping the NMR facilities running smoothly and for your help in getting experiments running. Thanks to Pat B for your help with crystal looping and for arranging the shoots. Thanks to Gemma H and David S for your help running the AUC experiments and to Michael L for your help in running the SAXS experiments.

Thank you to my Family for all your support and encouragement over the past few years and for patiently listening to my ramblings about work.

Annie, thanks for reminding me to “get a grip and stop being a stress head”. Thank you for making sure I took a break and for making my time off so much fun. I have really relied on your happy bubbly personality to keep me going.

Thanks to Sheffield for maintaining such a beautiful city and countryside – I couldn’t have asked for a better place to live these past few years.

Abstract

This thesis is primarily concerned with the characterisation of domain reorientation and hinge-bending motions in the conformational landscape of β-phosphoglucomutase (βPGM). βPGM comprises two domains connected by a flexible hinge region. The catalytic cycle involves dynamical exchange between an open, inactive conformation – to which the substrate can bind and from which the product is released – and a closed, catalytically competent conformation.

X-ray crystallography has been used successfully to describe key structures along the reaction coordinate of βPGM. A multidisciplinary study, combining molecular dynamics computer simulations with small angle x-ray scattering and hydrodynamic measurements, is described to establish the conformational landscape in solution. Several simulation artefacts are described highlighting the importance of recent forcefield optimisations in the study of collective motions in proteins. The experimentally validated MD ensemble is more open than the crystal structures and is stabilised by burial of the Y19 sidechain in a hydrophobic pocket within the hinge region. This mechanism may serve to stabilize the substrate-free, open conformer, facilitating product release.

In accordance with the phosphodianion-driven enzyme activation framework, domain closure is stimulated by recruitment of an inert phosphodianion group to the distal site. Using a combination of NMR, x-ray crystallography and steady state kinetics, a communication pathway between the distal phosphodianion binding site and the hinge has been characterised. This involves a hydrogen bonding relationship between a pair of carboxamides (N77 and N118) which in turn modifies the hydrogen bonding relationships within the 70s helix, and the backbone torsions of I84.

1 General Introduction

1.1 Enzyme catalysis

1.1.1 Transition state theory

A reaction coordinate is generally characterised by two ground states (reactants and products) separated by a high energy transition state. The reaction rate is dependent on the energy barrier separating the ground state from the transition state. This is quantified by the Eyring equation.

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

The proportion of activated molecules is given by the transition state, free energy barrier ($\Delta G^{\ddagger}$). The rate at which activated molecules are converted to product is determined by the bond vibrational frequency $\frac{k_B T}{h}$, where h is plank's constant. Not all activated molecules will necessarily proceed to the product state, some will revert to the reactant state. This is accounted for by the transmission coefficient (κ). A key principle in the catalytic theory of enzymes is that they must reduce the activation barrier ($\Delta G^{\ddagger}$), thereby increasing the proportion of molecules which can cross the energy barrier at a given temperature.

The reaction coordinate for an enzyme-catalysed reaction is made more complicated as it necessarily includes substrate binding and product dissociation steps, as well as, in some cases, binding, turnover and dissociation of intermediates. Consequently, transition state theory is often interpreted in the context of the Michaelis-Menten model, from which we obtain two key parameters k_{cat} and K_m (section 6.2).

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

There are two limiting cases in this model. Where the substrate concentration is sufficiently larger than K_m, the reaction is effectively unimolecular, and the activation barrier is defined by the energy difference between ES and $ES^{\ddagger}$ (k_{cat}).

$$[S] + K_m \approx [S] \qquad\qquad v \approx k_{cat}[E]_{TOT}$$

Where the substrate concentration is significantly lower than K_m, the reaction is bimolecular, and the activation barrier is defined by the energy difference between E+S and $ES^{\ddagger}$ ($k_{cat}[S]/K_m$)[1].

$$[S] + K_m \approx K_m \qquad\qquad v \approx \frac{[S]}{K_m} k_{cat}[E]_{TOT}$$

Using this framework, enzymatic rate-accelerations can be quantified relative to the equivalent reaction in a solvent cage. A compilation of data for 18 different enzyme catalysed reactions shows catalytic effects from 8 - 34 kcal/mol [2]. Many proposals have been put forward which describe how enzymes achieve such large rate-enhancements. These will be summarised in the following sections.

1.1.2 Electrostatic pre-organisation

The transition state can often be differentiated from the ground state by its electrostatics. Thus, the transition state can be specifically stabilised by tuning the electrostatic field within the active site in such a way that favours the transition state charge distribution over that of the ground state. This is manifested as a specific set of hydrogen bonding and electrostatic interactions between the enzyme and the reactive functional group(s) and in some cases is assisted by recruitment of metal ion cofactors [1]. This effect is more accurately described as a pre-organisation effect. Solvent dipoles, which would otherwise be randomly oriented, can reorganise themselves to optimally complement transition state electrostatics, but incur a large entropic penalty in doing so. Conversely, the folding energy of the enzyme creates an active site with an electrostatic distribution which is partially pre-organised to complement that of the transition state. This incurs a much smaller re-organisation cost [2]. EVB calculations have been used to demonstrate electrostatic pre-organisation in Ketosteroid Isomerase [3–5], catechol O-methyltransferase (COMT) [6], orotidine-monophosphate decarboxylase (OMPDC) [7]. *ab initio* QM/MM methods were used to identify significant electrostatic contributions of K12 to catalysis in Triose phosphate isomerase (TIM) [8]. The doubly anionic enolate intermediate is stabilised though the coordination by two catalytic magnesium ions [9]. Stabilisation of the charge distribution within a trigonal bipyramidal phosphoryl transfer transition state of tyrosine phosphatase is achieved through hydrogen bonding and electrostatic stabilisation [10].

1.1.3 The Spatiotemporal Hypothesis

The spatiotemporal hypothesis refers to a simple model in which enzyme-like rate-enhancements can be achieved by holding reactants at close contact with favourable geometry [11]. This hypothesis is supported by several organic systems. For example, the amide in N,N-diethyl-naphthalamic acid is destabilised by 17.4 kcal/mol in a derivative where it is surrounded by two appropriately positioned carboxylate moieties (2-carboxy-N,N-diethylnaph-thalamic acid). This mechanism was proposed to be analogous to that of aspartyl proteases [12]. Similar rate enhancements are observed for phosphodiester hydrolysis, where two imidazole moieties were attached to a scaffold placing them appropriately to activate a water molecule for nucleophilic attack [11]. Examples were identified in reactions which both relieve and generate geometric strain, thus excluding the possibility of ‘steric acceleration’ effects [13]. Examples are not limited to intramolecular catalysis – a pillar 5 arene macrocycle with 10 imidazole moieties produced a 10^4-fold rate improvement in phosphodiester hydrolysis. Whilst this rate-enhancement is small compared to that of a typical phosphohydrolase enzyme, this organic catalyst does not impose close-contact and optimal geometry to the same degree [14]. Whilst there is evidence of spatiotemporal effects in small organic models, they are not necessarily utilised by enzymes. However, it is suggested that the numerous crystal structures of enzyme active sites, in which reactants are placed in close contact with near-optimal geometry, provide sufficient evidence of spatiotemporal effects in enzymes [11]. This model has however been called into question. For example, the amide cleavage reaction modelled by N,N-diethyl-naphthalamic, may have a substantially altered TS charge distribution owing to delocalisation of the oxyanion charge across the naphthalene ring. Furthermore, the origin of catalysis was not considered in a quantitative manner, with respect to an appropriate, uncatalysed reference reaction [15].

1.1.4 Near Attack Conformers

Catalysis via stabilisation of near attack conformers (NACs) is essentially the same as the spatiotemporal argument. NACs are defined as conformations of the ES complex in which the reacting atoms are within van der Waals' contact and are appropriately aligned for bond formation/cleavage. These conformations are high energy and must also be stabilised by the enzyme. The activation barrier can be divided into two parts: the energy required to transition from the ground state to the NAC (ΔG_{NAC}); and the energy require to transition from the NAC to the transition state (ΔG_{chem}). It has been proposed that enzymes reduce the activation barrier by reducing ΔG_{NAC} rather than ΔG_{chem} (**Figure 1-1**) [16]. Multiple NACs have been structurally and spectroscopically characterised in β-phosphoglucomutase [17,18] (Section 1.2), although no comment has been made regarding their stabilisation and the contribution this makes to catalysis. The conformational landscape of chorismate was characterised using MD simulations. A diaxial NAC was identified from these simulations, which is stabilised substantially by the active site of chorismate mutase (CM) [19]. Free energy calculations made using a series of mutant CM complexes, shows a direct proportionality between $\Delta G^‡$ and ΔG_{NAC} [20]. However, it has been argued that these results have been misinterpreted and that the apparent NAC effect is simply a consequence of electrostatic stabilisation of the transition state [21,22]

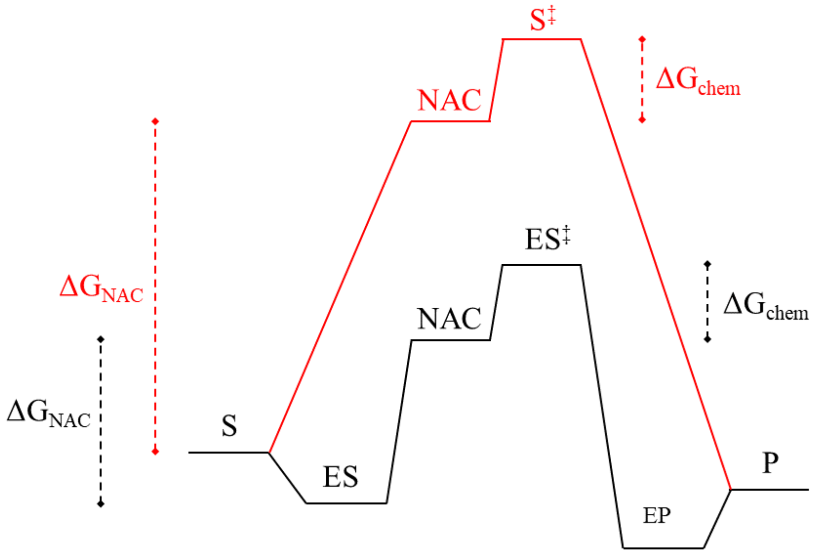


Figure 1-1 – Near Attack Conformer Model. The energy level diagrams for an uncatalysed reaction (red) and an enzyme-catalysed reaction (black). A high energy NAC, in which the reacting atoms are aligned and in close proximity, is stabilised in the enzyme catalysed reaction.

1.1.5 Ground State Destabilisation

Ground state destabilisation (GSD) refers to a reduction in the activation barrier – and therefore an increase in reaction rate – by increasing the energy of the ground state ES complex. Such mechanisms typically involve the introduction of geometric, steric or electrostatic strain into the substrate [23] (**Figure 1-2**). Early crystal structures of Hen Egg-White lysozyme provided a model of the peptidoglycan binding site [24]. The substrate could only be accommodated in this site if the fourth sugar was distorted away from the stable "chair" conformation towards a "half-chair" conformation which resembles a high-energy carbonium intermediate [25]. Substrate deformation has also been observed in other glycoside hydrolases [26]. A similar mechanism was described for chorismate mutase, in which the enzyme binds an unstable pseudo-diaxial chair-like conformation [27]. The conformations of three co-factors commonly found in protein structures (Adenosine 5'-Triphosphate (ATP), Nicotinamide-Adenine dinucleotide (NAD) and Flavin-Adenine dinucleotide (FAD)) were surveyed showing that in many cases torsion angles were perturbed significantly from their preferred positions [28]. An electrostatic destabilisation mechanism was also described for Alkaline phosphatase, in which the nucleophile (S102) destabilises the anionic substrate. Mutation of this residue to an alanine or glycine (at pH8) resulted in an order of magnitude increase in affinity for inorganic phosphate [29]. Similarly, in ketosteroid isomerase, the binding of ground state analogues to D38N and D38A mutants improved by 1-2 orders of magnitude [30]. A similar argument was made for the mechanism in Orotidine 5'-monophosphate decarboxylase, based on multiple structures from different species which place an aspartate adjacent to the substrate carboxylate [31]. Electrostatic strain has been suggested to contribute to catalysis by enoyl-CoA hydratase. An increase in magnetic shielding of the α-carbon atoms and de-shielding of the β- and carbonyl carbon atoms, in enoyl-CoA substrates bound to the active site of enoyl-CoA hydratase, demonstrates electronic polarisation and electrostatic strain within the substrate [32–34].

In general, enzymes can attain enhancements in rate by increasing the catalytic rate (k_{cat}) or by reducing the Michaelis constant (K_m) – thus the 'catalytic efficiency' (k_{cat} / K_m) is often cited as the target for optimisation by evolution. It has been pointed out that GSD mechanisms would (by definition) reduce the stability of ES relative to the unbound state (E+S) resulting in a compensatory effect on K_m. Thus, the overall catalytic efficiency (k_{cat}/K_m) would be largely unaffected [35,36]. The implication is that, whilst GSD mechanism do not themselves generate substantial rate improvements, they may be by-products of transition state stabilisation. Whilst GSD mechanisms may not have a primary role in catalysis, they are directly relevant to the prevention of substrate/product inhibition. (**Figure 1-3**).

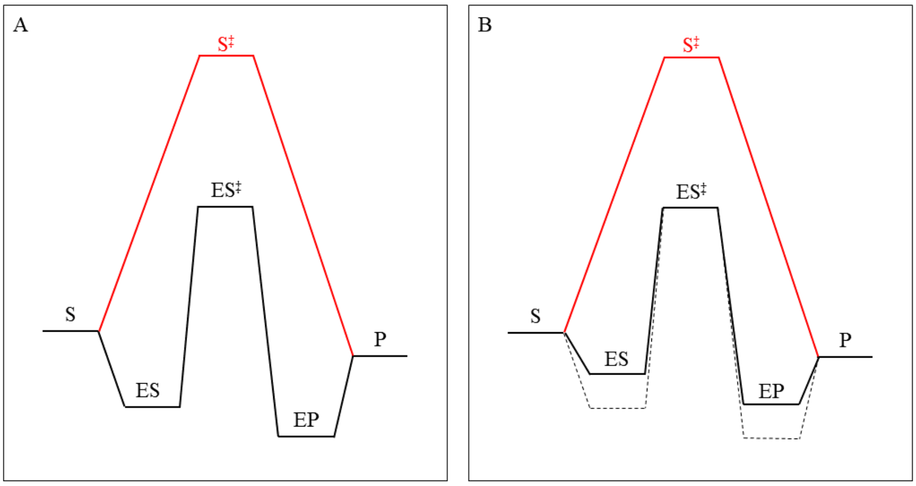


Figure 1-2 – Models of enzyme catalysis. The energy level diagrams for an uncatalysed reaction (red) and an enzyme-catalysed reaction (black). (A) Transition state stabilisation occurs via a ground state ES complex. (B) Destabilisation of the ground state EP complex prevents product inhibition, in which dissociation of EP complex becomes rate-limiting.

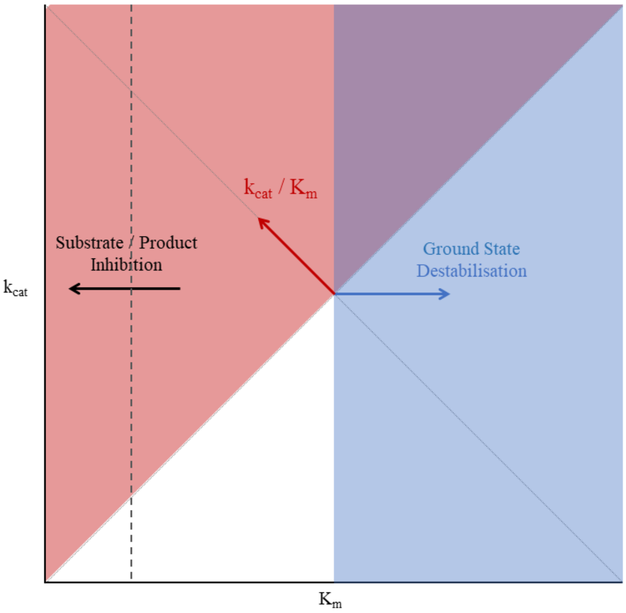


Figure 1-3 – Ground State Destabilisation. An abstract representation showing the relationships between catalytic efficiency (k_{cat} / K_m), substrate/product inhibition and ground state destabilisation. In general, any modification to an enzyme can be represented by a vector in this parameter space. The red region represents modifications which improve k_{cat}/K_m. The blue region represents modification which cause ground state destabilisation. The purple region represents modifications which meet both these criteria. Such modifications may be selected to improve catalytic efficiency, whilst preventing substrate/product inhibition.

1.1.6 Desolvation

Many chemical reactions are faster in organic solvent which has led to the suggestion, for some enzymes, that desolvation contributes to enzyme catalysis. More specifically, the low dielectric environment within the active site would enhance electrostatic stabilisation of the transition state [37,38]. Spectroscopic and biochemical measurements indicate a lower dielectric constant in proteins ($\epsilon < 20$) than in water ($\epsilon = 79$) [39,40]. However, simulation studies have been used to question the relevance and physical basis of this catalytic mechanism [41–43].

1.1.7 Conformational Dynamics

Conformational dynamics has been proposed to contribute to enzyme catalysis [44–48]. Several proposals identify dynamical motions, via a mixture of biophysical methods, which occur on the same timescale as catalysis. The suggestion is that conformational motions within the enzyme may somehow be coupled to the chemical step. However these dynamical proposals have been called into question [49,50]. Irrespective of the existence of a direct catalytic role for enzyme dynamics, many enzymes necessarily exhibit a conformational transition between open and closed conformations, originating from a set of ‘hinge’ residues. Such conformational transitions coincide with substantial changes in catalytic competency, substrate/product affinity and substrate accessibility. Thus, these conformational motions will be important in understanding how enzymes balances optimal TS stabilisation without over-stabilising ground-state complexes or introducing closed conformers which are inaccessible to the substrate.

1.2 β-Phosphoglucomutase

1.2.1 Introduction

β-Phosphoglucomutase (βPGM) from *Lactococcus lactis* is a member of the haloalkanoic acid dehalogenase (HAD) superfamily [51]. It has 222 residues with molecular weight 24.21 kDa. It has two domains connected by a flexible hinge region. The core domain (residues 1-15 and 88-221) has 6 parallel β-strands surrounded by α-helices. The cap domain (residues 16-87) comprises a bundle of 4 α-helicies (α1-α4). The domains are connected at two hinge regions, at the N-terminus of α1 (residues 14-16) and at the C-terminus of α4 (residues 85-90). The active site is found in the cleft between the two domains [52].

βPGM catalyses the isomerisation of β-Glucose-1-phosphate (βG1P) to Glucose-6-phosphate (G6P) via the β-Glucose-1,6-bosphosphate (βG16BP) intermediate. The enzyme is activated by phosphorylation of D8. βG1P binds with its 6'-hydroxyl oriented toward the aspartyl phosphate moiety and with the 1'-phosphate in the distal (inert) phosphate site. Phosphoryl transfer from D8 to βG1P produces the βG16BP intermediate, which is released and rebinds in the alternate orientation with the 1'-phosphate in the proximal site and 6'-phosphate in the distal site. Phosphoryl transfer from the 1'-phosphate to D8 yields the product (G6P) and regenerates the active phosphoenzyme [53] (**Figure 1-4**). βPGM is required for the catabolism of maltose and trehalose. Maltose is converted to glucose and βG1P by maltose phosphorylase. βPGM converts βG1P, to G6P which enters glycolysis. *L. lactis* βPGM mutants are unable to grow where maltose or trehalose are the sole carbon source [54].

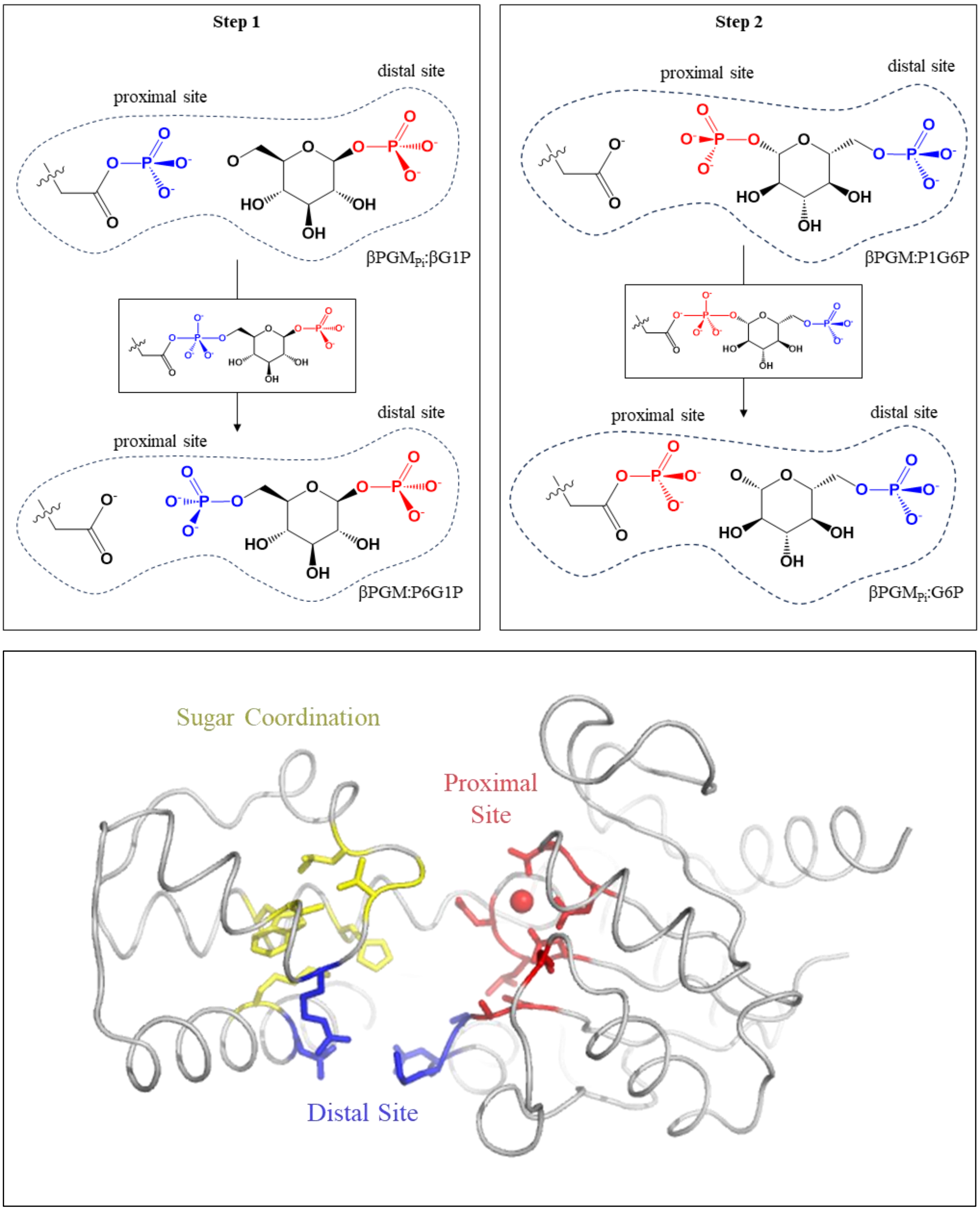


Figure 1-4 – β-Phosphoglucomutase isomerisation mechanism. (Top Panel) βG1P binds to βPGM$_{Pi}$ with the 1-Phosphate bound to the distal site and the 6-hydroxyl oriented toward the proximal site. Phosphoryl transfer from D8 to the 6-hydroxyl occurs via a trigonal bipyramidal transition state, yielding a βG16BP intermediate. βG16BP is released and rebinds with the 1-phosphate bound in the proximal site and the 6-phosphate bound in the distal site. Phosphoryl transfer from the 1-phosphate to D8 occurs yielding G6P and the active phosphoenzyme. (Bottom Panel) The global conformation of βPGM showing the architecture of the proximal (red), distal (blue) and sugar coordination (yellow) sites.

1.2.2 Metal Fluoride Complexes

Metal fluoride complexes have been used extensively to trap and characterise important complexes sampled along the reaction coordinate. Magnesium trifluoride (MgF_3^-) complexes have a trigonal bipyramidal (tbp) geometry with a net charge of 1-, making them isosteric and isoelectric transition state analogues (TSA) of phosphoryl transfer reactions. Aluminium tetrafluoride (AlF_4^-) complexes also act as transition state analogues but adopt an octahedral geometry to maintain a single negative charge. Beryllium trifluoride (BeF_3^-) complexes have a tetrahedral geometry and a net charge of -1, making them isosteric and isoelectric ground state phosphate analogues [55]. Other types of metal fluoride complex have been reported. In particular, MgF_3^- complexes are often misidentified as tbp AlF_3^0 complexes, where aluminium present in the sample precipitates as $Al(OH)_3$ at elevated pH, consequently MgF_3 becomes the dominant species [56]. However, a genuine AlF_3^0 complex (octahedral $AlF_3 \bullet H_2O$) has been identified in the K219A variant of Phosphoglycerate Kinase (PGK). The loss of a positive charge in the active site is complemented by the loss of a negative charge from the metal fluoride moiety to maintain charge balance [57]. The $\beta PGM_{WT}:MgF_3:G6P$ complex has also been misinterpreted as pentavalent phosphorane intermediate [58–60]. However, extensive computational, crystallographic and NMR studies have been used to confirm the identity of this complex as an MgF_3^- transition state analogue [61–64].

1.2.3 Step-2 TSA complexes

Many of the key functional interactions and residues can be identified from the $\beta PGM_{WT}:MgF_3:G6P$ transition state analogue complex (PDB: 2WF5). Residues of the proximal/catalytic phosphate site are found exclusively in the core domain. F_A from the MgF_3 moiety, receives three hydrogen bonds from the backbone amide of L9 and D10 and the sidechain hydroxyl of S114 and has the most downfield chemical shift (δ = -147.0 ppm) and the largest deuterium isotope effect ($\Delta\delta$ = 1.6 ppm) measured by ^{19}F NMR. F_B receives two hydrogen bonds from the backbone amide of A115 and the terminal amide of K145 and has an intermediate chemical shift (δ = -151.8 ppm) and isotope shift ($\Delta\delta$ = 1.4 ppm). F_C receives a single hydrogen bond from the 2'-hydroxyl of G6P and has the most upfield chemical shift (δ = -147.0 ppm) and smallest isotope shift ($\Delta\delta$ = 0.9 ppm). Hydrogen bond lengths measured in the crystal also correlated with X^1H-^{19}F scalar couplings. F_C is also coordinated by a magnesium ion cofactor, which is octahedrally coordinated by the backbone carbonyl of D10, the sidechain carboxylate of D8 and D170 and by two water molecules (**Figure 1-5**). In addition to the equatorial fluorine atoms, the MgF_3 moiety is axially coordinated by the side chain carboxylate of D8 (acceptor) and the 1'hydroxyl of G6P (donor). D10 acts as a proton donor/acceptor in the phosphoryl transfer mechanism – its sidechain coordinates the 1'-OH of G6P and the sidechain ($O\gamma_1$) of T16. The pyranose ring occupies a chair conformation and packs against His-20, with equatorial hydroxyl groups coordinated by water molecules buried in cap domain. The 2’OH is coordinated by the backbone amide of G46; the 3’OH is coordinated to S52-Oγ and L44-O indirectly via a water molecule; the 4’OH is coordinated directly to V47-O and to a water molecule which is in turn coordinated by K76 and the distal 6’-phosphate. The distal phosphate is coordinated by S116-Oγ N118-Oδ_2 and the backbone amide of K117 from the core domain. It is also coordinated directly by R49 guanidino side chain group and indirectly to the terminal amine of K76 via two water molecules, in a circular arrangement [61,64]. The structure of the βPGM$_{WT}$:AlF$_4$:G6P complex (PDB: 2WF6) is almost identical, except for the coordination of the AlF$_4$ moiety.

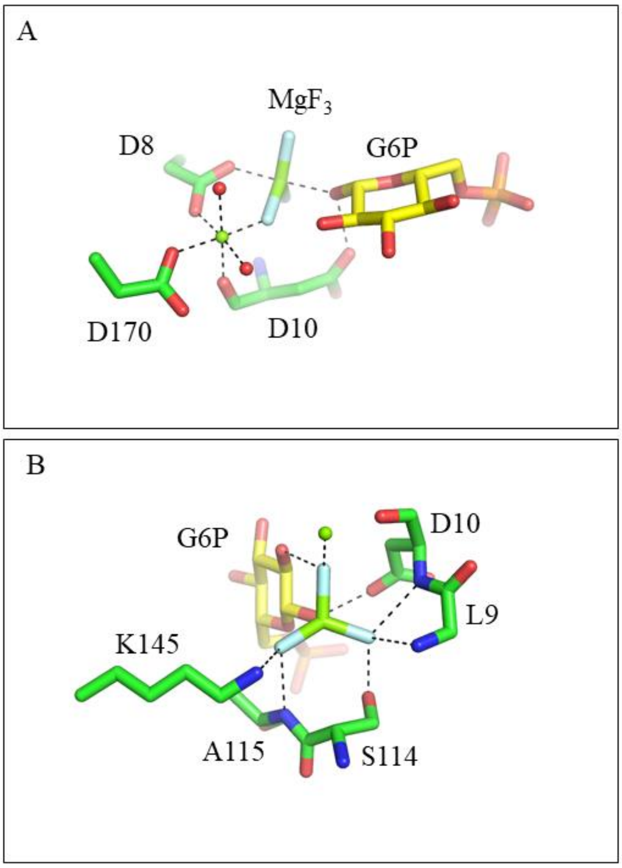


Figure 1-5 – βPGM$_{WT}$ active site in a Step-2 TSA complex. The architecture of the active site of the βPGM$_{WT}$:MgF$_3$:G6P TSA complex. **(A)** The catalytic magnesium ion (green sphere) is octahedrally coordinated by D10-O, D170-O$_\delta$, D8-O$_\delta$, F$_C$ and two water molecules. MgF$_3$ is a surrogate of the transferring PO$_3$ moiety – it is axially coordinated by the phosphate donor G6P-O$_1$ and the phosphate acceptor D8-O$_\delta$. **(B)** Coordination of the MgF$_3$ moiety: F$_A$ is coordinated by L9-N, D10-N and S115-O$_\gamma$; F$_B$ is coordinated by A115-N and K145-N$_\zeta$; and FC is coordinated by G6P-O$_2$ and the catalytic magnesium ion.

1.2.4 Step-1 TSA complexes

Step 1 TSA complexes βPGM:MgF_3:βG1P cannot be characterized structurally owing to incomplete inhibition by the MgF_3 complex, resulting in slow βG1P turnover. The rate of isomerisation is fast enough that the βPGM:MgF_3:βG1P is converted to a step 2 βPGM:MgF_3:G6P TSA complex during crystallisation. Step 1 TSA complexes were therefore obtained using phosphonomethylene and phoshphonofluoromethylene βG1P analogues (PDB: 2WF7, 4C4R, 4C4S, 4C4T). The conformation of the protein and coordination of the proximal and distal phosphate sites is virtually identical in these step-2 complexes. However, the orientation of the substrate is flipped (**Figure 1-6**). The 2’OH is directly coordinated by the terminal amine of K76; the 3’OH is coordinated directly to L44-O, S52-Oγ and W24-Nϵ; the 4’OH is coordinated directly to L44-O and G46-N. The α-face of the pyranose ring is packed against H20. The alternative orientation of the sugar is accommodated by rotation of the C5-C6 bond [65].

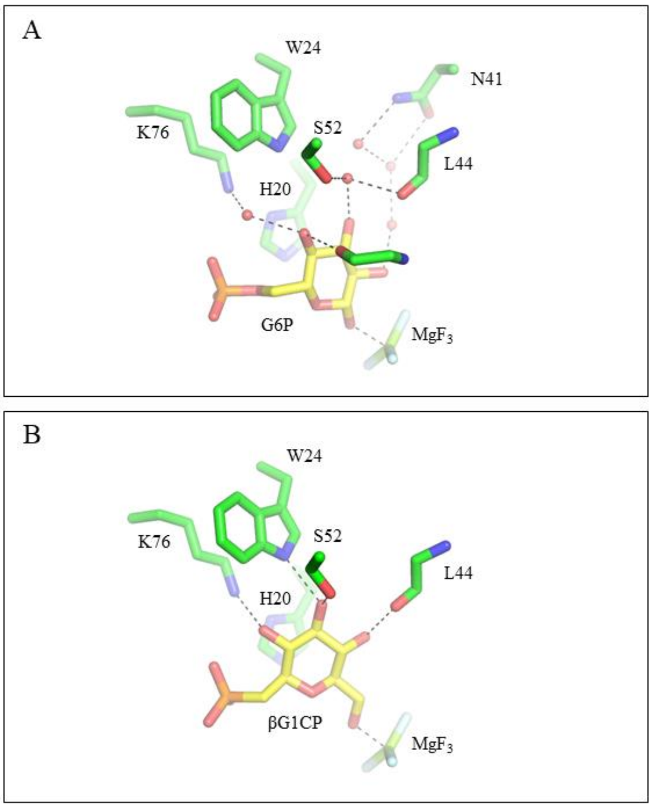


Figure 1-6 – Sugar coordination in βPGM TSA complexes. The interactions within the active site of (A) βPGM_{WT}:MgF_3:G6P step-2 TSA complex (B) βPGM_{WT}:MgF_3:βG1CP step-1 TSA complex. Water-mediated interactions between the pyranose ring and residues of the cap domain are made in the step-2 complex. Direct interactions between the pyranose ring and residues of the cap domain are made in the step-1 complex. The conformation of the protein is invariant between each complex.

1.2.5 Trifluoroberyllate complexes

The βPGM:BeF_3 complex acts as a substrate-free phosphoenzyme analogue (PDB: 2WFA). This complex has an *open* structure, resulting from a rotation of the cap domain by 33° around the hinge region. The unassigned resonances of the substrate-free complex – owing to millisecond exchange dynamics (D15-A17, L44-L53, S114-N118) – were assigned in the βPGM:BeF_3 complex. Wide-spread chemical shift changes are observed compared to the TSA complexes, consistent with a substantial domain reorientation and dissociation of the phospho-sugar. Coordination of the tetrahedral trifluoroberyllate moiety in the proximal site is equivalent to that of MgF_3 in the βPGM:MgF_3:G6P TSA complex, except D10 rotates about χ_1 to interact with the backbone amide of T16 (the 'out' position) and an additional water molecule coordinates F_A. ^{19}F spectra of this complex were assigned based on chemical shifts and deuterium isotope shifts. The assignments were consistent with the proximal coordination observed in the crystal structure. It was suggested that reorientation of the general base (D10) from the proximal site served to prevent activation of water molecules for nucleophilic attack of an aspartyl phosphate moiety. This was supported by the absence of any solvent atoms within 3.4 Å of the Be atom [17].

Addition of G6P to the βPGM:BeF_3 complex gives a βPGM:BeF_3:G6P product complex (PDB: 2WF9). This complex displays 50-60% of domain closure according to the chemical shift trajectories between the apo complex and the TSA complex, in agreement with the crystal structures which are closed but are rotated slightly relative to the TSA complexes. Coordination of the pyranose ring is similar to that of the βPGM:MgF_3:G6P and βPGM:AlF_4:G6P TSA complexes, except that both α and β-anomers are accommodated within the active site. D10 remains in the out position, interacting with T16-N and indirectly with the magnesium ion via a water molecule. Consequently, G6P-O_1 is protonated and forms a hydrogen bond with the BeF_3 moiety to reduce electrostatic repulsion. The donor-acceptor separation and the donor-Be-acceptor angle are only slightly perturbed from that of the TSA. This conformation was described as a near attack conformer (NAC I). Whilst this is a product complex, a similar enzyme conformation was predicted to be required, prior to the transition state, as an initial substrate alignment step for nucleophilic attack [17].

A second crystal form gave rise to a structure which much more closely resembled the TSA conformation and D10 is hydrogen bonded to the nucleophile (O1). The donor-acceptor distance in the G6P complex was comparable with that of the TSA complex. This complex was termed NAC II (PDB: 2WF8). The density showed two minor populations of a βG1P complex, one with the 6-OH and another with the 3-OH group oriented toward the BeF_3 moiety. However, these complexes have significantly longer donor-acceptor distances [17].

1.2.6 βPGM$_{D10N}$ complexes

βPGM$_{D10N}$ serves as a model of the WT enzyme where D10 is protonated. This variant copurifies with βG16BP which can only be removed by denaturation/refolding. βPGM$_{D10N}$ can be phosphorylated – confirmed using $^{1}H^{15}N$-TROSY NMR by comparison with the βPGM$_{D10N}$:BeF$_3$ and βPGM$_{WT}$:BeF$_3$ complexes. Phosphoenzyme hydrolysis was measured using ^{31}P NMR. The D10N mutation has only a modest effect on phosphoenzyme hydrolysis. Therefore, D10 is not involved in the activation of water molecules for nucleophilic attack of the aspartyl-phosphate moiety. Crystal structures also show that N10 is also in the out position with an equivalent hydrogen bonding arrangement to D10, in the βPGM$_{WT}$:BeF$_3$ complex. It is this conformation that was suggested to prevent D10 from activating water molecules for phosphoenzyme hydrolysis.

βPGM$_{D10N}$ has residual mutase activity in the presence of excess acetyl phosphate (AcP), which acts as a phosphorylating agent. However, the tightly bound βPGM$_{D10N}$:βG16BP complex is the dominant enzyme species. Two βG16BP complexes were characterised by x-ray crystallography: a βPGM$_{D10N}$:P1G6P complex (PDB: 5OK1) with the 1-phosphate in the proximal site and 6-phosphate in the distal site; and a βPGM$_{D10N}$:P6G1P complex (PDB: 5OK0) with the 6-phosphate in the proximal site and 1-phosphate in the distal site. The βPGM$_{D10N}$:P1G6P complex was shown to dominate in solution, using NMR methods. βPGM$_{D10N}$:P1G6P represents the enzyme conformation prior to the second phosphoryl transfer step. The nucleophile is aligned for attack, unlike the βPGM:BeF$_3$:G6P NAC I complex in which the nucleophile is hydrogen bonded to the fluorine atom. The D8-O phosphorus distance is within the sum of their van der Waals radii. Although, the donor-acceptor distance is slightly longer than that of the transition state analogue complexes. This complex was thus described as an aligned NAC – referred to here as NAC III. NAC III is more closed than NAC I but does not have a fully closed, TSA-like conformation. N10 coordinates the donor oxygen via Nδ2 and makes a bifurcated hydrogen bond with T16-N and T16-Oδ_2 – halfway between its canonical *out* position (found in the substrate-free complex, the βPGM:BeF$_3$ complex and in NAC I) and the catalytically active *in* position (found in the TSA complexes). The domain reorientations observed between NAC I, NAC III and the TSA complexes are largely dependent on torsion angles of the hinge residues D15 and T16. The βPGM$_{D10N}$:P6G1P complex represents the enzyme conformation which follows the first phosphoryl transfer step. Its conformation is very similar to that of the βPGM$_{D10N}$:P1G6P complex, except the alternative orientation of the sugar-phosphate gives rise to direct interactions with the cap domain, rather than water-mediated interactions. The same difference in sugar coordination is found between the step-2 TSA complexes and the step-1 α-fluorophosphonate TSA complexes. The βPGM$_{D10N}$:AlF$_4$:G6P TSA complex (PDB: 5OK2) is almost identical to the corresponding WT complex, except that N10 acts as the hydrogen bond donor to the deprotonated O1. Attempts to make a βPGM$_{D10N}$:AlF$_4$:βG1P step 1 TSA complex, surprisingly resulted in a βPGM$_{D10N}$:AlF$_4$:H$_2$O:βG1P complex (PDB: 5O6R), where the axial position in the AlF$_4$ complex was occupied by a water molecule rather than 6'OH of βG1P. It appears that the step 1 complexes are intrinsically less stable: The βPGM$_{D10N}$:P1G6P is favoured over the βPGM$_{D10N}$:P6G1P complex in solution; the βPGM$_{D10N}$:AlF$_4$:G6P complex makes a competent TSA, while the βPGM$_{D10N}$:AlF$_4$:βG1P reverts to a NAC III conformation; and G6P has a 5-fold higher affinity than βG1P for the βPGM:AlF$_4$ complex. The βPGM$_{D10N}$:P1G6P structure has a compromised Mg^{2+} coordination and a weaker magnesium affinity, compared to the βPGM$_{D10N}$:BeF$_3$ and βPGM$_{D10N}$:MgF$_3$:G6P complexes. Thus, Mg-coordination facilitates stabilisation of the transition state over the ground state [18,65].

1.2.7 Computational studies of proton transfer

EVB simulations of βPGM WT and D10N supported the role of D10 as the general acid/base, rather than alternative proposals of a substrate-assisted mechanism [66]. The analysis of QM simulations alongside crystal structures and NMR data, showed several significant features of the reaction coordinate in βPGM. The transition from a NAC III complex to a TSA complex, involves a concerted P–O bond dissociation, domain reorientation and proton transfer from D10 to the leaving group. The fact that domain orientational dynamics occur on a substantially slower timescale than bond vibrations suggests that domain reorientation instigates phosphoryl transfer. Simulations also show that the electrostatic repulsion between the transferring proton and the transferring phosphorus stabilises a planar transition state, which was experimentally corroborated by a subtle change in geometry in the AlF$_4^-$ moiety of the βPGM$_{D10N}$:AlF$_4$:G6P complex compared to WT [67].

1.2.8 Allomorphy

Substrate-free βPGM has two conformations (A and B) which exchange on a multi-second timescale, resulting from *cis-trans* isomerisation of P146 backbone amide, detected by ^{1}H^{15}N-TROSY NMR [68]. 30% of the assigned residues exist in slow chemical exchange, resulting in two spin systems for each backbone amide. These residues are found mostly in the core domain in the vicinity of P146. Five of these residues are missing in conformer B, owing to millisecond-exchange broadening (K145, A147, D149, I150 and Q176). The *cis* P146 isomer (conformer A) is adopted in the crystal structures, which places the adjacent residue K145 in the active site where it interacts with E169 and stabilises the negative charge of the proximal site phosphate (or phosphate analogue) [69].

The trans isomer (conformer B) is trapped in the βPGM$_{P146A}$ variant. A single species is observed in the ^{1}H^{15}N-TROSY spectra, which has chemical shifts which are similar to those of

βPGM$_{WT}$ conformer B. The crystal structure of the substrate-free βPGM$_{P146A}$ complex, shows a *trans* alanine with K145A displaced from the active site with the terminal amine oriented toward the cap domain. βPGM$_{P146A}$:MgF$_3$:G6P TSA complexes have almost identical chemical shifts to the corresponding WT complex. The crystal structure is almost identical to that of the WT structure and has a *cis* K145-A146 peptide bond with K145 side chain amine in the active site. Therefore, conformer A (*cis*) is catalytically active whilst conformer B (*trans*) has compromised activity. The steady state kinetic parameters (k_{cat}, $K_{m,\beta G1P}$ and $K_{m,\beta G16BP}$) are attenuated ~20-fold in βPGM$_{P146A}$ vs. βPGM$_{WT}$ [69].

Activation of the enzyme by phosphorylation of D8 can be achieved using different phosphorylating agents: Acetyl Phosphate (AcP), Fructose-1,6-bisphosphate (F16BP) or the reaction intermediate βG16BP. Activation using the latter results in conventional linear kinetics. AcP and F16BP both produces substantial lag-phases in their kinetics. The same pattern of behaviour is observed in βPGM$_{P146A}$. By monitoring the amide resonance of A113 (a well resolved reporter of conformer A and B, and their phosphorylated equivalents A^P and B^P) the effects of different phosphorylating agents (AcP, F16BP and βG16BP) on the population different phosphorylated βPGM species was measured. Critically, phosphorylation with F16BP and AcP results in population of a mixture of species including B^P, whilst phosphorylation with βG16BP does not generate any detectable B^P. A model was presented in which AcP and F16BP phosphorylate both conformers A and B whilst, βG16BP can couple phosphorylation with the transition to conformer A resulting in fast linear kinetics. The lag phase for AcP and F16BP is therefore a result of a significant population of the inactive B conformer. Only when βG16BP accumulates does the population of A start to dominate resulting in an increase in rate [69].

2 Characterisation of Hinge Dynamics in βPGM

2.1 Introduction

2.1.1 Conformational dynamics of globular domains

As described in Section 1.2, βPGM has two domains which open and close, to allow binding and release of substrates, during its catalytic cycle. Structural models describing the conformational dynamics of the globular domains in βPGM have come almost exclusively from x-ray crystallography. These models provide mechanistic insight as well as a structural basis for the interpretation of NMR data. It is therefore essential that we can understand the scope and limitations of these models. In particular, it is apparent from the structures of protein crystals that domain orientation is tightly coupled to the stability of the crystal lattice. It seems likely, at least in some cases, that domain orientation may shift to accommodate a more stable crystal lattice arrangement.

The dyndom databases [70–72] act as a repository of collective motions extracted from the protein data bank. A survey of this database revealed several other proteins which, like βPGM, are also comprised of two globular domains connected by a flexible linker which undergo a transition from a substrate-free open form to a substrate-bound closed form. From this survey, four proteins were identified which also had published data describing the conformational behaviour in solution. A case study of each system is described below to establish a precedence, if any, for the existence of crystallographic packing artefacts on domain orientation in proteins.

T4 Lysozyme

Lysozyme from the *Escherichia coli* T4 bacteriophage (T4L) belongs to the O-Glycosyl hydrolase family (EC. 3.2.1.17). It specifically hydrolyses the β1-4 glycosidic link between alternating N-Acetylglucosamine (NAG) and N-Acetylmuramic acid (NAM) residues within the peptidoglycan polymer of gram-positive bacteria. T4L is a small globular protein (18.6 kDa), comprising 164 residues. It has two domains: C-terminal domain (C-domain) and N-terminal domain (N-domain). The substrate binds in the active-site cleft between the two domains. The N domain has three contiguous antiparallel β-strands, connected by short loops followed by a helix. The C-domain comprises a bundle of helices. The two domains are connected in sequence at two points: within an N-terminal helix (α1) and at the C-terminal end of α3. There are over 1000 T4L structures deposited in the protein data bank. The vast majority populate a relatively compact state, irrespective of active site occupancy. It was suggested the enzyme would need to adopt a substantially more open form to allow the relatively large substrate to enter the active site cleft [73]. Different crystal isoforms of T4L were obtained, using varied crystallisation conditions and T4L variants, which showed substantially more open structures [74]. Furthermore, Crystallization of T4L fused to a polymeric scaffold yielded low resolution structural information which was more open than the predominant closed crystal isoform [75]. ^{1}H ^{15}N residual dipolar couplings (RDC) were measured to establish the ensemble average orientation in solution. [76] The rigid body transformation of the N-domain which best fit the solution data was expressed using Euler Angles. The results show a 17° rotation relative to the closed conformation (PDB: 3LZM). The best-fit orientation was close to that of the open crystal structure (PDB: 150L, chain C). RDCs and pseudo-contact shift (PCS) measurements on paramagnetically labelled T4L were fit to data calculated from a range of crystal structures. The best fitting structure was found to be 1SSY which is slightly more open than 150L(c). Furthermore, paramagnetic relaxation enhancements (PRE) predicted from closed crystal structure were not observed, ruling out an equilibrium between open and closed conformations [77]. Single molecule fluorescence resonant energy transfer (smFRET) experiments revealed anticorrelated fluctuations in donor-acceptor emission during enzymatic turnover of substrate, corresponding to conformational dynamics of the enzyme. No such motions were detected in the absence of substrate. This indicates a single open species is populated in the absence of substrate. However, the FRET signal is only semi-quantitative and could not be accurately related to structure [78]. “Single-walled carbon nanotube, field effect transistor” SWNT-FET sensors were also used to measure single molecule dynamics. The conductance of the SWNT-FET sensor is affected by conformational changes in a single bound lysozyme molecule. A featureless, monomodal signal was observed for acquisition in the absence of substrate or using inactive mutants (E11H, T26E), whereas a two-state switching was observed in the presence of different substrates [79].

Phosphoglycerate Kinase

Phosphoglycerate kinase is a glycolytic enzyme, catalysing phosphoryl transfer from the substrate 1,3-bisphosphoglycerate to adenosine diphosphate (ADP), generating 3-phosphglycerate and adenosine triphosphate (ATP). It is also an important processor of anti-retroviral pro-drugs. PGK has been studied using a variety of source organisms including human, plasmodium, staphylococcus, and mouse. These variants have a largely conserved tertiary structure. PGK has around 400 residues and is approximately 41 kDa. It has an N-terminal and C-terminal domain of similar size. Both domains have a central parallel β-sheet surrounded by helices. The domains are connected in two places, by a rigid helix and more flexible C-terminal loop. 1,3-bisphosphoglycerate binds to the N-domain whilst ADP binds to the C-domain. A range of human PGK complexes have been identified including apo, binary, ternary and TSA complexes [80] with a range of domain orientations. Open crystal structures are rotated by 20-30° relative to the fully closed TSA conformation. The domain orientation of a substrate free homology model were refined against small angle x-ray scattering data. The refined model was used to generate an electron density map. The substrate free structural model was then fit to the electron density map using ‘deformable elastic network’ restraints. The resulting structure was substantially more open, with a domain orientation rotated 56° relative to the TSA conformation [81]. Furthermore, molecular dynamics on the nanosecond timescale have interdomain centre-of-mass distances up to 39 Å which are greater than those of the open crystal structures (32 – 35 Å) [82]. Elastic network models were used to calculate normal modes which describe hinge bending in PGK. The amplitudes and timescales of these normal modes were fit to small angle neutron scattering (SANS) and neutron spin echo (NSE) data. The results indicate sampling of open and fully closed domain orientations with an average orientation which is similar to the crystal structures, on a nanosecond timescale. Substrate binding shifts the average orientation to a more closed complex and reduces the amplitude of fluctuations [83]. These observations are consistent with single molecule FRET experiments. FRET efficiency histograms for substrate-free PGK show two distinct populations of a more open and more compact conformation. These states do not exchange within the confocal dwell time (τ_{ex} > 10 millisecond timescale). The expanded state comprises a broad distribution of open conformations which also exhibit > 10ms exchange rates. These states are much more open that open crystal structures and may be partially unfolded. Conversely, the compact state is more similar to the open crystal structures but exhibits large amplitude fluctuations on the nanosecond timescale, with transient sampling of fully closed conformation [84].

Glutamate dehydrogenase

Glutamate dehydrogenase (GluDH) catalyses the reduction of glutamate to 2-oxoglutarate and ammonia using NADH as a reducing agent. GluDH has a hexameric structure. Each subunit (~46 kDa) has two domains of roughly equal size: a nucleotide-binding N-domain and a C-domain which contacts other subunits. The C-domain has a mixed (anti)parallel β-sheet surrounded by helices, whilst the N-domain has a twisted β-sheet with long strands packed against α helices on one side. The active site sits in the cleft between the two domains. Crystal structures of GluDH from *Thermococcus profundus* (PDB: 1EUZ) are typical of structures from other open GluDH variants. SAXS data showed significant deviations compared to the predicted scattering curves from the crystal structures. The Radius of Gyration of the solution structure is larger by ~2 Å. It was suggested that crystal packing forces bias the orientation of the N-domains to more a compact state [85]. A 200 ns molecular dynamics simulation of the hexameric GluDH complex illustrates ns-timescale fluctuations in domain orientation, spanning the full range of conformers sampled by the x-ray crystal structure (PDB: 1EUZ). AFM was used to measure the topography of a GluDH crystal, in which the N domain of subunit D was free from crystal contacts. An AFM height distribution showed two species with width σ ~3.4 Å separated by Δμ ~ 3.6 Å. An equivalent height distribution was obtained from the MD simulations, which showed two populations were sampled with a mean change in height Δμ ~ 2.9 Å, slightly less than that derived by AFM. [86]. Structures of the six C-domains by cryo-electron microscopy (cryoEM) were calculated with a resolution better than 3 Å, whereas the N-domains were worse than 3.2 Å with arc-shaped electron-density following the direction of domain reorientation. Distinct electron density maps were obtained with the N-domain in different orientations by re-classifying particles. A weighted sum of MD simulation structures (described above) could be fit to the electron density maps. Refined structures from each electron density map cover a similar range of motion compared to the crystal structures (PDB 1EUZ). [87].

Periplasmic binding proteins

Periplasmic binding proteins (PBPs) are a superfamily of proteins involved in chemotaxis and solute uptake in gram-negative bacteria. They are diverse in sequence but have a common tertiary structure, comprising two domains connected by a flexible linker. Solute binding occurs at the interface between the two domains, which causes substantial domain reorientation. A 500 ns MD simulation of ligand-free maltose binding protein (MBP) from *Thermotoga maritima* shows a single open population, which was fit to SAXS data using an implicit solvation model [88]. Crystal structures of ligand-free complex of MBP from *E.* coli were shown to be in close agreement with solution RDC data [89]. PREs are sensitive to the presence of minor species in fast exchange. Measurements of spin-labelled ligand-free MBP indicate a 5% population of a partially closed conformation [90]. This state was identified using accelerated molecular dynamics (aMD) [91]. Two state exchange was observed on a timescale of seconds, in smFRET experiments, which were attributed to this dynamic exchange [92]. PRE measurements in a closely related system (Glutamine binding protein) did not detect a minor, partially-closed species. [93,94]. Whereas crystal structures of a ligand-free closed complexes have been found for Galactose/Glucose binding protein (GGBP) [95] and choline binding protein (ChoX) [96]. A number of molecular dynamics studies have shown substantial domain reorientation in the ligand-free state compared to crystal structures [97–99]

Summary

Three key points can be understood from these case studies. 1) There is a precedence for biasing of domain orientation in crystal structures. Solution measurements are sometimes required to obtain a more accurate ensemble. 2) Experiments which probe domain orientation in solution can have conflicting results owing to low sensitivity. Careful calibration of experiments and calculations are required. Clear quantification and visibility of uncertainties are advisable to avoid such inconsistencies. 3) The balance between the open/closed state in ligand-free complexes is quite variable between these proteins. Some exclusively populate an open conformation, whilst others transiently (or otherwise) sample more closed conformations.

The following sections describe an investigation into the possible effects of crystal packing on domain orientation in βPGM. Two approaches are available. New structures can be derived from existing crystal structures through refinement against experimental data, e.g. the use of SAXS data to refine a deformable elastic network model [81,100] or RDC data to refine the rigid body domain orientation [76]. Alternatively, a new ensemble of structures can be derived using an MD forcefield, which is then validated against experimental data [101]. The performance of MD simulations has improved significantly in recent years owing to GPU-accelerated calculations, forcefield development and algorithmic improvements. Fluctuations in domain orientation occur largely on a ns-μs timescale, which are now routinely accessible by MD simulations. Therefore, it is this approach that we have chosen to use to investigate domain dynamics in βPGM.

2.1.2 MD Simulations – Theory and Developments

Theory

Molecular dynamics (MD) is a type of computer simulation used to predict the motions of atoms and molecules [102–105]. A system of interacting atoms and molecules are defined, and their trajectories are derived by numerically solving Newton's classical equations of motion. MD forcefields specify the interactions between the atoms in a molecular system. A forcefield has a functional form and a set of parameters. The functional form is a mathematical expression describing a general relationship between the potential energy and the atomic coordinates. The parameters define the specific quantitative relationships. There are a few standard classical potentials used to describe molecular systems .

Covalent bonding is typically described using a harmonic potential for bond distances and angles.

$$E_{\mathrm{length}} = \frac{1}{2}k(x - x_0)^2$$

$$E_{\mathrm{angle}} = \frac{1}{2}k(\theta - \theta_0)^2$$

A periodic cosine potential and the Ryckaert-Bellemans potential are used to describe torsions.

$$E_{\mathrm{torson}} = k(1 + \cos(n\theta - \theta_0))$$

$$E_{RB} = \sum_{i=0}^{5} c_i \cdot \cos^i(\theta)$$

Canonical dihedral angles are calculated between 4 consecutively bonded atoms. Improper dihedral angles are calculated between 3 atoms covalently linked by a 4th central atom. Improper dihedrals are useful for ensuring proper geometry for certain functional groups (e.g. maintaining planar peptide bonds).

Non-bonded interactions are described by the Lennard-Jones potential,

$$E_{LJ} = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right)$$

where r is the interatomic distance, σ is the interatomic distance where the potential is zero (this is often expressed as $r_{\mathrm{min}} = \sigma \cdot 2^{1/6}$), ϵ is the dispersion energy.

Electrostatic interactions are described by the coulomb potential,

$$E_Q = \frac{q_1 q_2}{4\pi\epsilon_0 r}$$

Where q_1 and q_2 are the charges of the interacting atoms and ϵ_0 is the dielectric constant.

Several integration methods exist to calculate trajectories using forcefields. Common examples include *Velocity Verlet* and *Langevin integrators*. An important setup step in MD simulation is minimisation of the initial coordinates. Typically coordinates will be derived from solved structures deposited in the protein data bank. These structures usually have atoms that are slightly too close or too far from neighbouring atoms (as defined by the forcefield) which will give rise to large forces. Evaluating large forces with long timesteps give rise to non-physical dynamics and 'explosions' in energy and temperature which will ultimately cause the simulation crash. To overcome this, structures are 'energy minimized'. The aim of this procedure is not to obtain an absolute energy minimum but to shift the structure away from conformations which give rise to very large derivatives in the potential. A commonly used algorithm is called 'steepest descent', whereby the gradient (m) of the potential is calculated, and each atom is incremented along the direction of minimum gradient, i.e. max(-m). After a finite number of increments the system should find a local minimum where the forces are close to zero. Minimisation usually only results in small structural changes [106].

Different types of ensembles can be collected in MD simulations. These are named according to the physical properties that are kept constant. The NVE (aka. Microcanonical) ensemble keeps the number of particles, the volume, and the energy of the system constant. This is not normally the best choice for MD simulations of large biomolecules, because conformational changes which reduce the potential energy of the system will coincide with a significant increase in kinetic energy, which can then give rise to different dynamic behaviours. A better alternative is the NVT ensemble, which keeps the temperature rather than the energy constant. Different computational thermostats work in different ways, but in simple terms this can be done by scaling the velocities of the particles in your system incrementally in response to changes in kinetic energy. Depending on how water molecules are placed around the protein, there may regions of high or low pressure within the system. The NPT ensemble involves the use of a computational barostat, which modifies the volume of the system to maintain a constant pressure. An NPT equilibration step is typically used to allow the protein to come to equilibrium with the solvent without creating abnormally high or low pressures. NVT and NPT equilibration steps are standard practise in the setup of MD simulations [103,104].

Periodic boundary conditions are frequently utilized to avoid surface tension effects in explicit solvent models or to avoid having very large water boxes with a very large number of atoms. They allow approximation of the behaviour of very large (infinite) systems without incurring a large computational cost. A simulation with PBCs has an infinite lattice of identical simulation boxes. If an atom reaches a boundary, effectively it will re-enter the same box on the opposite side of the cell. According to the minimum image convention, the box must be at least as large as the interaction cutoffs (see below) so that a molecule will only interact with one copy of itself in each dimension. Otherwise, self-interactions would cause the velocities to 'explode' [103,104].

Several methods are used to reduce the computational cost of MD simulations. In explicit solvent simulations of biomolecules the number of solvent atoms usually far exceeds the number of atoms in the biomolecule of interest. Consequently, water molecules are often made rigid (fixed bond lengths and bond angles) to minimise the cost of computing their dynamics.

Evaluating the forces is the most expensive part of each simulation timestep. Therefore, by increasing the timestep you reduce the number of timesteps you need to calculate and reduce the cost of the simulation. The timestep must be smaller than the fastest dynamical fluctuation of the simulation. The fastest fluctuations would normally be bond vibrations involving hydrogen atoms. However, by fixing the bond lengths and bond angles of water molecules, you remove these dynamics which allows you to increase the timestep of your simulation. Furthermore, masses of heteroatoms can be re-distributed to their bonded hydrogen atoms which reduces the bond vibration and bending frequencies, allowing timesteps up to 4 fs [107].

The long-range coulomb and LJ interactions are the most expensive part of the forcefield to evaluate. Coulombic interactions depend on $\frac{1}{r}$ resulting in slow convergence of the force as a function of the distance. Ewald summation methods are often used which separate the potential into a short-range part which is evaluated in real space and a long-range part which is evaluated in Fourier space [108]. Convergence of the potential in Fourier space is much faster allowing for a quick and accurate evaluation of the potential. Details of this calculation are described elsewhere [109].

Forcefields

Classical MD forcefields make many approximations. These models are atomistic – each atom is described as a particle with a fixed mass and charge. QM effects (e.g. polarization, quantum orbital effects) are not accounted for explicitly due to computational cost. Forcefields are based on empirical observation and are parameterized using experimental data and *ab initio* calculations . The information contained within such a model cannot be expected to accurately describe all varieties of molecular behaviour in all environments. Instead, it is designed to try and accurately account for the behaviour of specific types of molecules in specific environments and in specific structural/dynamical regimes. This is in analogy with the mathematical field of 'approximation theory' where complex functions can be reduced to simpler ones within specific limits or under certain assumptions. To increase the scope of such simulations, new bespoke parameter sets can be derived to describe the behaviours of different types of molecules cheaply and accurately (i.e. proteins, lipids, nucleic acids) [110]. Furthermore, within protein forcefields each amino acid has a unique set of parameters which helps describe the nuances of their behaviour. It has also been noticed that modern protein forcefields do not accurately describe the behaviours of intrinsically disordered proteins. New parameter sets are being developed to describe this alternative dynamic regime.

A few well-known families of forcefield exist: Amber, CHARMM, GROMOS, OPLS. Amber's ff14sb protein forcefield [111] is one of the most recent forcefields at the end of a long history of development and optimisation [112]. Some notable developments included the gradual improvement of the representation of secondary structure elements and side chain rotamer preferences ultimately through fitting to more sophisticated and thorough *ab initio* models and to more diverse and extensive experimental data sets. ff14sb has recently been succeeded by ff19sb which highlights an inherent underestimation of helical propensity which is inexactly compensated for by the propensity of the tip3p water model to bias toward more compact structures. The ff19sb model claims to capture more sequence specific behaviours [113].

Typically, forcefield development has focussed on optimising specific sets of torsion angles whilst the bond interaction parameters, non-bonded interaction parameters and partial charges are often left unoptimized. The ff15ipq forcefield is the second in a new generation of Amber forcefields, which have rederived implicitly polarised atomic charges using the "IPolQ" method. This method fits charges to QM-derived charge distributions calculated in a vacuum and in a field of point charges derived from an MD simulation using a specific water model. 60 new bond angle parameters and 900 new torsion parameters were derived, with several new atom types to decouple the behaviours of different residue types. Finally, new empirically derived Lennard-Jones radii were calculated for polar hydrogen atoms bonded to a nitrogen atom, which better predicted the propensity to form salt bridges compared to contemporary forcefields. This forcefield has been demonstrated to provide better agreement with NMR data [114] and is better able to reproduce hydrogen bonding networks from crystal structures [115].

Over-binding of metal ions to proteins is another well-documented problem in classical MD forcefields. One approach to address this problem is the 'electronic continuum correction' which implicitly accounts for electronic polarization effects in a mean-field way for ions in solution. Typically this is introduced either through charge scaling or by modification of the dielectric constant. This has proved useful for the modelling of ion transporters [116] and calcium binding proteins [117,118]. Charge scaling has also been used to improve the representation of salt bridges in proteins [119].

Sampling

Simulation performance depends upon the size of the system being simulated, the platform on which the simulation software is being run, the available computational resources, the complexity of the forcefield being evaluated and the number of time-saving approximations being used. The availability of fully programmable, massively parallel graphics processing units (GPUs) has improved performance in many areas of scientific computation. GPUs have enabled routine sampling of µs dynamics in MD simulations [120]. With speeds of 360 ns/day it would take about one month to run a 10 µs simulation. Access to these speeds can also be attributed to software/algorithmic improvements [121]. Millisecond timescale unbiased simulations are becoming accessible for smaller simulation systems [122].

Many functionally relevant dynamical processes occur on millisecond – second timescales. Sampling of dynamics on these timescales would require vast computational resources, using conventional molecular dynamics for larger systems. Enhanced sampling algorithms are a diverse collection of methods used to try and accelerate sampling of conformational dynamics so that they occur on computationally accessible timescales. Adaptive-bias methods such as metadynamics are amongst the most widely used [123]. During a metadynamics simulation, gaussian potentials are deposited to push the system away from previously visited states. These potentials are applied to a user-defined set of “collective variables” (CVs) i.e. some function of the coordinates. The aim is to select a CV which discriminates between each of the metastable states within the conformational landscape and the barriers between them. These may be torsions, angles, distances, global properties such as radius of gyration, a linear combination of multiple CVs such as those identified by PCA [124] or some measurable property such as a SAXS curve [125]. If a system is trapped in a stable conformation the potential will accumulate, gradually destabilising the conformation until transition to other conformations becomes kinetically accessible. The metadynamics bias is described by the following equation.

$$V(s,t) = \sum_{\kappa\tau<t} W \cdot e^{-\frac{(s-s(\vec{r}(\kappa\tau))^2}{2\sigma^2}}$$

Where s is the collective variable(s), t is simulation time, τ is the deposition timestep, κ is the number of gaussian kernels deposited, σ is the width of the gaussian bias and W is the gaussian height. The bias at time t is therefore taken as a sum of gaussian kernels centred on the CV coordinates defined by the simulation. Well-Tempered Metadynamics is a variant in which the height of the deposited bias depends on the size of the bias already deposited at the current position in CV-space.

$$W = W_0 e^{-\frac{v(s,t)}{k_B T}}$$

$$\mathrm{T_S = T + \Delta T}$$

$$\mathrm{B = \frac{T + \Delta T}{T}}$$

The bias factor B determines the effective temperature at which the CVs are explored. This provides a mechanism for tuning the degree of acceleration to apply to the simulation. The biased potential energy landscape converges to a landscape where the transition barriers are attenuated. Metadynamics has been used successfully in many systems, however there is a limit on the number of CVs that can be efficiently biased.

2.2 Results: MD Simulations

2.2.1 βPGM Crystal Contacts

Crystal lattice contact mapping was used to investigate the influence of crystal packing on the conformational landscape of βPGM. Pymol's `symexp` function was used to reconstruct neighbouring asymmetric units for each of the βPGM *Lactococcus lactis* PDB crystal structures. A Euclidean distance matrix was calculated between the coordinates of central copy and those of the neighbouring molecules in the crystal lattice. Distances less than 3.5 Å were counted, cumulatively, for each residue and for each structure. These are represented on a contact map in **Figure 2-1**. The vast majority of open βPGM crystal structures have either a $P2_12_12_1$ symmetry with one chain in the asymmetric unit, or a $P12_11$ symmetry with two chains in the asymmetric unit. A representative example of each is given by 2WHE and 6HDI, respectively. Any residues of neighbouring molecules containing atoms within 5 Å of the central copy of the asymmetric unit are highlighted in **Figure 2-2**. This analysis shows both structures have extensive contacts around the cap and core domains, which restrict their accessible range of motion. This does not necessarily mean that the observed orientations are biased, only that that an alternative orientation would require a different crystal packing arrangement. Notably, contacts are also made with the hinge region (S88-I84) and with residues in the 70s helix including residues N77 and N118 which interact with K109-N_Z and G159-O from a neighbouring molecule.

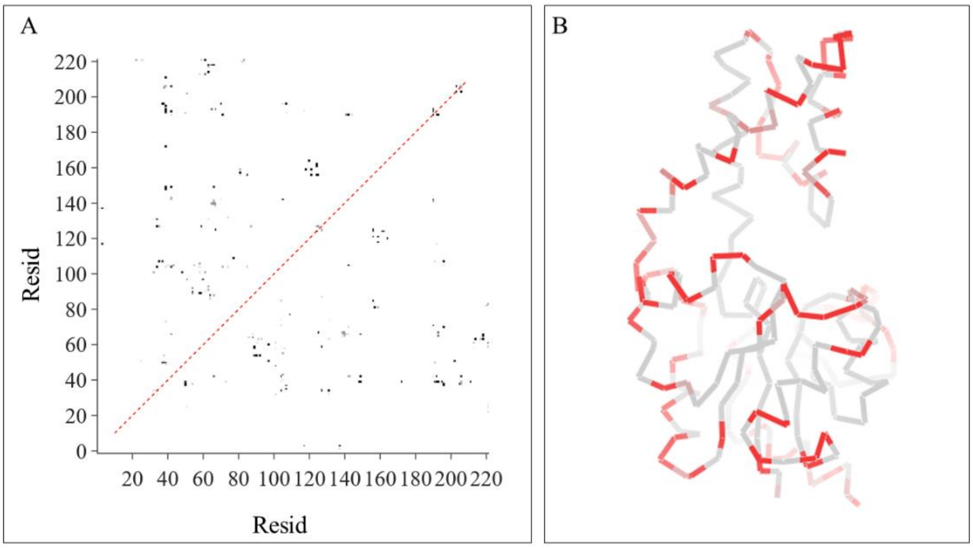


Figure 2-1 – βPGM crystal packing contacts. (A) A per-residue, crystal lattice contact map of open βPGM structures. Residues with atoms within 3.5 Å of one another are highlighted in the contact map. The greyscale represents the number of occurrences of a given residue-residue contact. (B) The total number of contacts per residue are plotted on an open βPGM crystal structure (PDB: 2WHE).

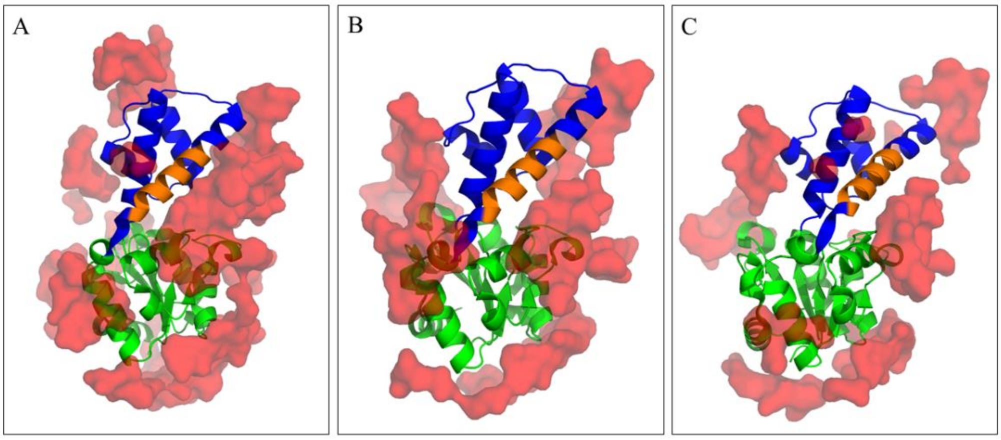


Figure 2-2 – βPGM crystal packing contact examples. βPGM crystal structures (A) 2WHE (B) 6HDI chain A and (C) 6HDI and chain B are displayed as a cartoon representation with cap and core domains coloured blue and green respectively. The C-terminal portion of the 70s helix, between N77 and the hinge is coloured orange. All residues from a neighbouring asymmetric unit within 5 Å are shown as a translucent red surface, illustrating the crystal contacts which limit the accessible range of motion of the cap and core domains.

2.2.2 Molecular Dynamics

An extensive series of MD simulations were run with a cumulative total of 40 μs simulation time, to investigate solution hinge-bending dynamics in βPGM. Two forcefields were used: the ff14sb protein forcefield paired with tip3p solvent and the ff15ipq protein forcefield paired with SPEC/E_b solvent. In some cases, an electronic continuum correction applied to SPEC/E_b monoatomic ions. This was implemented as 75% scaled charges ($Mg^{1.5+}$, $Na^{0.75+}$, $Cl^{0.75-}$) with corresponding Lennard-Jones parameters (ϵ, σ) empirically determined using neutron scattering data [126–128]. All but one of the simulations were initiated using coordinates from the βPGM_{WT}:MgF_3:G6P transition state analogue complex (PDB: 2WF5), with ligands and selected water molecules removed. The other was initiated from a substrate-free βPGM_{WT} structure (PDB: 2WHE). All but one of the simulations were collected as a single production run, between 5 and 10 μs. One ensemble was calculated as 5 × 1 μs production runs initiated from randomly selected coordinates and velocities. An overview of the simulations is given in **Table 2-1**.

Simulation Name	Initial Coordinates	Protein Force Field	Solvent Model	ECC	P146 (cis/trans)	Length (μs)
ff14sb-tip3p-1	2WF5	FF14SB	TIP3P	✗	cis	10
ff14sb-tip3p-2	2WF5	FF14SB	TIP3P	✗	cis	5
ff15ipq-spceb	2WF5	FF15ipq	SPC/Eb	✗	cis	5
ff15ipq-spceb-5×	2WF5	FF15ipq	SPC/Eb	✗	cis	5×1
ff15ipq-spceb-ecc	2WF5	FF15ipq	SPC/Eb	✓	cis	5
ff15ipq-spceb-ecc-meta	2WF5	FF15ipq	SPC/Eb	✓	cis/trans	5
ff15ipq-spceb-ecc-2whe	2WHE	FF15ipq	SPC/Eb	✓	cis	5

Table 2-1 – Summary of molecular dynamics simulations. The initial coordinates sets are defined using the protein data bank ID code. Amber’s protein forcefields were used with their corresponding solvent model. For some simulations an electronic continuum correction (ECC), implemented as 75% charge scaling and modified Lennard-Jones parameters, was applied for monoatomic Mg, Na and Cl ions. One of the simulations included a metadynamics bias to introduce cis/trans isomerisation of the K145-P146 peptide bond. Simulations were implemented as either single long production run, or multiple short production runs.

The intrinsic Euler angles (yaw, pitch, roll) describing domain orientation were calculated throughout each trajectory. The pitch angle corresponds to the opening/closing motion of the cap domain; the yaw angle corresponds to a ‘left-to-right shearing motion’ and the roll angle corresponds to a twisting motion around the long axis of the cap domain. Euler angles were calculated relative to the fully closed βPGM$_{WT}$:MgF$_3$:G6P transition state analogue conformation (PDB: 2WF5). Full details of the calculation are described in Section 5.2.4. The yaw-pitch angle distributions for all the simulations combined are shown in **Figure 2-3**. Four conformations (c_1, c_2, c_3 & c_4) were identified from these distributions (**Table 2-2**). These simulations suggest the solution ensemble exist in multi-state μs-exchange between conformations that are all more open than the crystal structures.

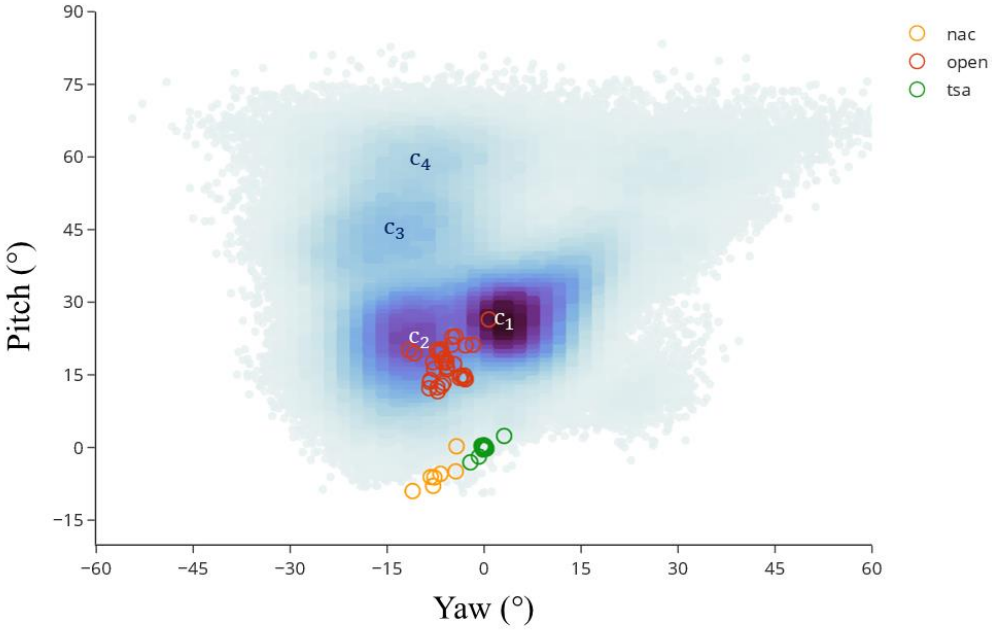


Figure 2-3 – Cap Domain Euler Angles. The yaw and pitch angles calculated from all MD simulations specified in Table 2-1 were separated into 100×100 discrete bins and coloured according to the frame count for each bin, with lower counts coloured light blue and higher counts coloured dark purple. The angles calculated for the open, partially closed and fully closed TSA-like crystal structures are represented with open-circles and are coloured red, orange and green, respectively. The data are shown for each simulation individually.

Conformer	**R_g (Å)**	**Pitch (°)**	**Yaw (°)**
xtal	18.7	17	-6
c_1	19.0	26	3
c_2	19.2	26	-13
c_3	20.0	45	-15
c_4	20.0	61	-9

Table 2-2 – Cap Domain Euler Angles. The average Radius of Gyration, Pitch angle and Yaw angle calculated for the open βPGM crystal structures (xtal) and for each of the conformers identified from MD simulations (c_1 – c_4).

2.2.3 Conformer c_2

The ff15ipq-spceb-ecc-2whe simulation is dominated by conformer c_2. During the simulation the structure relaxes to a new conformation which is, on average, rotated by ~9° in pitch angle away from the open crystal structures. A distinct feature of this conformational change involves the packing interactions of Y19. An internal cavity is present in the crystal structures formed by residues from the core domain (L9, T14, L122, M126), the cap domain (Y19, Y80, I84) and the hinge (D15, T16, V87). In the ff15ipq-spceb-ecc-2whe simulation, this cavity is filled by the sidechain of Y19 (**Figure 2-4**). The distance between Y19-C_ζ and L122-C_γ reports on the occupancy of the hinge cavity by Y19 sidechain and is reduced from 8.3 Å in 2WHE to 5 Å in the MD ensemble (**Figure 7-2**). Y19 sidechain packing interaction is initially satisfied by a hinge bending motion, resulting in a ~40° pitch angle, which is unstable and does not persist for more than 100 ns. A more stable conformation is found after ~1.5 μs, involving a subtle rearrangement of multiple hinge residues and a shift to the c_2 domain orientation. Significant changes are observed in the backbone torsions of T14, D15, T16, V81 and K82. Conversely, the dihedral angles for M83-S88 overlay closely with the crystal structures (**Figure 2-5**). Furthermore, the backbone carbonyl of K117 flips to form a hydrogen bond with Y80-O_η, which has a stabilising effect on the domain orientation.

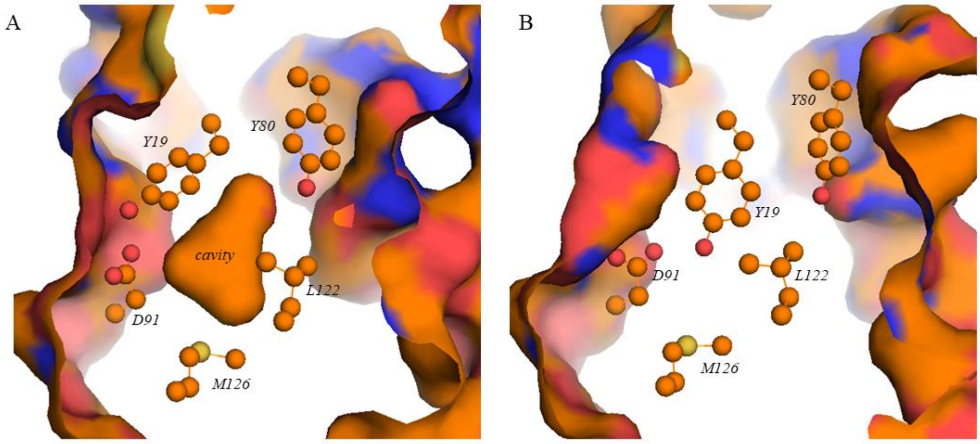


Figure 2-4 – A hydrophobic cavity in substrate-free βPGM. (A) Shows a slice through the Van Der Waals' surface of the βPGM_{WT} substrate free crystal structure (PDB: 2WHE), revealing a hydrophobic cavity within the hinge region defined by the side chain atoms of Y19, Y80, D91, L122 and M126 (ball-and-stick representation) and L9, T14, D15, T16, I84 and V87 (not shown). (B) Shows the same slice from a representative structure of the c_2 domain orientation from the ff15ipq-spceb-ecc-2whe simulation. The cavity is occupied by the sidechain of Y19.

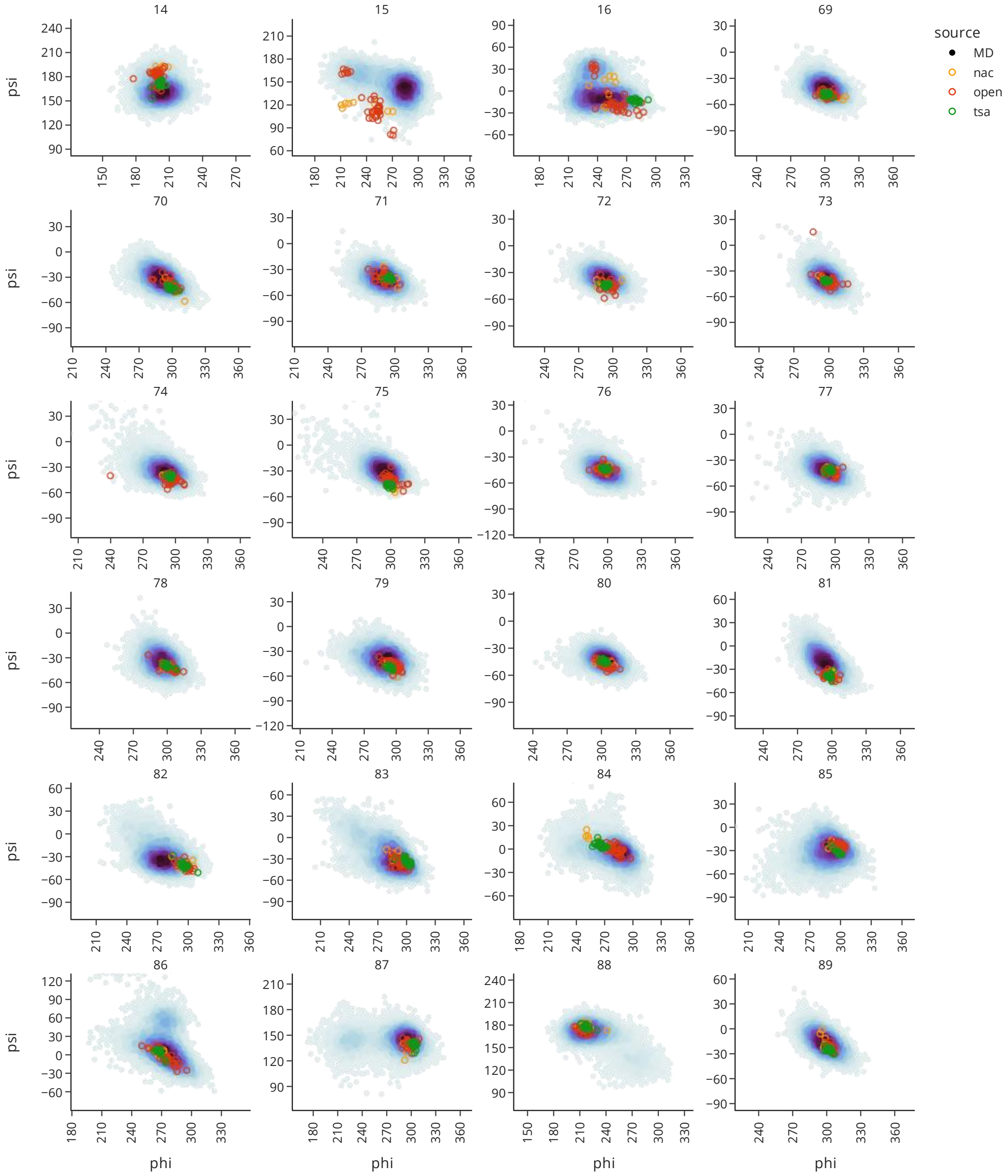


Figure 2-5 – c₂ Ramachandran distributions. Ramachandran (ϕ, ψ) distributions for hinge residues (T14, D15, T16, V81, K82, M83, I84, Q85, D86, V87, S88 and P89) and for residues at the base of the 70s helix from F69 – Y80. Simulation frames were separated into 100×100 discrete bins and coloured according to the frame count for each bin, with lower counts coloured light blue and higher counts coloured dark purple. The angles calculated for the open, partially closed and fully closed TSA-like crystal structures are represented with open-circles and are coloured red, orange and green, respectively. Angles are expressed in an atypical periodic boundary to aid visualisation.

The open crystal structures comprise two distinct hinge conformations, resulting from rotations about D15-ϕ and T16-ψ (**Figure 2-5**). The dominant crystal conformer is characterised by a direct hydrogen bond between D10-O_δ and T16-N and a more *compact* conformation of the backbone between D15 and T16. The minor conformer (PDB: 1LVH, 5OLW, 6H8U, 6H8V, 6YDL) has a more *extended* backbone conformation, resulting in a water-mediated hydrogen bond between D10-O_δ and T16-N (**Figure 2-6**). This conformational change does not have a substantial effect on the domain orientational Euler angles. Instead it appears to cause a subtle shear/translational motion of the cap domain. Also, the extended conformation places the side chain of T16 into the hydrophobic cavity within the hinge region.

An equivalent behaviour is observed in the MD simulations. The major/compact conformation has similar T16 backbone torsions to the corresponding crystal structures but has distinct D15 torsions, associated with a shift to the c_2 domain orientation. Transition to the extended conformation stretches the hydrogen bond between D10-O_δ and T16-N. However, the interaction is not maintained via a water molecule. Instead, without hydrogen bonding constraints, D10 can freely rotate about χ_1 and coordinate the catalytic magnesium ion. Thereafter, it remains trapped. The D15-ϕ and T16-ψ backbone torsions can collapse back to the *compact* arrangement, with a shift in T16-ϕ resulting from the lack of hydrogen bonding restraints imposed by D10. A similar magnesium-bound D10 rotamer is adopted if the simulation is initiated from a closed complex (2WF5). In this case, D10-O_δ is oriented to coordinate G6P 1-OH group which has been stripped from the active site prior to the start of the simulation. This conformation is not stable and depending on the initial velocities it may adopt a canonical *out* conformation or it will bind irreversibly to the catalytic magnesium. Using initial coordinates from an open, substrate-free crystal structure (PDB: 2WHE), D10-O_δ starts in its canonical *out* position, which persists for several μs before rotating to interact with the catalytic Mg ion.

No experimental evidence has been found to support the existence of a direct interaction between D10-O_δ and the catalytic magnesium. The existence of this conformation in MD simulations is not surprising due to the tendency of contemporary forcefields to exaggerate electrostatics, particularly for atomic ions which retain their full formal charge and are not subject to charge transfer or polarisation effects. Adoption of the D10 *out* or Mg-bound rotamer appears largely independent of forcefield. Use of the FF15IPQ forcefield with an electronic continuum correction did not preclude the adoption of the magnesium bound rotamer. Whilst the Mg-bound rotamer would likely be destabilised in a simulation with scaled charges, sampling of this conformational change in our μs-timescale simulations was insufficient to quantify this effect. This behaviour is likely an artefact of the MD forcefield and is associated with other dubious hinge conformations.

2.2.4 Conformer c_1

The hinge is defined by an extensive network of packing interactions between three distinct regions of the primary sequence: the N-terminus of α1 (L9, D10, T14, D15, T16, Y19, H20); the C-terminus of α4 (Y80, V81, I84, Q85, V87, D91); and α6 (N118, F121, L122, R125, M126). In particular, I84-C_δ normally sits in a pocket formed by N118, F121, L122. However, the free space yielded from the displacement of D10 sidechain, allows a shearing motion of the cap domain, corresponding to a 15° rotation in yaw angle, which coincides with a rearrangement of the packing interactions between the two domains. I84 sidechain rotates around χ_1 and C_δ drops into a new pocket between L122, T16, and V87, displacing Y19 (**Figure 2-7**). This conformational change has characteristic backbone torsion angles of I84, Q85 and V87. The largest change is exhibited by I84, which rotates by -20° in ϕ and ψ. Subtle torsional changes also occur in T14, D15 and T16. Notably, the heterogeneity in D15 and T16 Ramachandran-space is lost as the extended hinge conformation is precluded (**Figure 2-8**).

Adoption of the Mg-bound D10 rotamer appears to be necessary but not sufficient for transition to c_1. Each of the four simulations which populate conformer c_1 also adopt the Mg-bound D10 rotamer – the packing arrangement in the hinge region for c_1 is incompatible with a D10 *out* conformation due to steric clashes. The ff14sb-tip3p-2 simulation demonstrates multiple attempts to transition to c_1 which are unsuccessful because D10 is in the *out* position. Conversely, whilst c_2 is associated with the D10 *out* conformation, adoption of the Mg-bound D10 rotamer does not always lead to the adoption of c_1.

This conformation is stable on a μs-timescale. It is populated in three simulations: ff14sb-tip3p-1, ff15ipq-spceb-ecc and ff15ipq-spceb-ecc-meta (**Table 2-1**). Transition from c_1 back to c_2 requires a shift in the dihedral angles of the hinge residues to lift I84-C_δ out of the cleft between L122, T16, and V87. Furthermore, I84 sidechain must rotate back around χ_1 to allow C_δ to drop back into the pocket formed by N118, F121, L122. This transition often coincides with the population of the more open c_3 conformation.

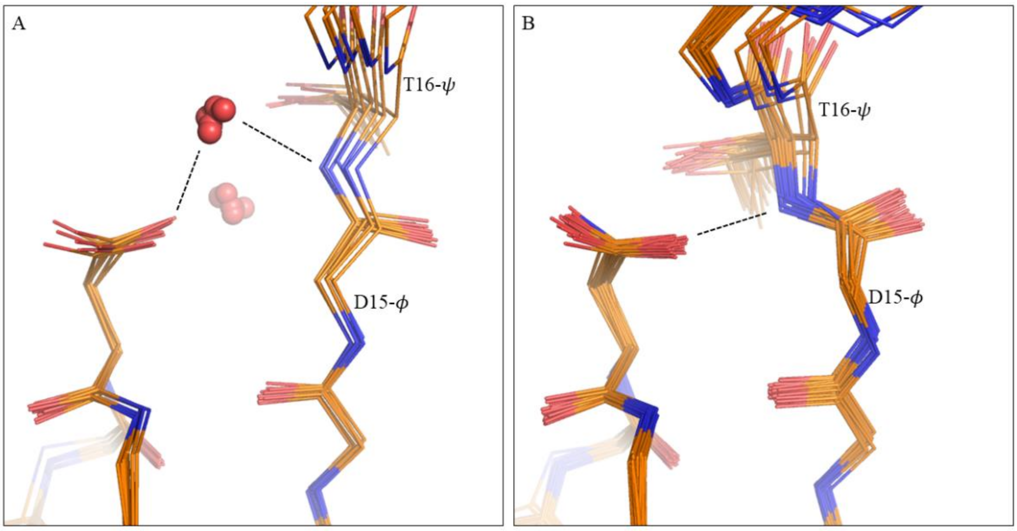


Figure 2-6 – Hinge conformations in βPGM. (A) Open crystal structures of βPGM populating an extended hinge conformation characterised by a water-mediated hydrogen bond between D10-O_δ and T16-N. (B) Open crystal structures of βPGM populating a compact hinge conformation characterised by a direct hydrogen bond between D10-O_δ and T16-N. The torsions which give rise to this conformation change (D15-ϕ and T16-ψ) are labelled.

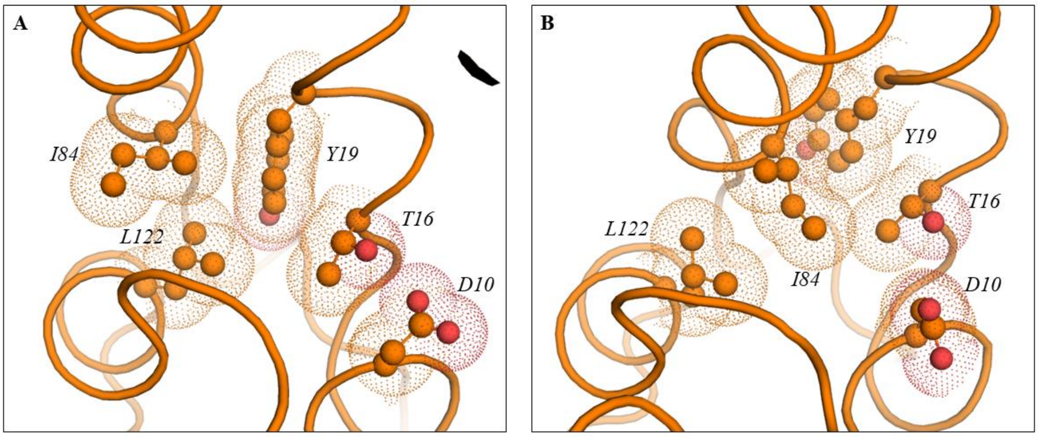


Figure 2-7 – c_1 Hinge conformation. The packing arrangement of hinge residue sidechains in (A) conformer c_2 and (B) conformer c_1. The backbone is displayed as an orange ribbon and sidechains are shown as ball-and-sticks with dots representing their Van Der Waals' radii. During the transition from c_2 to c_1, I84 displaces Y19 and packs against T16.

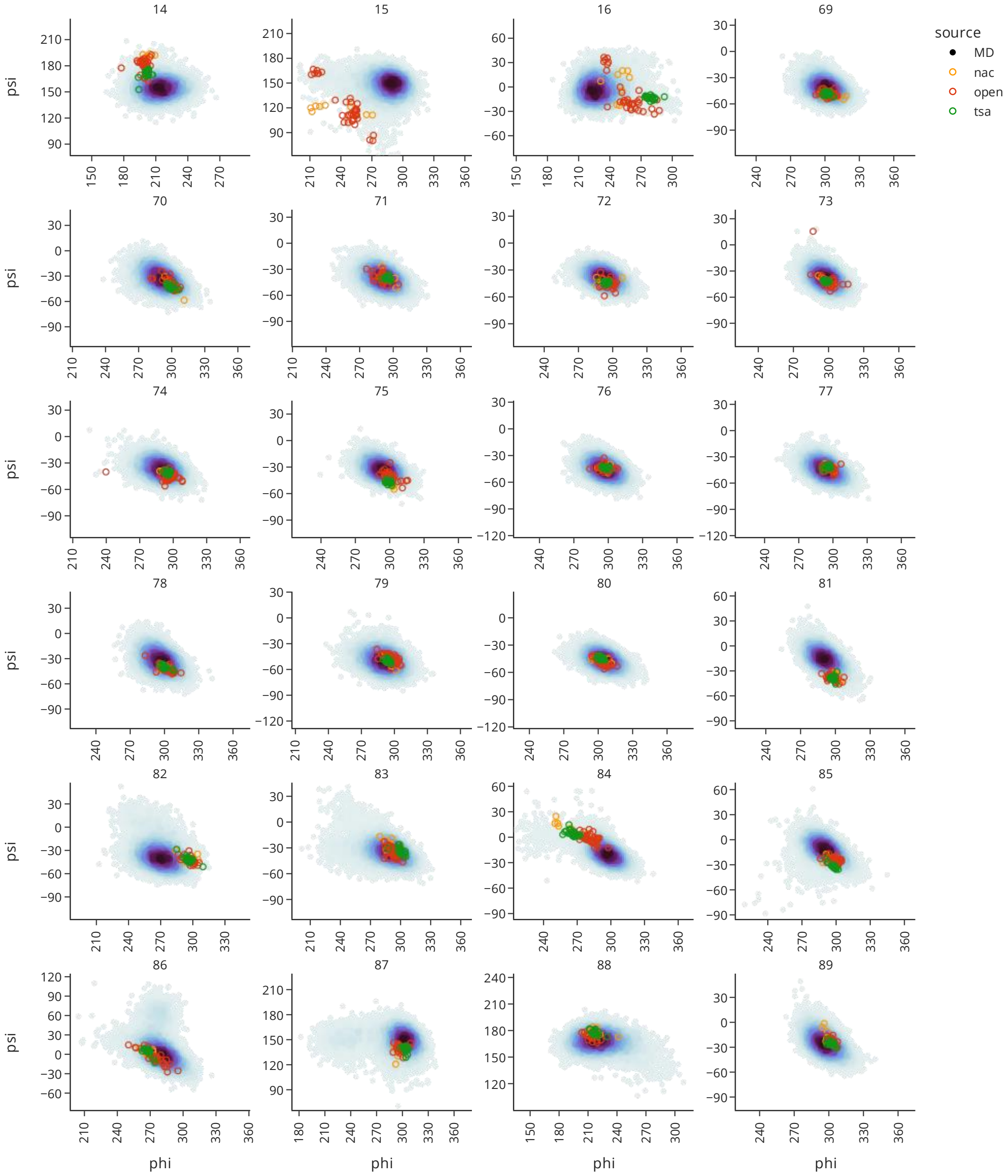


Figure 2-8 – c₁ Ramachandran distributions. Ramachandran (ϕ, ψ) distributions for hinge residues (T14, D15, T16, V81, K82, M83, I84, Q85, D86, V87, S88 and P89) and for residues at the base of α4 from F69 – Y80. Simulation frames were separated into 100×100 discrete bins and coloured according to the frame count for each bin, with lower counts coloured light blue and higher counts coloured dark purple. The angles calculated for the open, partially closed and fully closed TSA-like crystal structures are represented with open-circles and are coloured red, orange and green, respectively. Angles are expressed in an atypical periodic boundary.

2.2.5 Conformer c_3

In the crystal structures and in conformers c_1 and c_2, α4 runs from S65 through to I84, with the final α-helical hydrogen bond (O_i – HN_{i+4}) found between Y80 and I84. This is followed by a short 3_{10} helical turn (O_i – HN_{i+3}). V81-O hydrogen bonds with Q85-$N_{\epsilon 2}$; K82-O forms a long hydrogen bond with Q85-N; M83-O hydrogen bonds with D86-N; and I84-O hydrogen bonds with V87-N. The same arrangement is found in c_1 and c_2. Conformer c_3 is characterised by a rotation about the ψ-angle of I84, which converts the 3_{10}-helical turn to an α-helical turn. The backbone carbonyls of V81, K82 and M83 form hydrogen bonds with the backbone amides of Q85, D86 and V87, respectively (**Figure 2-9**). This motion causes an increase in pitch angle, which brings D86 and R22 into proximity allowing a salt bridge to form and requires burial of Y19 sidechain into the hinge cavity. The sidechain of I84 is lifted out of the pocket formed by N118, F121 and L122; the side chains of Y80 and H20 are pulled away from T16 and L122; and the hydrogen bond between H20-N and T16-O is stretched. Thus, the adoption of this conformation depends upon the balance between a collection of antagonistic interactions within and around the hinge.

In the ff15ipq-spceb-ecc-2whe simulation, the balance of forces lies in favour of the c_2 domain orientation, with only transient sampling of c_3. However, after D10 rotates away to interact with the magnesium ion, sampling of c_3 increases. Furthermore, the magnesium bound D10 rotamer is populated for the latter half of the ff15ipq-spceb simulation, during which c_3 dominates. The correlation between D10 sidechain orientation and c_3 sampling is also observed in the ff15ipq-spceb-5X simulations. This correlation may, at least in part, be a consequence of the disrupted packing interactions which result from reorientation of D10’s sidechain. This effect is illustrated in **Figure 2-10**. A contiguous packing arrangement is found between the sidechains of D10, T16, H20 and Y80, in conformer c_2. However, when D10’s side chain is reoriented, this packing arrangement is compromised which can be resolved either by adoption of an extended hinge conformation (with a c_2 domain orientation) which separates T16 from D10 or more often by adoption of c_3 which separates H20 and Y80 from T16 and D10.

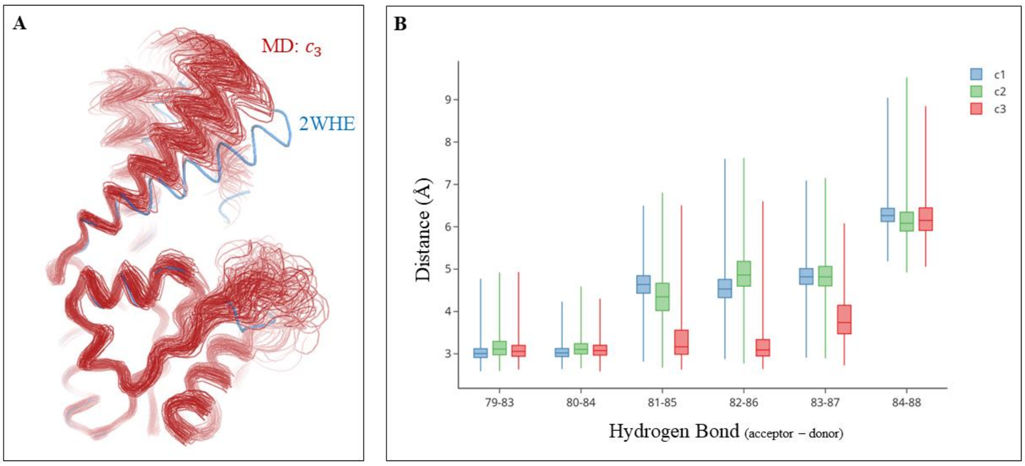


Figure 2-9 – c_3 domain orientation and hydrogen bonding. (A) Comparison of the cap domain orientation of the MD conformation c_3 (red) and the crystal structure 2WHE (blue). (B) Comparison of the hydrogen bond lengths of the N-terminal portion of the 70s helix. Residues V81-V87 adopt a 3_{10} helix in the MD conformers c_1 and c_2, resulting in long $O_i – N_{i+4}$ hydrogen bonds. Residues adopt an α-helix in conformer c_3, resulting in shorter $O_i – N_{i+4}$ hydrogen bonds.

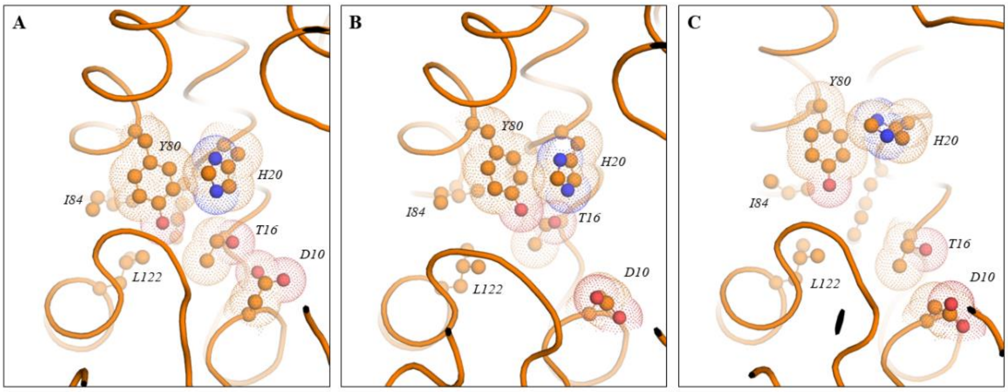


Figure 2-10 – c_3 hinge conformations. The packing arrangement of hinge residue sidechains in (A) conformer c_2 with D10 *out* and a compact hinge conformation (B) conformer c_2 with Mg-coordinated D10 and an extended hinge conformation (C) conformer c_3 with Mg-coordinated D10 and a compact hinge conformation. The backbone is displayed as an orange ribbon and sidechains are shown as ball-and-sticks with dots representing their Van Der Waals' radii.

2.2.6 Conformer c_4

The c_4 conformation is closely related to c_3 except that the salt bridge between D86 and R22 is replaced by a salt bridge between R22 and D91, which requires a further increase in the pitch angle. The formation of this salt bridge is strictly correlated with the transition to c_4. The anti-correlated trajectories of pitch angle and the R22–D91 side chain distances are shown in **Figure 2-11**. This prompted the use of the FF15IPQ forcefield which was developed to better improve the representation of salt bridges for proteins [129]. Based on our observations of R22 and D91 and the numerous long range electrostatic interactions between the two domains, it was expected that this would reduce the occupancy of c_4 and possibly shift orientation of the dominant conformations in solution. Indeed, the distributions in **Figure 7-1** show that conformations c_4 is not occupied in these simulations.

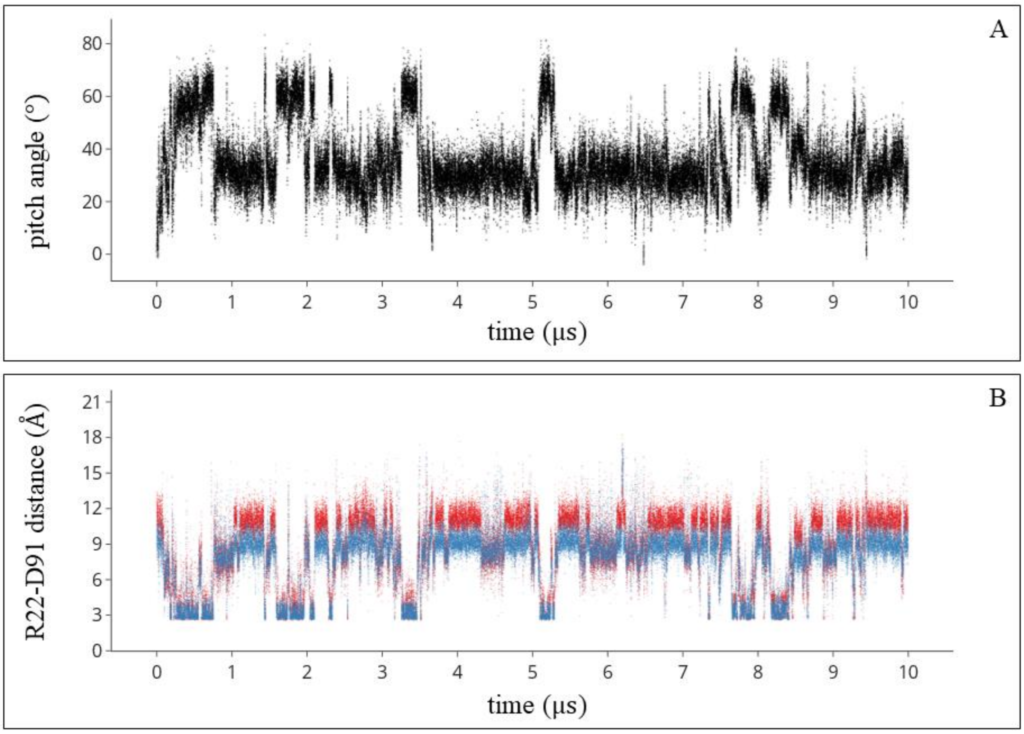


Figure 2-11 – A salt bridge biases the c_4 conformer. (A) ff14sb-tip3p-1 cap domain pitch angle trajectory reporting on domain orientation. (B) ff14sb-tip3p-1 trajectory of the distance between R22-NH1 and D91-OD$_1$ (red) R22-NH$_2$ and D91-OD$_2$ (blue) which reports on formation of a salt bridge.

2.2.7 P146 isomerisation effects

Substrate-free βPGM has two conformations (A and B) which exchange on a multi-second timescale, resulting from *cis*-*trans* isomerisation of P146 backbone amide [69] (Section 1.2.8). However, the energy barrier for P146 isomerisation is much too large to be sampled by conventional MD. All simulations were initiated from crystal structures with *cis*-P146, which was stable throughout. In crystal structures, the *cis* isomer is concomitant with recruitment of K145 into the active site, where it interacts with E169, which indirectly coordinates the catalytic magnesium via a water molecule. In our simulations, an alternative competing arrangement is sampled, where E169 adopts a direct mono- or bi-dentate coordination of the catalytic magnesium ion and K145 dissociates from the active site. This competition is shifted slightly in favour of the canonical crystallographic conformation when simulations are calculated under the FF15IPQ forcefield. However, as is standard in fixed-charge atomistic forcefields, the catalytic magnesium ion carries a 2+ charge which undoubtedly biases the conformations of the residues in its vicinity. To attenuate this effect, a simulation was run under the FF15IPQ forcefield with a further electronic continuum correction. Consequently, the canonical crystallographic interactions between K145-E169 and the catalytic magnesium were retained for the full length of the simulation (**Figure 2-12**).

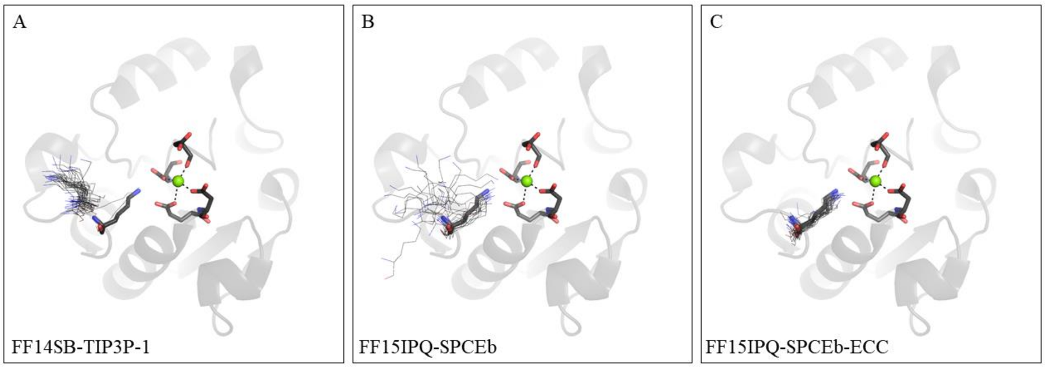


Figure 2-12 – Forcefield electrostatics bias K145 recruitment. The cartoon backbone of the core domain of the open crystal structure 2WHE with residues D8, D10, K145, E169 and D170 shown as sticks. Each of three trajectories (A) ff14sb-tip3p-1 (B) ff15ipq-spceb (C) ff15ipq-spceb-ecc, were aligned to the core domain backbone atoms and the K145 side chain positions are shown as lines.

To emulate the solution conditions more closely, a metadynamics bias was employed to rapidly sample P146 cis/trans isomerisation during a 5 μs simulation. An external, history-dependent periodic gaussian potential was applied to the K145 ω-angle (indirectly via the dihedral angle ζ defined by atoms: K145-C_α, K145-O, P146-C_δ and P146-C_α) and the P146 ψ-angle. Introduction of this bias resulted in rapid *cis/trans* isomerisation of the K145-P146 peptide bond (**Figure 2-13**). Furthermore, there was a substantial increase in the conformational dynamics in the vicinity of P146 (residues 138-148), indicating that the *trans* isomer destabilises the conformation of the surrounding loop. The convergence of these dynamics was evaluated using a dihedral categorisation strategy (see materials and methods). After 5 μs the P146 loop was still sampling new conformations meaning that the resulting ensemble was not fully converged and was not projected to do within a computationally accessible timeframe (**Figure 7-3**). Nevertheless, the canonical *cis* conformation was re-visited a few times throughout the simulation, albeit transiently due to trapping of E169 by an interaction with the magnesium ion. Despite unequilibrated dynamics in this loop, these simulations still serve as a model to evaluate the effects of P146 isomerism on domain dynamics.

Application of the metadynamics bias under the ff14sb forcefield resulted in a significant reduction in domain orientation. The disordered P146 loop can make contacts with the cap domain and stabilise a more closed conformation (**Figure 2-14)**. However, this effect is absent under the ff15ipq forcefield. In this simulation c_1 and c_3 conformations are populated, which is consistent with the distributions acquired in the absence of the P146 metadynamics bias. We conclude that the exaggerated electrostatics of the FF14SB forcefield biases the domain orientation to a more closed state, in simulations with a P146 ζ-angle metadynamics bias.

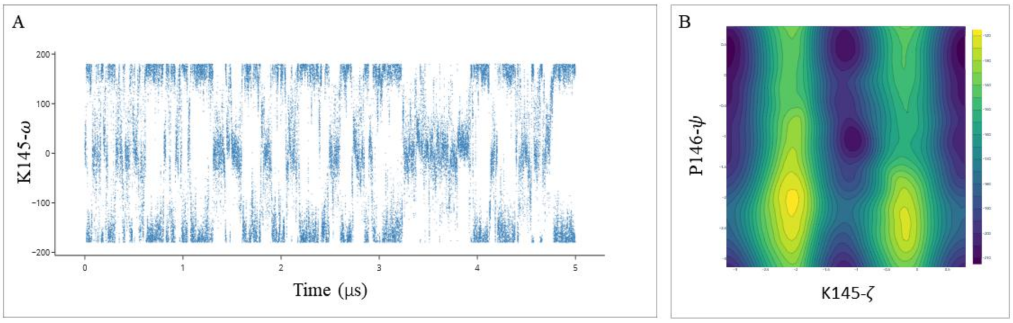


Figure 2-13 – Metadynamics sampling of K145/P146 cis/trans isomerism. (A) Trajectory of the K145-P146 ω angle during metadynamics simulation. (B) Contour plot of the final cumulative metadynamics bias in ζ / ψ space as an approximation of the free energy landscape for proline isomerisation. High energy regions (ζ = ± 90 °) are coloured yellow and low energy regions (ζ = ± 180 °) are coloured blue.

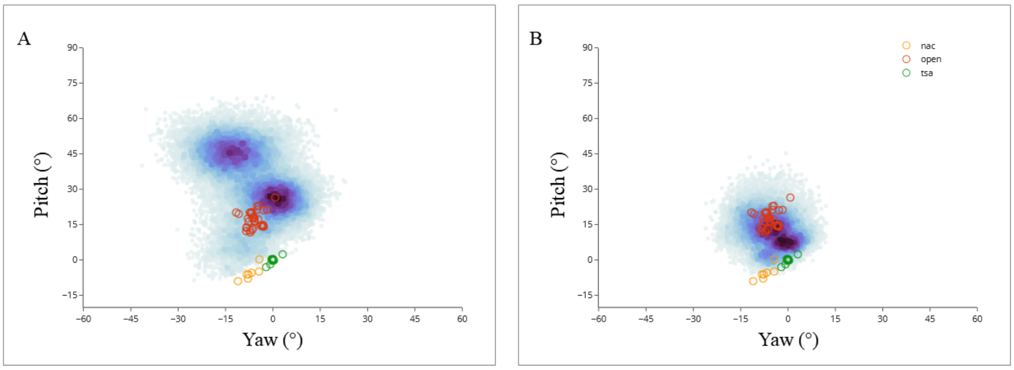


Figure 2-14 – Forcefield electrostatics bias domain orientation. The cap domain Euler angles for each of two simulations (A) ff15ipq-spceb-meta (B) ff14sb-tip3p-meta with a metadynamics bias acting on the K145-P146 ζ-angle were separated into 100 × 100 discrete bins and coloured according to the frame count for each bin, with lower counts coloured light blue and higher counts coloured dark purple. The angles calculated for the open, partially closed and fully closed TSA-like crystal structures are represented with open-circles and are coloured red, orange and green, respectively.

2.2.8 A partially closed, substrate-free conformer

Prior to running metadynamics simulations on the P146 ζ-angle, a conventional MD simulation using the ff14sb/tip3p forcefield was initiated using modified coordinates where P146 was switched to the *trans* isomer. As with the metadynamics simulations, this introduced substantial conformational dynamics in the vicinity of P146 and resulted in transient interactions with the cap domain. A further consequence of this, which was not observed in the metadynamics simulations, was trapping of a partially-closed, substrate-free conformation (**Figure 2-15**). This conformation is characterised by a drastic change in the roll-angle (-38 °). Once formed, this domain orientation did not continue to be stabilised via interactions with the P146 loop, which found a stable helical conformation. Instead, this conformation is stabilised by electrostatic interactions between R190, E41 and K45 as well as an unusual hydrogen bonding arrangement within the hinge between the backbone carbonyls of L9 and D10 and the sidechains of Y80 and H20, respectively.

The canonical pathway for substrate-induced closure has been previously characterised by crystallographic studies of metal fluoride ground state and transition state analogue complexes as well as D10N-trapped βG16BP complexes (Section 1.2). In this pathway, recruitment of the phosphodianion to the distal site; coordination of the sugar; and recruitment of D10 to coordinate the donor/acceptor axial ligand, results in a controlled change in domain orientation toward a catalytically competent TS-like conformation. The closure event, sampled in this simulation, is also associated with the displacement of D10 (which is coordinated by the catalytic magnesium) but the absence of sugar-phosphate results in an off-pathway conformation.

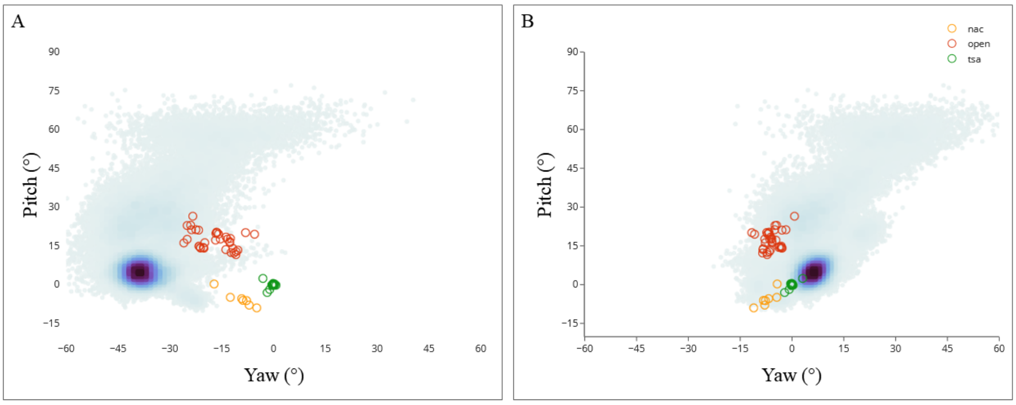


Figure 2-15 – Sampling of a substrate-free closed conformation. The cap domain Euler angles of the ff14sb-tip3p-trans simulation (A) pitch *vs.* yaw (B) pitch *vs.* roll. Euler angles were separated into 100 × 100 discrete bins and coloured according to the frame count for each bin, with lower counts coloured light blue and higher counts coloured dark purple. The angles calculated for the open, partially closed and fully closed TSA-like crystal structures are represented with open-circles and are coloured red, orange and green, respectively.

2.3 Results: Experimental Validation

2.3.1 Introduction

In the previous chapter, several distinct conformations obtained by MD simulation were described. Whilst MD simulations attempt to replicate the solution conditions, thereby alleviating any biases imposed by crystallographic packing forces, they also suffer from forcefield and sampling biases. It is therefore important to provide validation using experimental data acquired in solution. In the following section, Small Angle X-ray Scattering (SAXS), Analytical Ultracentrifugation (AUC) and Nuclear Spin Relaxation (NSR) measurements are described along with the calculations necessary to accurately relate them to the structural ensembles.

2.3.2 Small Angle X-ray Scattering

The following samples were prepared for SAXS data collection. βPGM_{WT} was exchanged into TSA buffer (50 mM HEPES pH 7.2, 10 mM $MgCl_2$, 20 mM NaF, 10 mM G6P) to generate a βPGM_{WT}:MgF_3:G6P TSA complex. 0.82 mM βPGM_{D10N} was incubated with ~16 mM AcP and ~8.2 mM G6P, resulting in the formation of a trapped βPGM_{D10N}:βG16BP:Mg complex, which was exchanged into 50 mM K+ HEPES pH 7.2, 5 mM $MgCl_2$, 2 mM NaN_3. A Magnesium-depleted βPGM_{D10N}:βG16BP complex was prepared by buffer exchange into 50 mM K+ HEPES pH 7.2, 2 mM NaN_3. Together these samples represent various closed conformations.

Substrate free samples were prepared by exchanging βPGM_{WT} into each of two buffers differing in their concentration of chloride: (50 mM HEPES pH 7.2, 5 mM $MgCl_2$, 2 mM NaN_3) or (50 mM HEPES pH 7.2, 100 mM $MgCl_2$, 2 mM NaN_3). Low chloride results in a mixture of A and B conformers, whilst conformer A is exclusively populated at high chloride concentrations. A βPGM_{P146A} substrate free sample was prepared by buffer exchange into 50 mM HEPES pH 7.2, 5 mM $MgCl_2$, 2 mM NaN_3. βPGM_{P146A} exclusively populates conformer B. Together these samples should elucidate any relationships between P146 isomerism and domain orientation.

SAXS data was collected using the B21 Beamline (Diamond Light Source, UK) with HPLC separation to ensure monodispersity. Radially averaged, reference-subtracted, normalised scattering data is plotted in **Figure 2-16**. The intensity at $0.8 \leq q \leq 0.14$ is particularly sensitive to the change in R_g between open and closed complexes.

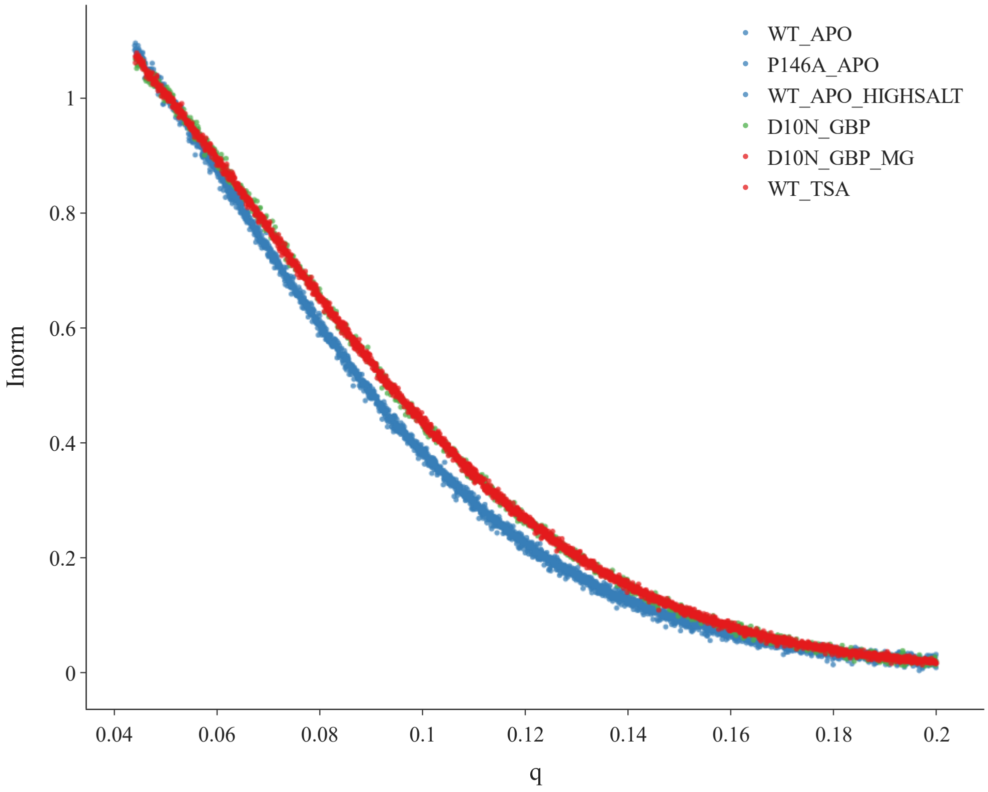


Figure 2-16 – SAXS curves for open and closed βPGM complexes. A plot of normalised, radially averaged scattering intensity against the scattering vector *q* for 1) βPGM_{WT} substrate free complex measured with 5 mM $MgCl_2$ 2) βPGM_{P146A} substrate free complex with 5 mM $MgCl_2$ 3) βPGM_{WT} substrate free complex with 100 mM $MgCl_2$ 4) βPGM_{D10N}:βG16BP complex 5) βPGM_{D10N}:βG16BP:Mg complex and 6) βPGM_{WT}:MgF_3:G6P TSA complex. Substrate-free open complexes are coloured blue. The closed βPGM_{D10N}:βG16BP NAC complex is coloured green. The closed βPGM_{D10N}:βG16BP:Mg and βPGM_{WT}:MgF_3:G6P TSA complexes are coloured red.

The relationship between structure and solution SAXS data is well understood [130]. The density of solvent at the surface of the protein is normally slightly higher than that of the bulk solvent. Consequently, the apparent radius of gyration of the molecule is slightly higher than the true value. This effect can be accounted for implicitly or explicitly. The parameters of an implicit model are often optimized by least-squares non-linear regression within set boundaries (ρ_{bulk} = 334 e $Å^{-1}$, $\delta\rho < 10$ %, $r_0 \pm 5$ %). This approach effectively reduces the capacity of the data to distinguish ensembles with subtly different R_g values, i.e. two structures with different R_g-values can fit equally well to the same experimental data by adjusting these parameters. Consequently, the boundaries applied to the fitted parameters directly correspond with the uncertainty in R_g. The open crystal structures of βPGM have ~6 % larger R_g than the TSA complexes. The MD simulations have 8 – 13 % larger R_g than the TSA complexes. To determine which ensemble fits the experimental data more accurately, $\delta\rho$ and r_0 must be fixed. Fitting these parameters within the specified bounds would result in both ensembles fitting equally well to the experimental data. Alternatively, an explicit solvent MD simulation can be used to calculate the contrast of the solvation layer. This is a parameter-free method and has shown to be largely independent of the MD forcefield and is implemented by the WAXSIS sever [131]. Unfortunately, this method is computationally expensive and can preclude analysis of several SAXS datasets for thousands of simulated structures. To overcome this issue, a uniform sample of ~50 structures across a wide range of R_g values were submitted to the WAXSIS server. The χ^2– profile as a function of R_g for each structure has a well-defined minimum. This same procedure was repeated using an implicit solvation model, with different fixed values for r_0 and $\delta\rho$. Each iteration gives a different χ^2 profile, where lower contrasts skew the χ^2 profile to higher R_g values and higher contrast skew the χ^2 profile to lower R_g values (**Figure 2-17**). The parameters which best match the profile generated using the explicit solvent model were rolled out to multiple datasets for thousands of simulated structures. The normalised χ^2 profiles calculated for each data set using the crystal structures and 1000 frames from the ff15ipq-spceb-ecc-2whe simulation are shown in **Figure 2-18**. The closed crystal structures are in good agreement with the SAXS data derived from closed complexes. The minimum χ^2 is found at 19.2 ± 0.2 Å for the datasets derived from substrate-free complexes. The distribution of R_g-values for each of the ff15ipq-spceb MD simulations are shown in **Figure 2-19**, **Table 2-3**. Conformer c_2 provides the best agreement with SAXS data, whereas the crystal structures are too compact.

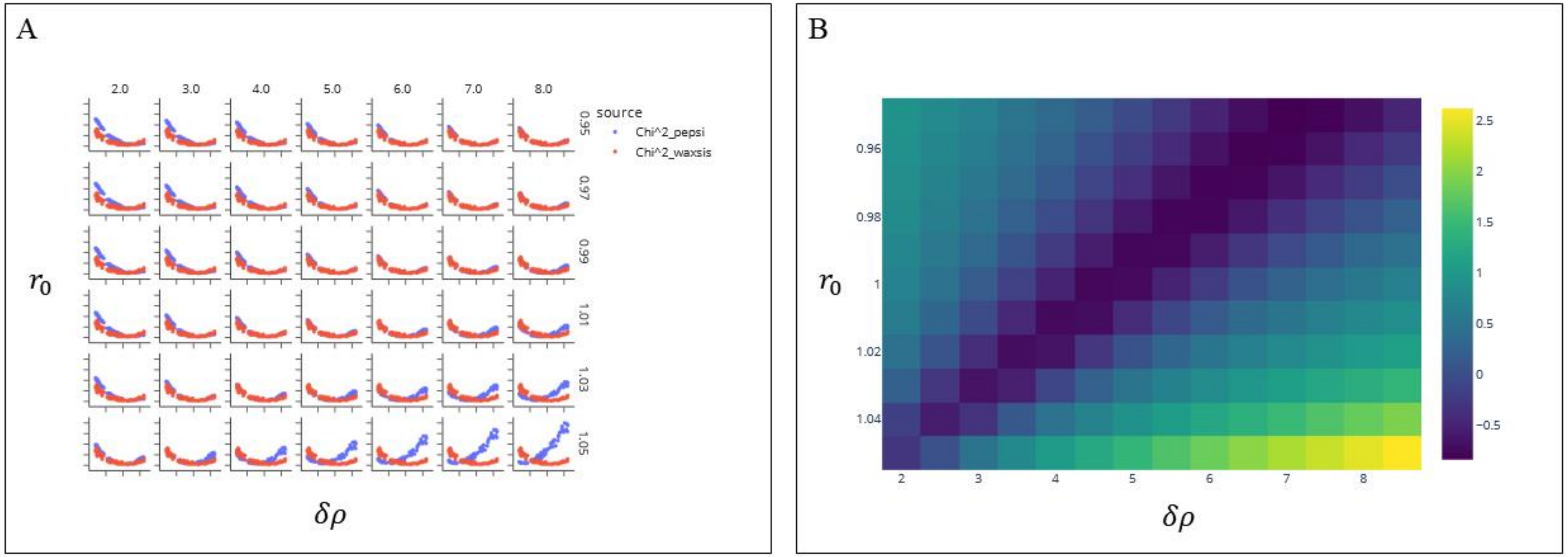


Figure 2-17 – Model optimisation for SAXS data analysis. (A) Each subplot shows the relationship between the χ^2 (y-axis) – derived from either an implicit solvation model (blue) or an explicit solvation model (red) – and the radius of gyration (x-axis). In each subplot, the parameters of the implicit solvent model ($\delta\rho$, r_0) are varied; the results of the parameter-free explicit solvent model are identical. (B) The log-difference in the χ^2 *vs.* R_g profiles between implicit and explicit solvent models, for different values of $\delta\rho$ and r_0 is plotted as an RMSD heatmap, with similar profiles coloured purple and dissimilar profiles coloured green/yellow.

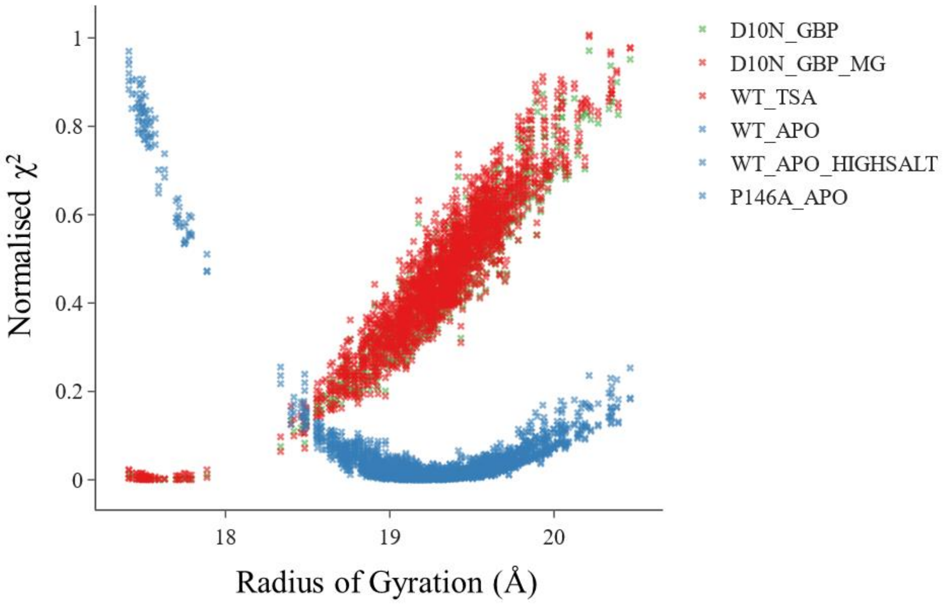


Figure 2-18 - χ^2 vs. R_g profiles from an optimised implicit solvation model. χ^2-values were derived from an implicit solvent analysis of each SAXS dataset, using the x-ray crystal structures and each of 1000 frames from the ff15ipq-spceb-ecc-2whe MD simulation. The implicit solvent parameters ($\delta\rho$ = 5 % and r_0 = 0.99) were optimised against an explicit solvent model. χ^2-values are plotted against the radius of gyration of each structure. SAXS data from substrate-free open complexes are coloured blue. Data from closed βPGM$_{D10N}$: βG16BP NAC complex is coloured green. Data from closed βPGM$_{D10N}$: βG16BP:Mg and βPGM$_{WT}$:MgF$_3$:G6P TSA complexes are coloured red.

Source	conformation	R_g
Crystal Structures	TSA	17.55 ± 0.07
	NAC	17.81 ± 0.06
	Open	18.73 ± 0.13
MD Simulation	c1	19.02 ± 0.17
	c2	19.30 ± 0.19
	c3	19.92 ± 0.15

Table 2-3 – βPGM R_g values for crystal structures and MD conformers. R_g values for different βPGM conformations derived from X-ray crystallography and MD simulations. The mean value and standard deviation is reported for each distribution.

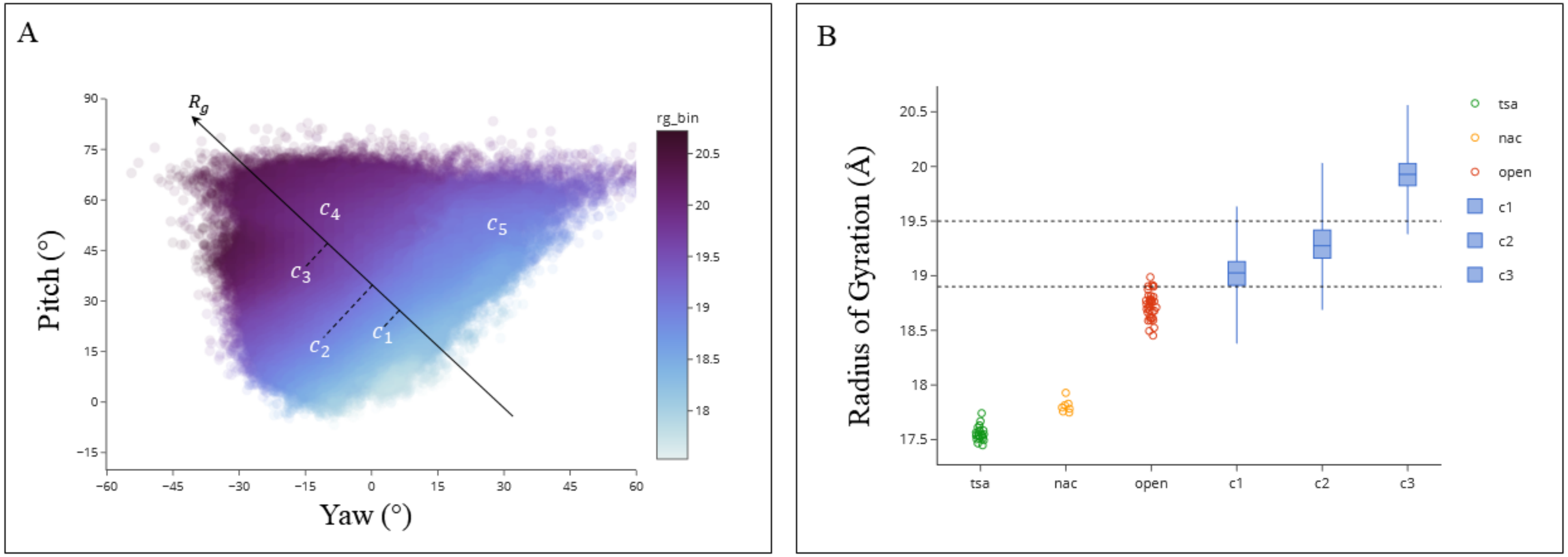


Figure 2-19 - R_g Distributions. (A) Euler Angle distributions (pitch *vs.* yaw) coloured according to the local average R_g value. (B) Radius of gyration distributions for different structures derived from x-ray crystallography and MD simulation. Fully closed, TSA-like structures are coloured green; partially closed NACs are coloured orange; substrate-free open complex are coloured red; MD simulation conformers are coloured blue and depicted using boxplots.

2.3.3 Analytical ultracentrifugation

Sedimentation and Hydrodynamic Modelling

Sedimentation can be described as a balance between gravitational/centrifugal (F_c), buoyant (F_b), and frictional (F_f) forces. During an AUC sedimentation velocity experiment, a constant (terminal) velocity is attained, with zero net force.

$$F_c - \left(F_b + F_f\right) = 0$$

$$F_c = m_p\omega^2 r \quad F_b = m_s\omega^2 r = m_p\bar{v}\rho \cdot \omega^2 r \quad F_f = f_t v$$

The buoyant force is the force required to move the displaced solvent against the centrifugal field. The term ($m_p\bar{v}\rho$) is equivalent to the mass of displaced solvent. The sedimentation coefficient is defined by the ratio of the sedimentation velocity to the angular acceleration within the centrifuge. The preceding equations can be rearranged to give an expression for the sedimentation coefficient in terms of the physical properties of the solution and solute – i.e. the Svedberg equation.

$$\mathrm{m}\omega^2\mathrm{r}(1-\rho\bar{\mathrm{v}}) - \mathrm{fv} = 0$$

$$\mathrm{s} = \frac{\mathrm{v}}{\omega^2\mathrm{r}} = \frac{\mathrm{m}(1-\rho\bar{\mathrm{v}})}{\mathrm{f_t}}$$

The frictional coefficient depends on the solvent viscosity (η_0) and a value which depends on the size and shape of the molecule which is normally expressed as the radius of a sphere with equivalent hydrodynamic behaviour – the Radius of Hydration (R_H) [132]

$$\mathrm{f_t} = 6\pi\mathrm{R_H}\eta_0$$

R_H is one of the primary outputs of the hydrodynamic bead models implemented in the HYDRO software suite. HYDROPRO10 was used to calculate R_H, using a residue-level primary hydrodynamic model comprised of beads centred on each C_α atom with a radius of 6.1 Å . This was preferable to the use of atomic resolution models, which were more computationally expensive, particularly when applying to thousands of simulated structures. The bead radius used in the primary hydrodynamic model was optimised against a diverse dataset of molecular structures and hydrodynamic measurements, with an error of 4-6% [133]. Due to this uncertainty, the ΔR_H between TSA and substrate free βPGM_{WT} complexes are compared. Thus assuming the βPGM:MgF_3:G6P TSA crystal structures accurately represent their structures in solution. R_H distributions for each of the simulations are shown in **Table 2-4**.

Source	Conformation	R_H (Å)	ΔR_H (Å)
MD Simulation	c3	24.97 ± 0.17	1.24 ± 0.19
	c2	24.68 ± 0.14	0.95 ± 0.16
	c1	24.48 ± 0.13	0.75 ± 0.15
Crystal Structures	open	24.31 ± 0.08	0.58 ± 0.11
	nac	23.82 ± 0.04	0.09 ± 0.09
	tsa	23.73 ± 0.08	N/A

Table 2-4 – βPGM R_H values for crystal structures and MD conformers. R_H values, predicted using HYDROPRO10, for different βPGM conformations derived from X-ray crystallography and MD simulations. The mean value and standard deviation is reported for each distribution. The difference in predicted R_H between the TSA complex crystal structures and each other conformation (ΔR_H) is also calculated with root-sum-squared errors.

Measuring Sedimentation Coefficients

The following samples were prepared for AUC data collection and measurement of sedimentation coefficients. A $\beta PGM_{WT}MgF_3$:G6P TSA complex was prepared by buffer exchange into 50 mM HEPES pH 7.2, 5 mM $MgCl_2$, 5 mM G6P, 20 mM NaF, 2 mM NaN_3. Substrate-free βPGM_{WT} was buffer exchanged into 50 mM HEPES pH 7.2, 5 mM $MgCl_2$, 2 mM NaN_3. Data was collected at 1, 0.5 and 0.25 mg/mL.

The Svedberg equation is insufficient to describe the results of an AUC experiment. Approximate or numerical solutions of the Lamm equation are normally used, which account for diffusion against the concentration gradient which is established during sedimentation. Sedimentation coefficients were derived a whole-boundary model [134], implemented in the SVEDBERG analysis software package (version 7.1.2) by John Philo. The fitted sedimentation coefficients were linearly extrapolated to zero-concentration using Monte-Carlo error estimation with a 10% uncertainty in protein concentration (**Figure 2-20**).

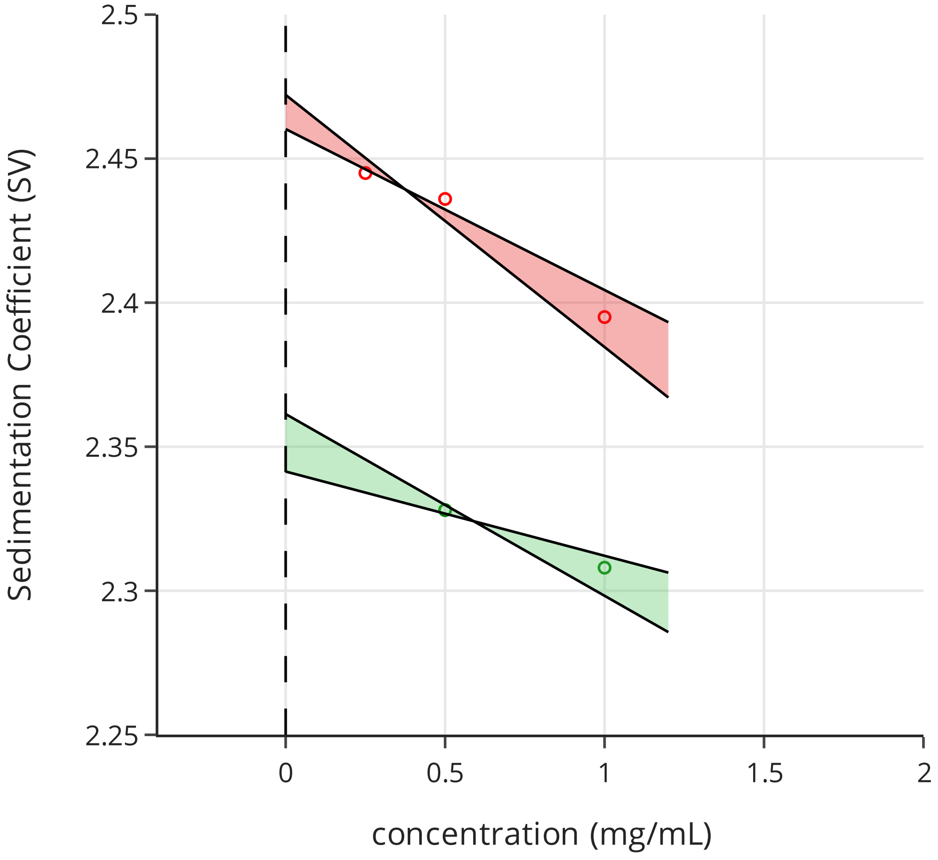


Figure 2-20 – Linearly extrapolated Sedimentation coefficients. The measured sedimentation coefficients for βPGM:Mg (green) and βPGM:MgF$_3$:G6P TSA complexes (red) are plotted against the concentrations at which they were measured. Least-squares linear regression with MC-sampling of protein concentration (assuming 10% error) resulted in a distribution of linear parameters. Two lines representing the 5th and 95th percentile of this distribution are shown for each complex.

The R_H was the calculated from the sedimentation coefficient (s) derived for each dataset using the Svedberg equation.

$$R_H = \frac{M(1 - \rho\bar{v})}{6\pi N_A s \eta_0}$$

The partial specific volume ($\bar{v}$) and molecular weight (M) were calculated from primary sequence of βPGM_{WT} using SEDFIT. High precision density and viscosity measurements were made for each buffer. All values and errors used in the calculation are given in **Table 2-5**. The calculated R_H values and numerically propagated errors at 2σ are: βPGM substrate free = 22.26 ± 0.24 Å; $\beta PGM:MgF_3:G6P$ TSA = 21.49 ± 0.22 Å and the difference: $\Delta R_H = 0.77 \pm 0.32$ Å. By comparison with HYDROPRO predictions the data is consistent with the findings of the SAXS experiments, however the uncertainty is too large to independently validate the simulated ensembles – i.e. they are also consistent with a solution ensemble which resembles the crystal structures (**Figure 2-21**).

	βPGM_{WT}:Mg	**$\beta PGM_{WT}:MgF_3$:G6P**
Sedimentation Coefficient (s) ($\times 10^{-13}$ s)	2.348 ± 0.0036	2.464 ± 0.0036
Partial Specific Volume ($\bar{v}$) (mL/g)	0.7442	0.7442
Buffer Density (ρ) (g/mL)	1.003126 ± 0.000005	1.002849 ± 0.000005
Buffer Viscosity (η) (cP)	1.0343 ± 0.0052	1.0362 ± 0.0052
Molecular Weight (kDa)	24.21 ± 0.01	24.64 ± 0.01

Table 2-5 – **Hydrodynamic variables for R_H calculations.** The values used in the calculation of R_H from sedimentation coefficients. Uncertainties at 1σ for sedimentation coefficients were derived from linear extrapolation with Monte Carlo sampling. Uncertainties in the measured buffer density and viscosity are derived from the DMA 5000M densometer & Lovis 2000ME viscometer manufacturer specifications.

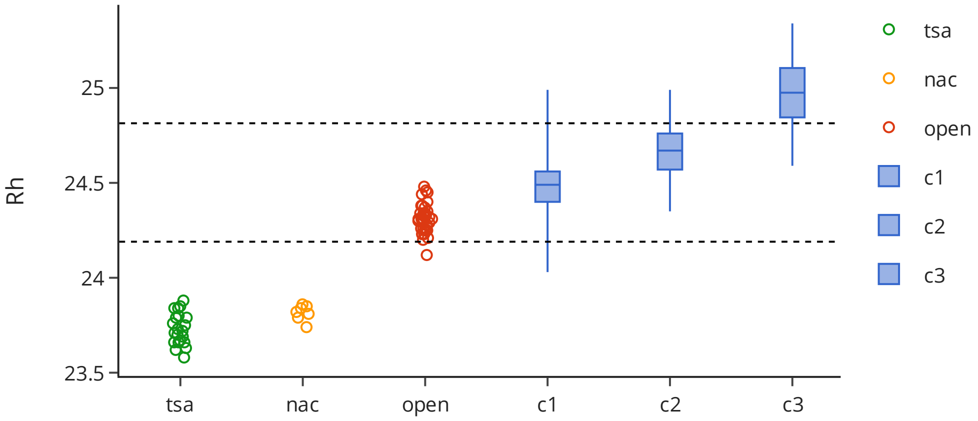


Figure 2-21 – Radius of Hydration distributions. Radius of hydration distributions for different structures, predicted using HYDRPRO, derived from x-ray crystallography and MD simulation. Fully closed, TSA-like structures are coloured green; partially closed NACs are coloured orange; substrate-free open crystal structures are coloured red; MD simulation conformers are coloured blue and depicted using boxplots. Black-dashed lines indicate the 95% confidence interval in the measurement of R_H for the βPGM substrate-free complex compared to the βPGM:MgF_3:G6P TSA complex.

2.3.4 NMR Relaxation

Rotational diffusion from Relaxation measurements

Longitudinal (R_1), transverse (R_2) and cross-relaxation rates (σ_{NOE}) are sensitive to specific frequency components of the local electromagnetic fields associated with random molecular motion. The relationship between the spectral density J(ω) and relaxation rates are stated in the following equations,

$$R_1 = d\big(J(\omega_H - \omega_X) + 3J(\omega_X) + 6J(\omega_H + \omega_X)\big) + cJ(\omega_x)$$

$$R_2 = \frac{d}{2}(4J(0) + J(\omega_H - \omega_X) + 3J(\omega_X) + 6J(\omega_H) + 6J(\omega_H + \omega_X) + \frac{c}{6}\big(4J(0) + 3J(\omega_X)\big) + R_{ex}$$

$$\sigma_{NOE} = d\big(6J(\omega_H + \omega_X) - J(\omega_H - \omega_X)\big)$$

where c and d are the chemical shift anisotropy and dipolar constants. It follows that the relaxation rates are dependent on five distinct frequencies (ω_H, ω_X, $\omega_H + \omega_X$, $\omega_H - \omega_X$, 0) which relate to the Larmor frequencies of the ^1H and X (e.g. ^{15}N, ^{13}C) nuclei. By collecting data at multiple field strengths, a larger number of frequencies are sampled which can be fit to a suitable 'model' describing molecular tumbling and conformation dynamics.

The 'model-free formalism' [135] and extensions thereof relate spectral density to dynamical parameters without the need for a specific dynamical model. The contributions from rotational diffusions are described by a single correlation time (τ_m). The amplitude and frequencies of local motions are represented by the generalised order parameter S^2 and effective correlation time τ_e, respectively. More complex models for local dynamics exist which separate the contributions from dynamics on multiple timescales (picosecond – nanosecond) and account for chemical exchange.

Anisotropic diffusion

If a macromolecule undergoes anisotropic rotational diffusion (i.e. the diffusion rate is dependent on the axis of rotation), the relaxation rate for a dipole X-H will depend on its orientation within the macromolecule. Most molecules do not have a perfectly symmetric, spherical mass distribution and thus have anisotropic rotational inertia. This is completely described using the moment of inertia tensor (3 × 3 matrix), whose *eigenvectors* specify three principal axes of inertia and *eigenvalues* state the inertia associated with rotation about each of the three corresponding principal axes. The inertia tensor is well-defined from the atomic coordinates of a protein. Anisotropic rotational diffusion depends upon the rotational inertia but must also account for hydrodynamic/frictional effects. These effects are more difficult to calculate but can be approximated using bead/shell models such as those implemented in HYDROPRO10[136].

Five distinct diffusion models may be assumed when modelling relaxation data.

1. The "local τ_m" model fits a single, independent correlation time for each residue.
2. A global, isotropic diffusion model fits a single, global correlation time shared by all residues in the molecule.

$$D_{xx} = D_{yy} = D_{zz}$$

3. A partially anisotropic oblate spheroid, in which rotation about one of the principal axes is slower than the others.

$$D_{xx} < D_{yy} = D_{zz}$$

4. A partially anisotropic prolate spheroid, in which rotation about one of the principal axes is faster than the others.

$$D_{xx} > D_{yy} = D_{zz}$$

5. A fully anisotropic *ellipsoid*, where rotational diffusion is different for all three principal axes.

$$D_{xx} \neq D_{yy} \neq D_{zz}$$

HYDROPRO10 was used to calculate the anisotropic rotational diffusion tensor (D_{rr}) and the global correlation time (τ_c) for each of the crystal structures and for each of the MD simulations. The eigenvalues of the diffusion tensor (D_{xx}, D_{yy} and D_{zz}) were calculated and are plotted alongside the global correlation times in **Figure 2-22**. The eigenvalue distributions for both substrate-free and closed, TSA samples the show a single fast diffusion component ($D_{xx} \approx 1.45 \times 10^7$) and two slow diffusion components ($D_{yy} \approx D_{zz} \approx 0.97\times10^7$). Therefore, a prolate diffusion model is most suitable for both the substrate-free complex and TSA complexes. D_{xx} is relatively insensitive to the transition from open to closed, whilst D_{yy} and D_{zz} increase significantly. The average affect is captured by the global correlation time (τ_m) which decreases. As described in Section 2.3.3, the change in hydrodynamic parameters (rather than their absolute values) have been calculated between the TSA structures and all other conformations and are stated in **Table 2-6**.

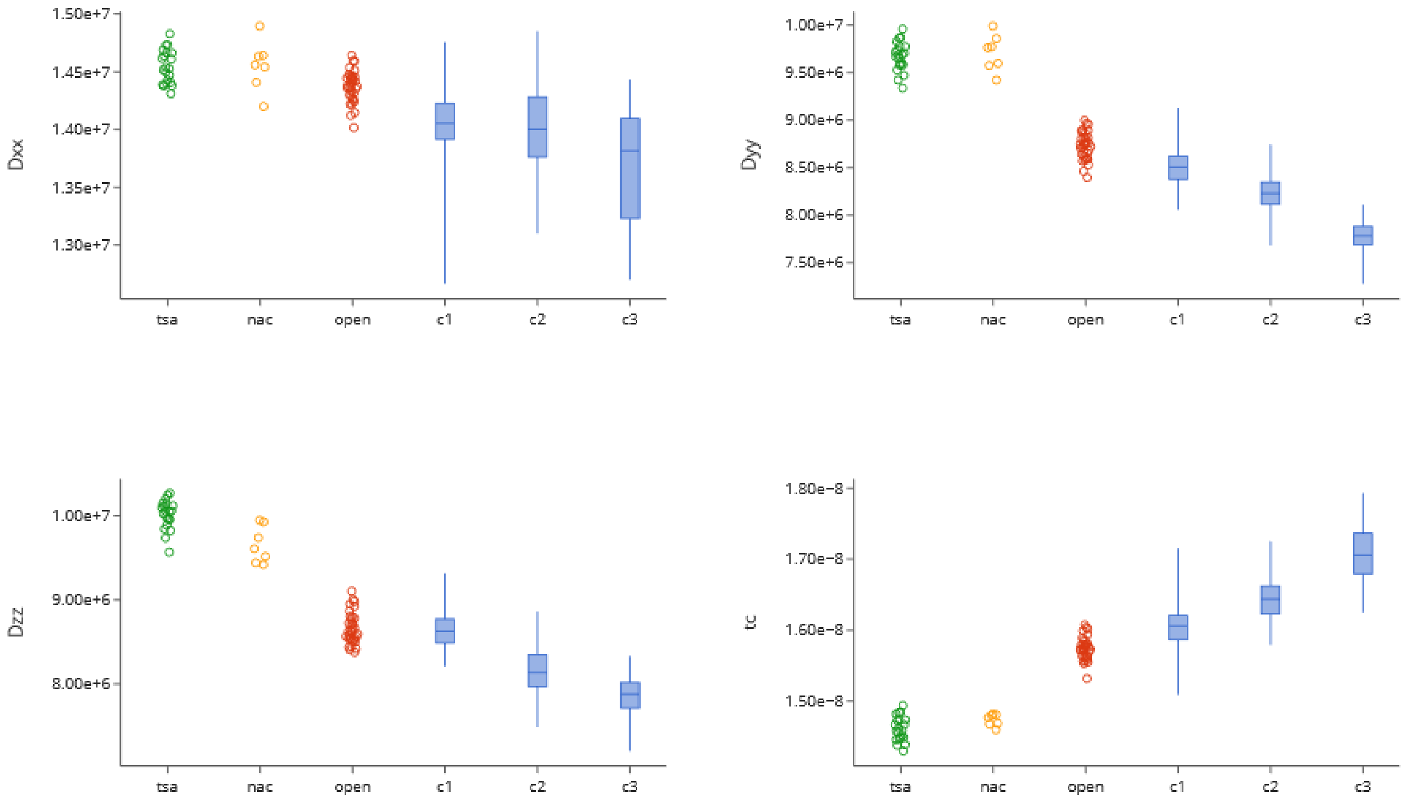


Figure 2-22 – Correlation time (τ_m) distributions. Rotational correlation time distributions calculated using HYDROPRO, for different structures derived from x-ray crystallography and MD simulations. Fully closed, TSA-like structures are coloured green; partially closed NACs are coloured orange; substrate-free open complex are coloured red; MD simulation conformers are coloured blue and depicted using boxplots.

Source	Conformation	τ_c (ns)	D_{xx} (rad/μs)	D_{yy} (rad/μs)	D_{zz} (rad/μs)
MD Simulation	c3	17.06 ± 0.39	13.69 ± 0.49	7.78 ± 0.16	7.85 ± 0.23
	c2	16.45 ± 0.31	14.02 ± 0.36	8.22 ± 0.17	8.16 ± 0.26
	c1	16.03 ± 0.27	14.06 ± 0.26	8.51 ± 0.19	8.64 ± 0.2
Crystal Structures	open	15.74 ± 0.16	14.37 ± 0.14	8.73 ± 0.14	8.66 ± 0.18
	nac	14.74 ± 0.08	14.55 ± 0.22	9.71 ± 0.19	9.66 ± 0.22
	tsa	14.61 ± 0.17	14.54 ± 0.14	9.67 ± 0.15	10.01 ± 0.17

.Source	Conformation	$\Delta\tau_c$ (ns)	ΔD_{xx} (rad/μs)	ΔD_{yy} (rad/μs)	ΔD_{zz} (rad/μs)
MD Simulation	c3	2.45 ± 0.42	-0.86 ± 0.51	-1.89 ± 0.22	-2.16 ± 0.29
	c2	1.84 ± 0.35	-0.53 ± 0.39	-1.44 ± 0.23	-1.85 ± 0.31
	c1	1.42 ± 0.32	-0.49 ± 0.29	-1.16 ± 0.24	-1.38 ± 0.27
Crystal Structures	open	1.13 ± 0.24	-0.17 ± 0.2	-0.94 ± 0.21	-1.35 ± 0.25
	nac	0.13 ± 0.19	0.01 ± 0.26	0.04 ± 0.25	-0.36 ± 0.28
	tsa	N/A	N/A	N/A	N/A

Table 2-6 – **Diffusion tensor eigenvalue distributions.** predicted using HYDROPRO10, for different βPGM conformations derived from X-ray crystallography and MD simulations. The mean and standard deviation are reported for each distribution. The difference in predicted values between the TSA complex crystal structures and each other conformation is also calculated with root-sum-squared deviation.

Measurement of R_1, R_2 and cross-relaxation rates

Substrate free βPGM$_{WT}$ sample was prepared containing 1 mM βPGM$_{WT}$ in 50 mM K^+ HEPES pH 7.2 + 5 mM $MgCl_2$ + 2 mM NaN_3. A βPGM$_{WT}$:MgF_3:G6P TSA complexes were prepared by addition of 10 mM NaF and 10 mM G6P. TROSY-based R_1, $R_{1\rho}$ and ^{15}N-{^{1}H} NOE experiments [137] were used to measure ps-ns dynamics for each complex.

Data fitting with a local τ_m diffusion model

Model Free analysis was carried out using RELAX [138]. A local τ_m model was initially used to calculate a single correlation time for each spin system. The local τ_m distributions for the βPGM$_{WT}$ substrate free and βPGM$_{WT}$:MgF_3:G6P complexes are shown in **Figure 2-23**. The combined lack of sensitivity and large uncertainties, mean that the data from this analysis cannot be compared with the theoretical correlation times and therefore cannot be used to validate a structural model of the substrate-free complex.

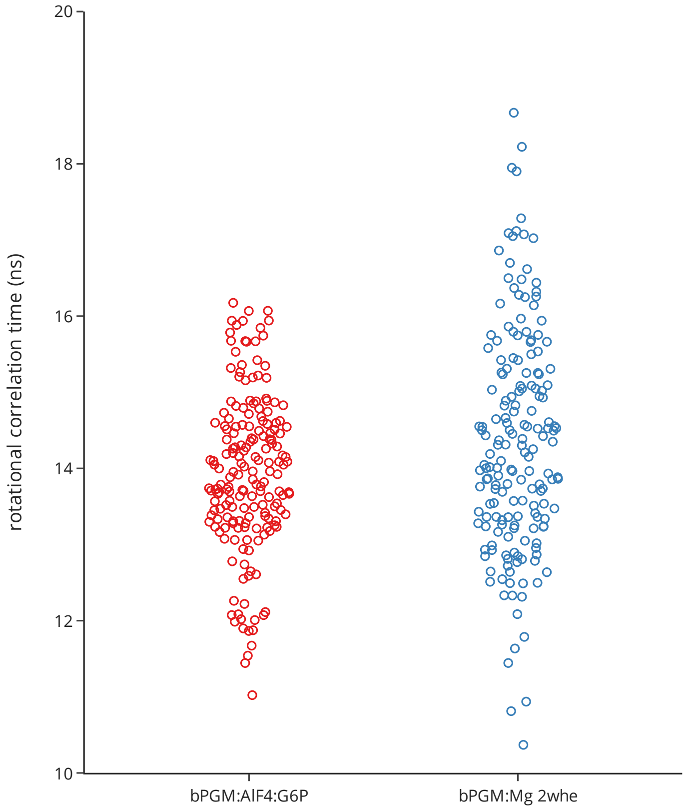


Figure 2-23 – local τ_m distributions. Strip chart showing the distribution of locally fit Brownian rotational diffusion correlation times (local τ_m) measured for a βPGM substrate free complex (blue) and for a βPGM:AlF_4:G6P TSA complex (red).

Data fitting with a prolate diffusion model

The large variation in local τ_m values may result from anisotropic rotational diffusion. Therefore, the relaxation data were fit using the dual-optimisation methodology, implemented by RELAX, for optimisation of model-free parameters and a prolate global diffusion tensor. Briefly, an initial estimate of the global prolate diffusion tensor is derived from the local τ_m values. The diffusion tensor is fixed during model-free model fitting, elimination, and selection. The selected model is then optimised alongside the global diffusion tensor. This process is repeated until parameter convergence is obtained. This protocol requires user-input atomic coordinates to define the relative orientations of the spins-systems. Coordinates from the $\beta PGM_{WT}:AlF_4:G6P$ (PDB: 2WF6) crystal structure was used for the analysis of the corresponding relaxation data. Analysis of the substrate free dataset was run using either an open βPGM_{WT} crystal structure (PDB:6YDL) or the average coordinates of the ff15ipq-spceb-ecc-2whe MD simulation. Note: calculations using 2WHE became stuck in a non-convergent optimisation loop. The results of the analysis are given in **Table 2-7.** Domain opening, as modelled by HYDROPRO, results in a small decrease in D_{xx} and a larger decrease in D_{yy} and D_{zz}, thus resulting in an overall increase in τ_m. Conversely, according to the diffusion tensors optimised using relaxation data, the D_{xx} component *increases* substantially and the $D_{yy} = D_{zz}$ components decrease slightly, resulting in a small increase in τ_m. The diffusion model derived from the relaxation data is inconsistent and incomparable with HYDROPRO calculations. Ultimately, the relaxation data is not sufficiently sensitive to the change in global correlation time τ_m to assess the validity of the MD ensemble.

Source	**Relaxation Data**	**τ_c (ns)**	**D_{xx} (rad/µs)**	**D_{yy} ΔD_{zz} (rad/µs)**
ff15ipq-spceb-ecc-2whe	APO	14.1	15.4	10.1
6YDL	APO	14.0	15.8	9.9
2WF6	TSA	13.8	14.6	10.8

Structure	**Relaxation Data**	**$\Delta\tau_c$ (ns)**	**ΔD_{xx} (rad/µs)**	**$\Delta D_{yy} = \Delta D_{zz}$ (rad/µs)**
ff15ipq-spceb-ecc-2whe average	APO	0.3	0.8	-0.7
6YDL	APO	0.2	1.2	-0.9
2WF6	TSA	N/A	N/A	N/A

Table 2-7 – Prolate diffusion models of relaxation data. The absolute values of the diffusion tensor components and global correlation times derived from the fitting of: TSA relaxation data to a prolate diffusion model using 2WF6; APO relaxation data to a prolate diffusion model using 6YDL; and APO relaxation data to a prolate diffusion model using the average coordinates from the ff15ipq-spceb-ecc-2whe MD simulation. The Δ values are calculated relative to those of the TSA dataset.

2.4 Discussion

2.4.1 Describing Domain Orientation in βPGM

Several metrics have been reported which describe changes in domain orientation. The simplest methods involve calculation of select interatomic distances/angles; which are sensitive to specific conformational changes. The backbone RMSD and Radius of Gyration (R_g) are more generally applicable metrics which are frequently reported. However, whilst these methods are relatively interpretable, they suffer from some information loss, i.e. distinct conformations may be assigned the same value of a given metric. This is particularly problematic for conformationally diverse ensembles, such as those derived from an MD trajectory. A more information rich approach describes domain reorientation as a rigid-body transformation, calculated as an alignment tensor, which can be represented in multiple ways. The DYNDOM server converts the alignment tensor to a screw-axis transformation, with five parameters representing the position and orientation of the screw-axis along with a translation along the screw axis (d) and a rotation around the screw-axis (θ). However, describing an ensemble using just one of these parameters (typically θ) results in information loss. An alternative method takes one of these underdetermined metrics and uses them to calculate a pairwise distance matrix which can then be submitted to a clustering algorithm. The backbone RMSD would typically be used as a distance metric for structural clustering of proteins, however a previous analysis of the *β*PGM crystal structures has used the DYNDOM screw-axis angle (θ) for calculation of a distance matrix, which was used for manual classification of structures [139]. Here we convert the domain alignment tensor to a set of three Euler angles. To generate angles with a consistent interpretation, the principal axes of the cap domain of 2WF5 are used as an initial coordinate system for calculations using an x-y’-z’’ convention (aka. nautical yaw, pitch & roll angles). The yaw angle captures the left-to-right shear motion of the cap domain. The pitch captures the clam-shell-like opening and closing mechanism. The roll captures the twisting motion around the long axis of the cap domain. Hierarchical clustering using the backbone RMSD as a distance metric identified three high-level clusters corresponding to the fully closed canonical TSA-like conformation; the partially closed conformations comprised of various NACs; and the substrate-free open complexes. Euler angles clearly discriminate between each of these structures. These angles proved to be a useful tool for visualising the potential energy landscape of domain orientations in βPGM and aided characterisation of distinct hinge conformations.

2.4.2 Experimental Validation

Many experimental approaches cannot unambiguously define a structural model. MD simulations are often used to complement experiments by adding forcefield information to narrow down the conformational ensemble. The value of the simulation, in this context, is therefore limited by the experimental uncertainty and by the uncertainty in the theoretical relationships between atomic coordinates and experimental observables. The relationship between SAXS data and molecular structure (i.e. Radius of Gyration) is very well understood. However, in the absence of an explicit model of the solvent effects, the uncertainty in this relationship can be prohibitively large. Conformational changes which affect the R_g by < 15% are difficult to validate using SAXS because the magnitude of the change in R_g is of the same order as the uncertainty in the density of the solvation layer. This solvation effect is usually accounted for implicitly by a parameter which adds a uniform excess of electron density ($\delta\rho$) around the surface of the protein and via a free parameter (r_0) which define the volume of displaced solvent [140]. The model implemented by WAXSIS resolves this issue using MD simulations to explicitly model the density of the solvation layer and the volume of displaced solvent. Thus, the only parameters which are fit are a scaling factor (I_0) and a baseline offset (c). This model has been shown to be forcefield-independent and thus serves as a robust tool for validating conformational ensembles using SAXS data [141]. The βPGM$_{WT}$:AlF$_4$:G6P and βPGM$_{WT}$:MgF$_3$:G6P TSA structures are stable, trapped complexes, with a compact global conformation which is less susceptible to crystallographic packing artefacts than the open, substrate-free structures. These structures show excellent agreement with the solution SAXS data, providing an independent validation of the explicit solvent analysis implemented by WAXSIS, and therefore greater confidence in the conclusions made about the conformation of the substrate-free enzyme.

Prediction of hydrodynamic parameters from atomic coordinates also has inherent uncertainty. HYDROPRO evaluates frictional coefficients of a primary hydrodynamic model (PHM), where the radii of the primary elements of the model were fit to minimise the deviation between the experimentally measured hydrodynamic radii and those predicted from a corresponding crystal structure [133]. The typical error associated with these predictions is ~4%, which is similar in magnitude to the conformational change between the open and closed βPGM complexes. Therefore, the change in hydrodynamic parameters were compared rather than the absolute values. However, the model implemented in HYDROPRO were fit to data derived from structures which varied between 6 – 466 kDa and was not necessarily designed to accurately quantify the changes associated with a subtle domain reorientation. In addition to the apparent lack of sensitivity of the relaxation experiments and the cumulative error in the sedimentation data, it appears that the use of hydrodynamic measurements is not suitable for validation of domain orientational dynamics in βPGM and other enzymes with dynamics of a similar amplitude.

2.4.3 D10 artefacts

The population of a magnesium coordination state, which has no experimental validation, points to a well-documented problem with classical MD simulations of metalloproteins. The standard model describes a monoatomic ion as a particle with full nominal charge and bespoke Lennard-Jones parameters. However, the binding and coordination characteristics of metalloproteins using such models is often inaccurate, which has resulted in the development multi-site models and more computationally expensive polarisable models [142–144]. Furthermore, water molecules cannot be ionised in conventional MD simulations which make them more likely to be displaced by nearby ions. The simulations described here illustrate the propagation of local coordination artefacts to global conformational changes. Three distinct domain orientations were identified which are biased by aberrant displacement of the general acid/base D10. **1)** A shear domain motion, represented by a change in the Yaw angle (c_1), occurs exclusively in simulations populating a direct interaction between D10 side chain and the catalytic magnesium. Furthermore, the experimentally observed *out* rotamer sterically precludes the repositioning of I84 sidechain. **2)** Conformer c_3 is significantly more open and is stabilised by the disruption of the packing interactions within the hinge, resulting from the displacement of D10. **3)** A substrate-free closed conformation was observed which is sterically precluded by the canonical D10 *out* rotamer. Whilst sampling of these states is, in general, undesirable because they deviate from the on-pathway ensemble and hinder sampling of the functionally relevant regions of the conformational landscape, they do point to an auxiliary roll for D10 in destabilising off-pathway open conformations.

2.4.4 Salt-bridge artefacts

The stability of solvent-exposed salt bridges in proteins has been studied using the interactions between guanidinium, butylammonium, imidazolium and acetate ions as a model for the interactions between arginine, lysine, histidine, glutamate and aspartate sidechains. The stability of different types of salt bridge were measured by potentiometric titration [145]. The populations inferred from these measurements were compared with those calculated from MD simulations using a range of contemporary forcefields from Amber, CHARMM and OPLS [146], which showed that all forcefields overestimate their populations significantly. The ff15ipq forcefield was later developed and which gave better agreement with experiments. For example, salt bridges found in crystal structures of protein GB1 domain were not detectable by NMR [147]. These salt bridges were found in MD simulations run using several conventional forcefields, are only weakly formed under the ff15ipq forcefield [148].

The formation of the R22-D91 salt bridge in βPGM is strictly coupled to a substantial increase in the cap domain pitch angle (conformer c_4). In the ff14sb-tip3p-1 simulation, D10 coordinates the catalytic magnesium and conformer c_1 dominates the ensemble. This conformation separates R22 from D91, weakening their electrostatic attraction. Nonetheless, conformer c_4 still represents ~20% of the ensemble. In the ff14sb-tip3p-2 simulation, D10 adopts the canonical *out* conformation and conformer c_1 is sterically precluded. Under these conditions, conformer c_4

dominates the ensemble (>50%) alongside a small population of c_2 and c_3 (**Figure 7-1**). Our SAXS data indicate that conformer c_4, and therefore the R22-D91 salt bridge, is scarcely populated in solution. Thus, the ff14sb/tip3p forcefield over-stabilises the R22-D91 salt bridge resulting in a biased domain orientation. This does not occur under the ff15ipq/spceb forcefield.

Simulations run with a metadynamics bias on the K145-P146 ζ-angle result in augmented dynamics in the P146-loop, and reorientation of E140 toward the cap domain. Simulations run under the ff14sb/tip3p forcefield adopt a more closed conformation stabilised by a salt bridge between E140 and R49. Conversely, the domain orientation in simulations run under the ff15ipq/spceb forcefield are unaffected. Our SAXS data indicates that there is not a significant change in domain orientation between conformer A and B. Thus, stabilisation of a more closed conformation by the E140-R49 salt bridge (with K145-P146 ζ-angle metadynamics-enhanced sampling) is another example of conformational biasing by the ff14sb/tip3p forcefield.

Together, these results provide a general warning to users of conventional MD forcefields, that protein dynamics (particularly domain orientation) can be biased substantially by overpopulated salt bridges.

2.4.5 A hydrophobic cavity facilitates product release.

Progression from the substrate-free open complex to the transition state involves multiple steps, with incremental changes in: the domain orientation; the backbone torsions of the hinge region; the hydrogen bonding network between D10, T16 and the substrate; and the volume of a hydrophobic cavity in the hinge. Our validated MD simulations show that the substrate-free open complex is more accurately described by an ensemble with a more open conformation which places the side chain of Y19 into a hydrophobic cavity within the hinge. The crystal structures have a slightly more compact conformation, biased by crystal packing forces, where Y19 and the hydrophobic cavity are exposed to solvent. An alternative hinge conformation was also identified in both the crystal structures and MD simulations which involves rotations about D15-ϕ and T16-ψ, which inserts T16 sidechain into the hydrophobic cavity. The crystal structure of the βPGM:BeF_3:G6P NAC-I complex (PDB: 2WF9 [17]) shows that substrate binding and domain closure necessarily lifts Y19 sidechain from this cavity, exposing both Y19 sidechain and the hydrophobic cavity to solvent. Two T16 sidechain rotamers are populated in this crystal structure owing to the free space within the hinge. D10 remains in the *out* position, with D10-$O\delta_1$ hydrogen bonded to T16-N. The βPGM_{D10N}:βG16BP NAC III complex (PDB: 5OK1, [18]) is arrested prior to proton transfer owing to the D10N mutation. This complex is more closed than the NAC I complex – torsional changes within the hinge region cause a decrease in pitch angle of the cap domain. N10 hydrogen bonds to βG16BP-O1 and displaces T16 sidechain which moves into the hydrophobic cavity – reducing its size. In the TSA complex, D10 makes a shorter hydrogen bond to G6P-O1 which makes space for a change in domain orientation which causes the residual space within the hydrophobic cavity to collapse – there is a closer packing between the sidechains of L9, T14 and T16 (**Figure 2-24**). In summary, the substrate free enzyme has a stable configuration which, sequesters the hydrophobic sidechain of Y19 into the hydrophobic cavity within the hinge. The NAC I complex is destabilised as the initial mode of domain closure lifts Y19 sidechain out of its hydrophobic pocket, exposing both the sidechain and the internal cavity to solvent (**Movie S1**). Progression to the transition state requires changes in the hydrogen bonding network, which sees that the cavity is filled by T16 and collapses to a more compact conformation. These observations imply a specific destabilisation of the closed ground state NAC I conformation, via a loosely packed, solvent-exposed, hydrophobic cavity. The WT Michaelis complex is closed [139] and is likely a NAC I conformation (Section 3.3.1). Therefore, destabilisation of this complex would result in a decrease in K_m which would be compensated for by an increase in k_{cat} – the overall catalytic efficiency (k_{cat}/K_m) would be largely unaffected. The putative role of this architecture would thus serve to provide thermodynamic stabilisation of the open complex, facilitating product release. It appears that this mechanism would not diminish the effect of substrate inhibition. The substrate-inhibited complex is modelled by the βPGM_{D170N}:βG1P complex (6HDG) [139] and can achieve near full domain closure with a collapsed hydrophobic cavity.

Y19 is moderately conserved amongst βPGM sequences (**Figure 7-9**, **Section 5.2.14**). Tyrosine has a 54.7 % occupancy at this position with the majority of other sequences having bulky hydrophobic and/or aromatic residues (Phenylalanine = 20.9 %, Leucine = 17.3 %, or Histidine = 1.7 %) which could fulfil the same functionality. The presence of a bulky hydrophobic at the base of the hinge adjacent to an internal cavity may be a partially redundant mechanism for ground state destabilisation resulting in weaker conservation than, for example, residues in the distal-phosphate, proximal-phosphate, sugar-coordination and magnesium-binding sites.

A related mechanism was suggested for 3’-phosphoglycerate kinase (PGK). Small angle x-ray scattering data was used to generate a solution model of the substrate-free open complex. The refined structure was significantly more open than the crystal structure. The domain orientation measured compared to the fully closed active conformation was 55°, as opposed to a maximum 33° rotation observed in the crystal structures. A patch of hydrophobic residues is buried in this new structure, which are otherwise exposed to solvent in the closed complexes. Thus, the authors proposed a “spring-loaded” release mechanism, whereby the open conformation is thermodynamically stabilised through burial of hydrophobic residues, facilitating product release and efficient substrate binding [81]. An implication of this mechanism is that the exposure of hydrophobic residues to solvent also affects the stability of the closed transition state. Conversely, the equivalent mechanism in βPGM is able to temper the strong stabilisation of the closed product complex by the distal-site phosphodianion interaction, without compromising the stability of the transition state.

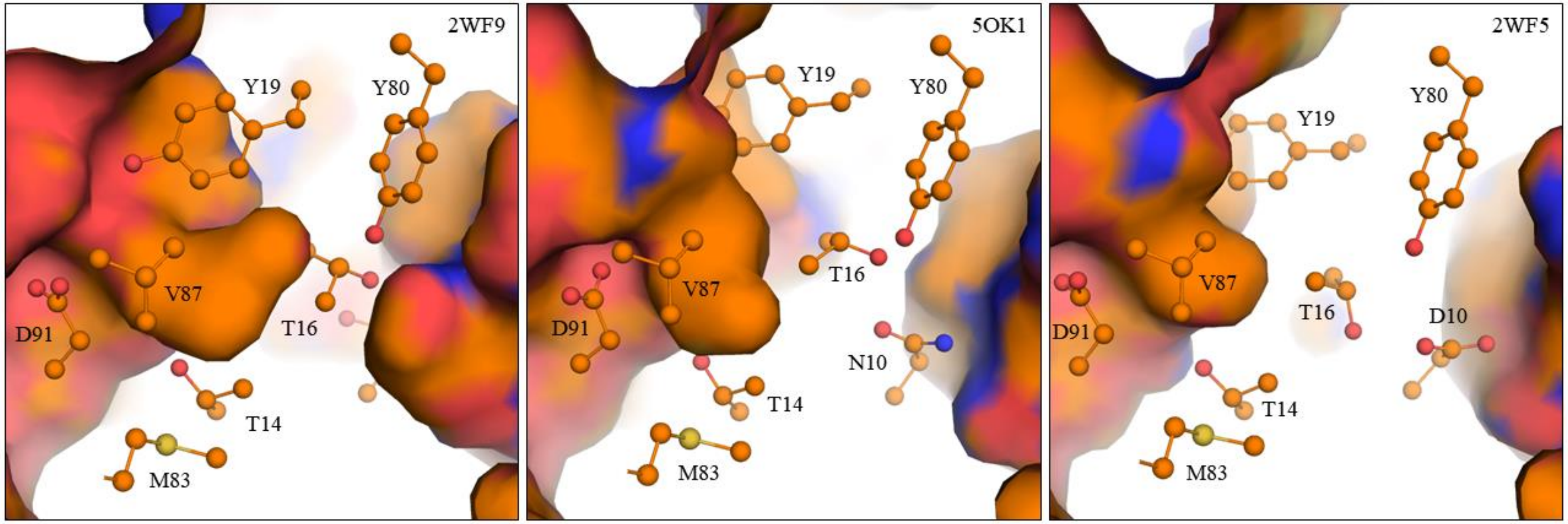


Figure 2-24 – Collapse of a hydrophobic cavity in the hinge of βPGM. Shows a slice through the Van Der Waals' surface of (2WF9) βPGM_{WT}:BeF_3:G6P NAC I crystal structure; (5OK1) βPGM_{D10N}:βG16BP NAC III crystal structure and (2WF5) βPGM_{WT}:MgF_3:G6P TSA crystal structure. The hydrophobic cavity – defined by the side chain atoms of D10, T14, T16, Y19, Y80, M83, V87 and D91 (represented as ball and stick) – gets progressively smaller on approach to the transition state. D10 displaces T16 which occupies the hydrophobic cavity in the NAC III complex. As the domain closes further in the TSA conformation, the hydrophobic packing within the hinge becomes more compact.

2.5 Conclusions

Small Angle X-ray Scattering data indicate that the crystal structures of substrate-free βPGM are biased to a more compact conformation by crystal packing forces. Structural ensembles obtained using MD simulations adopt a more open conformation which better describes the solution SAXS data. This ensemble stabilises a more open conformation by burial of the hydrophobic side chain of Y19 into a hydrophobic cavity within the hinge region. This reveals a mechanism in which an unstable hydrophobic cavity found in the product complex stabilises domain opening and product release. MD simulations also highlight the propagation of a local artefact, namely the coordination of Mg by D10, to large scale conformational changes within the enzyme. Furthermore, domain orientational dynamics in βPGM are biased substantially by forcefields which over-stabilise salt bridges.

3 Describing thermodynamic coupling between phosphodianion binding and hinge bending

3.1 Introduction

3.1.1 The phosphodianion-driven activation framework

In general, enzymes operate by sequestering the substrate in an environment which stabilises the transition state via extensive electrostatic interactions. Such an environment is established via substrate-induced conformational transition from an inactive-open conformation to an active-closed conformation, which may involve whole domain reorientation or more subtle rearrangements of short loops. The phosphodianion-driven enzyme activation framework states that the intrinsic binding energy, derived from the interaction between an inert phosphodianion group from the substrate and a positively charged distal site in the enzyme, is used to offset the energy required to transition from an open-inactive conformation to a closed-active conformation. A general approach to study this effect involves the comparison of reaction kinetics for the whole substate (SP), for the phosphodianion-truncated substrate (S) and for the phosphite-activated turnover of S (S·HPi). The total intrinsic phosphodianion binding energy (IBE_T) is given by the ratio of the second-order rate constants (k_{cat}/K_m) for the reactions involving S and SP.

$$\mathrm{IBE_T} = \mathrm{RTln}\left(\frac{(\mathrm{k_{cat}}/\mathrm{K_m})_{\mathrm{SP}}}{(\mathrm{k_{cat}}/\mathrm{K_m})_{\mathrm{S}}}\right)$$

A third-order rate constant (which has the general form $k_{cat}/(K_S K_{HPi})$) can be derived from initial velocity data collected at multiple concentrations of S and HPi. The activation barrier associated with the turnover of S is derived from the second-order rate constant for that reaction. The smaller barrier, associated with the phosphite-activated turnover of S, can be derived from the third-order rate constant for that reaction. The change in energy barrier represents the stabilisation of the transition state by phosphite binding. This is typically much larger than the intrinsic binding energy of the phosphite alone, which is offset against the energy required to form a closed-active conformation. This approach has been applied to several enzymes involved in central metabolic pathways, including glucose-6-phosphate isomerase (GPI), glucose-6-phosphate dehydrogenase (G6PDH) , 6-phosphogluconate dehydrogenase (6GPDH) [149]; Triose phosphate isomerase (TIM) [150]; orotidine 5-monophosphate (OMPDC) [151]; glycerol 3-phosphate dehydrogenase (GPDH) [152]; and in βPGM [153]. Together these studies show that the IBE_T is typically between 11-13 kcal/mol, of which ~50% can be accounted for by the IBE of HPi [154].

A second approach is to measure the kinetic consequences of distal site alanine mutations. The R269 contributes 9.1 kcal/mol to transition state stabilisation in GPDH. 6.7 kcal/mol can be recovered by the addition of 1M guanidine [155]. R235 contributes 5.6 kcal/mol to the transition state in OMPDC for the decarboxylation reaction and 7.2 kcal/mol to that of the deuterium exchange reaction [156]. K12 in TIM contributes 7.8 kcal/mol to the TS stabilisation in TIM, but also makes a smaller contribution to turnover of the truncated substrate, which suggests an additional role in stabilisation of the enolate anion [157]. These experiments show that the majority of the phosphodianion binding energy is delivered through interactions with a single arginine or lysine residue in the distal site.

3.1.2 Distal site mutants in βPGM

In βPGM two distal site mutants were investigated, βPGM_{R49A} and βPGM_{R49K}. These mutations had only a small local effect in the conformation of the cap domain in the substrate free enzyme. Coordination of the 6'phosphate in the βPGM_{WT}:MgF_3:G6P complex, occurs via N118-Nδ_2, K117-N, S116-Oγ and two hydrogen bonds to R49 sidechain guanidino group. In the βPGM_{R49K}:MgF_3:G6P complex K49 can only make one hydrogen bond to the 6'phosphate, whilst in the βPGM_{R49A}:MgF_3:G6P complex K117-Nζ is recruited to partially compensate for the loss of R49 sidechain. Otherwise these distal site mutant TSA complexes are very similar to their equivalent WT complexes. These observations are corroborated by the backbone $^{1}H^{15}N$ and ^{19}F chemical shifts.

A substantial decrease in stability of the transition state complex was measured via G6P titration, monitored by ^{1}H NMR (300-fold for R49K and 2000-fold for R49A). The Michaelis-Menten kinetic parameters were reported for WT (k_{cat} = 24.5 ± 0.7 s^{-1}, K_m = 92 ± 6 μM). However, the K_m values for the R49K and R49A were increased beyond the accessible range of βG1P concentrations. The lower limit on k_{cat} and K_m can be specified (R49A: k_{cat} > 12 s^{-1} , K_m > 600 μM, derived from [139] Fig S8) meaning the maximum k_{cat} effect is ~2-fold. To account for the 2000-fold change in the stability of the TSA complex, the K_m must be significantly larger than 600 μM.

In general, destabilisation of the closed complexes (ES_C and $ES_C^{\ddagger}$) can have either of two consequences. 1) If the Michaelis complex is closed, then its destabilisation will have a substantial K_m effect. In the case of a large destabilisation, the Michaelis complex switches to an open conformer and domain closure becomes part of the rate limiting step resulting in a small k_{cat} effect. 2) If the Michaelis complex is dominated by an open conformation, then its destabilisation will have a large k_{cat} effect and a small K_m effect. Thus, distal site binding energy is utilised to deliver rate enhancements either through 1) an increase substrate affinity or 2) a reduced activation barrier (**Figure 3-1**). βPGM_{R49A} and βPGM_{R49K} exhibit a dominant K_m effect and thus βPGM$_{WT}$ has a closed Michaelis complex which switches to a more open conformer in the distal site mutants.

An implicit assumption of this framework is that the distal site mutation affects equally the stability of the closed ground state complex (ES$_C$) complex and the closed transition state complex (ES$_C$‡). Consequently, the measured destabilisation of the βPGM$_{WT}$:MgF$_3$:G6P TSA (ES$_C$‡) would also apply to ES$_C$. This assumption was supported by the assembly of canonical transition state analogue complexes, with negligible chemical shift changes. If this assumption were false, and the phosphodianion interaction provided an additional specific stabilisation of the TS, then the TSA complex would be compromised by the mutation resulting in structural and/or chemical shift changes – none of which were observed.

A relationship was uncovered whereby enzymes with lower catalytic efficiencies have an open Michaelis complex and distal site binding energy primarily contributes to lowering of the activation barrier. On the other hand, enzymes with higher catalytic proficiencies have a closed Michaelis complex and distal site binding energy primarily contributes to higher substrate affinity. In doing so, this introduces the potential for substrate and/or product inhibition. A structural model of the inhibition of βPGM by βG1P was obtained using a D170N mutant, which reduces the activity sufficiently to allow crystallisation of a βPGM$_{D170N}$:βG1P complex. The substrate-inhibited complex achieves TS-like domain closure without proximal site occupancy. The coordination of the distal phosphate is equivalent to that of the WT closed complexes. Furthermore, the structure of a βPGM$_{WT}$:Pi complex was obtained by x-ray crystallography. The enzyme adopts an open conformation, the inorganic phosphate is bound to the distal site via R49 sidechain guanidino group and K117 alkylammonium sidechain. The length and flexibility of the sidechain torsions allow K117 to coordinate the phosphate whilst maintaining an open conformation. Interactions between the inorganic phosphate and residues from the core domain (N117-N, N118-Nδ$_2$ and S116-Oγ) cannot be maintained by the open conformer. This complex represents an initial mode of substrate binding to the open conformer (**Figure 3-2**).

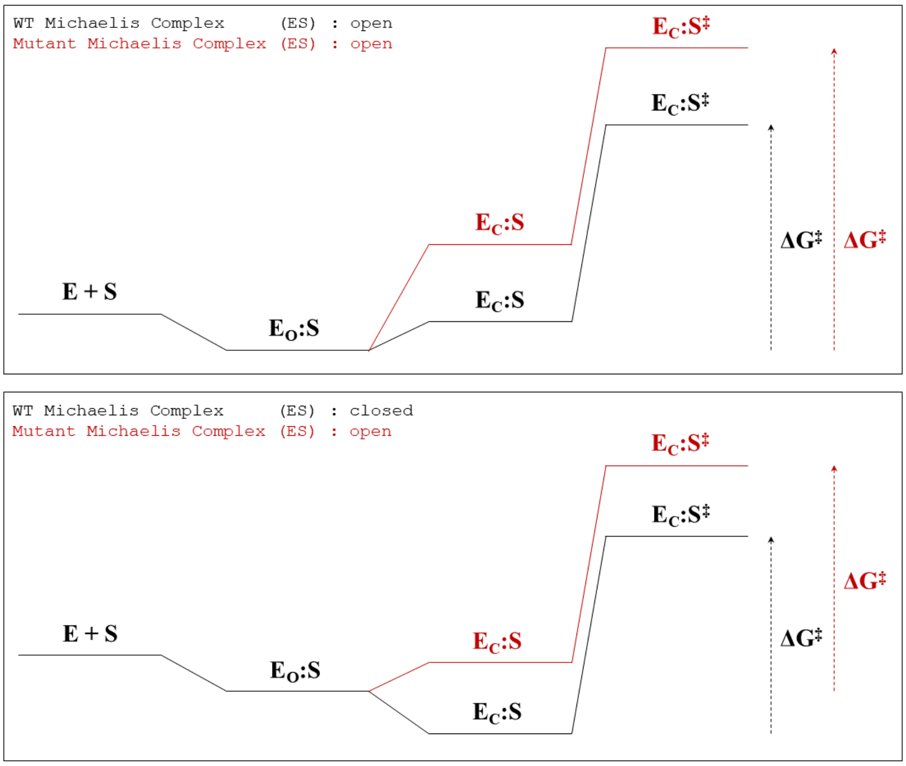


Figure 3-1 – Phosphodianion-driven activation framework. Energy level diagrams showing the consequences of distal site mutation on enzymes which have (TOP) an open Michaelis complex (BOTTOM) a closed Michaelis complex. The WT energy levels are shown in black, and the distal site mutant levels are shown in red. An assumption of this framework is that the closed ground state complex (E_C:S) and transition state complex (E_C:S‡) are affected equally by the distal site mutation, i.e. the phosphodianion binding energy is not utilised to specifically stabilise the transition state E_C:S‡. A k_{cat} effect is only observed if E_C:S is destabilised beyond E_O:S such that domain closure becomes part of the rate-limiting step. The K_m is determined by the difference in free energy between E+S and the lowest energy ES complex, therefore the K_m effect drops off after E_C:S is destabilised beyond E_O:S. Hence, the top panel represents a large k_{cat} effect and small K_m effect, and the bottom panel represents a small k_{cat} effect and a large K_m effect.

3.1.3 How is phosphodianion binding coupled to domain closure?

In summary, βPGM and several other enzymes utilise the intrinsic binding energy between an inert phosphodianion group and a positively charged distal site to stabilise a closed, catalytically competent conformation. However, the mechanism by which these two thermodynamic processes are coupled is not fully understood. In general, biomolecular energy transduction depends on the ability of proteins to couple thermodynamic processes which require energy with others that release energy. A classic example is the coupling of proton translocation to ATP synthesis by the ATP synthase [158]. However, in the case of βPGM and the phosphodianion-driven enzyme activation framework, the question becomes: how is domain closure coupled to the R49-phosphodianion interaction? That is, how is energy transferred/communicated from the distal site to the hinge? Coordination of the phosphodianion by R49 is present in the open βPGM$_{\mathrm{WT}}$:P$_{\mathrm{i}}$ complex. The interaction with the core domain is mediated by the side chain of K117, rather than by S116-Oγ, N117-N and N118-Nδ$_2$ (**Figure 3-2**). A more specific question can thus be posed: why is the R49-phosphodianion interaction energy not fully realised until the enzyme closes?

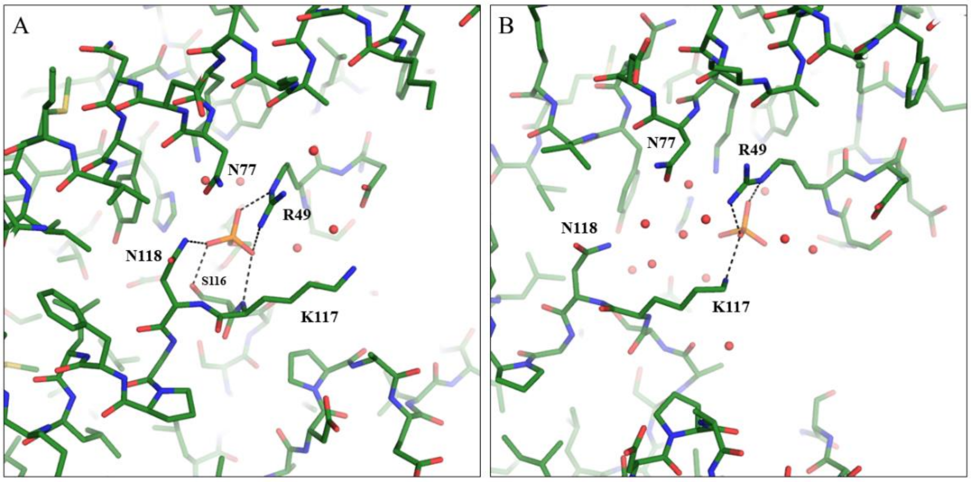


Figure 3-2 – Distal site coordination in open and closed conformations. Distal site interactions in (A) the closed βPGM$_{\mathrm{WT}}$:MgF$_3$:G6P complex (PDB: 2WF5) and (B) the open βPGM$_{\mathrm{WT}}$:P$_{\mathrm{i}}$ complex (PDB: 6H93). Hydrogen bonding to the distal phosphate is illustrated with black dashed lines. Water molecules are represented as red spheres. The hydration of Pi in the open βPGM$_{\mathrm{WT}}$:P$_{\mathrm{i}}$ complex is greater than in the closed βPGM$_{\mathrm{WT}}$:MgF$_3$:G6P complex.

3.2 Results

3.2.1 A communication pathway between the hinge and the distal site

A library of NMR resonance assignments of 22 distinct ^2H,^{13}C,^{15}N-labelled βPGM complexes has been collected over the past 20 years (**Table 7-1**). This chemical shift library was used to identify residues which would potentially be involved in coupling distal site phosphodianion binding to a conformational change in the hinge. Backbone ^1H-^{15}N chemical shifts for a selection of residues are plotted for each complex in **Figure 3-3**. Residues in the hinge region (D15, T16, I84, S88 and A90) are specifically sensitive to the transition between open, partially closed NAC and fully closed TSA conformations. However, the chemical shifts of M83, V81, N79, D78, N77, K74 and A73 also act as distinct signatures of closure. These residues are found within the helix which connect the hinge to the distal site via N77 which forms a carboxamide pairing with N118, and a planar stacking interaction with R49 (**Figure 7-7**). The ^1H chemical shift changes inform on hydrogen bond lengths within the helix. A shorter hydrogen bond results in polarisation of the N–H bond, de-shielding of the proton and a downfield chemical shift and *vice versa*. The chemical shifts therefore reveal a concerted hydrogen bond shortening in N77/V81 and hydrogen bond lengthening closer to the hinge in N79/M83, upon closure.

It is less obvious from the crystal structures that any structural coupling exists between the hinge and N77 – the backbone atoms from M83-N77 appear virtually superimposable (**Figure 3-5A**). However, shiftX2 chemical shift predictions, calculated using the βPGM crystal structures, do not accurately reproduce the measured chemical shifts changes (**Figure 7-8**). Therefore, the structural changes which occur during domain closure in this region, are greater in solution than suggested by the crystal structures. This is not surprising owing to the extensive contacts with neighbouring molecules within the crystal (**Section 2.2.1**). Nevertheless, some relevant signatures of closure can be detected. The A73/N77 and N77/V81 hydrogen bonds are compressed whilst the R75/N79 and Y80/I84 hydrogen bonds are stretched (**Figure 3-4**). Relaxation of the Y80/I84 hydrogen bond will affect the torsions of I84 which are key determinants of domain orientation.

Furthermore, Linear Discriminant Analysis (LDA, **Section 6.5**) was used to identify structural signatures of domain closure from the backbone coordinates between N77 and M83. For each of the 64 βPGM crystal structures, interatomic distances were calculated between all pairs of backbone atoms between N77 and M83. The distances calculated between atoms found within the same helical turn (between O_i and N_{i+4}) and not from within the same residue (except between N_i and O_i) were taken forward for LDA. Two linear discriminants (LD1 and LD2) can be derived for a data set with three classes (open, NAC, TSA). The open complex is separated from the closed (NAC & TSA) complexes by LD1. The NAC and TSA complex are separated by LD2.

(**Figure 3-5B**). The LDA coefficients which are predictive of the open/close transition, receive small contributions from many distances derived from across the full length of the helix between N77 and M83. In particular, the distances between the backbone atoms of N77 and V81 are signatures of domain closure (**Figure 3-5C**).

Taken together, the crystal structures and chemical shift data indicate that the helical backbone between N77 and M83 is sensitive to domain orientation and may serve as a pathway for communication of distal site binding energy to the hinge via N77 sidechain. To test this model, N77 was mutated to an alanine which would break the chain of communication and decouple the hinge conformation from occupancy of the distal phosphate site.

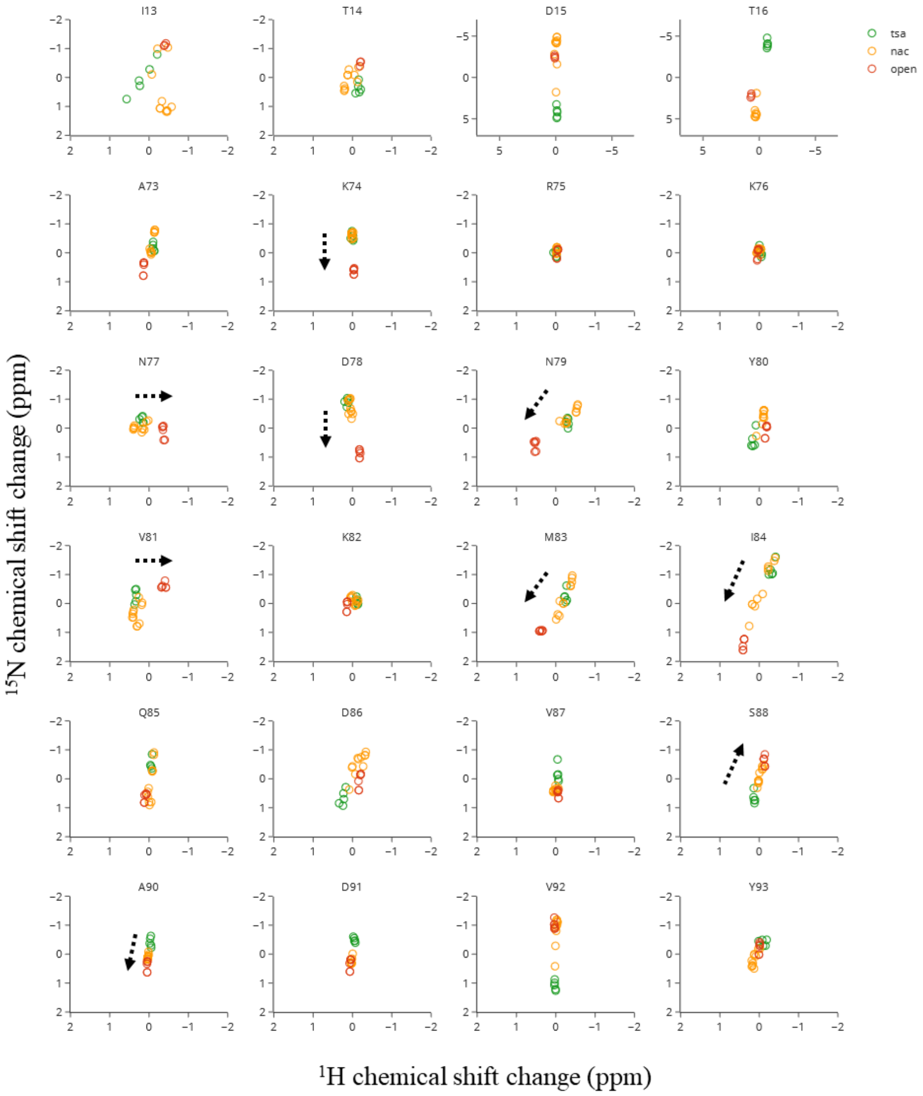


Figure 3-3 – βPGM Chemical Shift Library. Each subplot shows the ^{1}H and ^{15}N chemical shift changes in ppm, for a series of βPGM open, NAC, and TSA complexes, coloured red, orange and green, respectively. Chemical shift changes are calculated relative to the midpoint of the data for each residue. The x and y-axis ranges are the same for all residues (± 2 ppm), except D15 and T16 which show particularly large chemical shift changes (± 5 ppm).

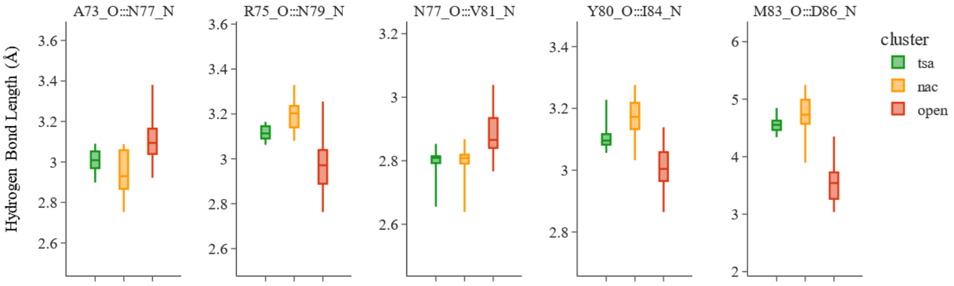


Figure 3-4 – Communication between N77 and I84 via the Y80/V81 peptide bond. Backbone hydrogen bond lengths in the 70s helix are shown for the open (red) NAC (yellow) and TSA (green) crystal structures. Domain closure causes a shortening of the N77-V81 hydrogen bond and a concomitant lengthening of the Y80-I84 hydrogen bond.

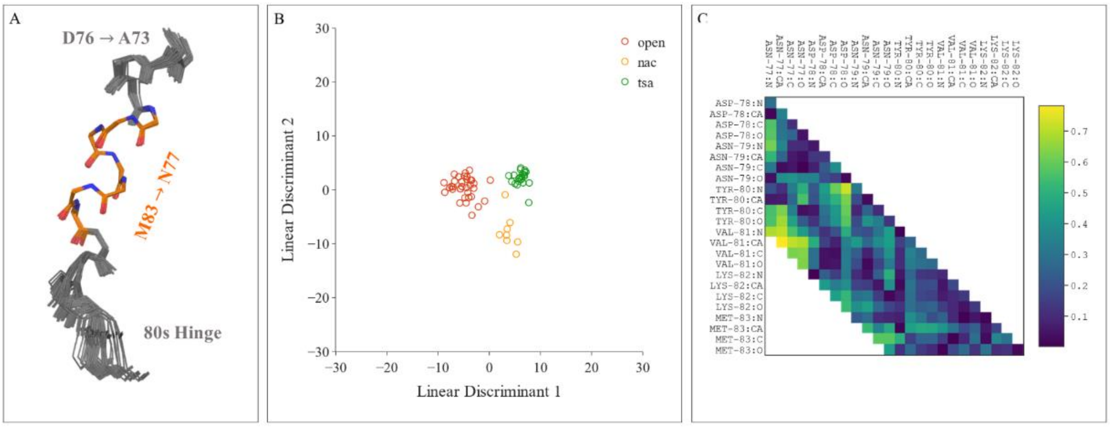


Figure 3-5 – Conformation of backbone from M83-N77 is sensitive to domain closure. (A) Backbone atoms of residues S88 – A73 for all crystal structures of βPGM deposited to the PDB. Structures are aligned on residues M83 – N77. C, N and O atoms are coloured orange, blue and red, respectively. (B) The linear discriminants of the interatomic atomic distances, calculated between all pairs of backbone atoms which occupy the same helical turn from residues M83 to N77. Open structures are coloured red, partially closed NACs are coloured yellow and fully closed TSA structures are coloured green. (C) A heat map showing the contribution made by each interatomic distance to LD1 and LD2. The colour indicates the sum of the weights for LD1 and LD2 for each atom pair. The weights of all intra-residue distances (except between N and O) and for all atoms which were NOT found within the same helical turn (between O_i and N_{i+4}) were also set to zero.

3.2.2 Solution properties of the βPGM_{N77A}:MgF_3:G6P TSA complex

The effect of the N77A mutation on the stability of the βPGM_{N77A}:MgF_3:G6P TSA complex was determined by serial titration of G6P against 0.5 mM βPGM_{N77A} in 50 mM K+ HEPES pH 7.2, 5 mM MgCl2, 20 mM NaF and 2 mM NaN3. ^{1}H NMR was used to monitor the decrease in intensity W24 indole resonance of the βPGM_{N77A} substrate-free complex and the increase in intensity of the K117 backbone amide resonance of the βPGM_{N77A}:MgF_3:G6P complex. These resonances are well-resolved and report on G6P binding and formation of the fully closed βPGM_{N77A}:MgF_3:G6P complex (**Figure 3-6**). The pseudo dissociation constant (K_D) for βPGM_{N77A} was derived by fitting the change in intensity to the following equation.

$$I = I_0 \frac{([G6P] + [E]_0 + K_D) - \sqrt{([G6P] + [E]_0 + K_D)^2 - 4[E]_0[G6P]}}{2[E]_0}$$

$$[G6P] = \text{dilution} \times [G6P]_0$$

The titration data was linear up to ~80% saturation, indicating approach to the tight-binding limit. Therefore, Monte Carlo sampling was used to calculate parameter uncertainties more rigorously from the estimated experimental uncertainties. A 5 % uncertainty in $[E]_0$ and in serial pipetting volumes, corresponded to a ~17 % uncertainty in $K_D = 30 \pm 5$ μM . The equivalent value for the βPGM_{WT}:MgF_3:G6P complex was measured by isothermal titration calorimetry (K_D = 1 μM) [64]. Therefore, the stability of the βPGM_{N77A}:MgF_3:G6P is attenuated 30-fold compared to WT. The contribution of N77 sidechain to transition state stability can be estimated using the following equation.

$$\Delta G^{\ddagger} = RT\ln\left(\frac{K_{D,N77A}}{K_{D,WT}}\right) = 2.0 \text{ kcal/mol}$$

^{1}H^{15}N TROSY NMR spectra of the βPGM_{N77A}:MgF_3:G6P complex were analysed to investigate the behaviour of this complex in solution. Resonances assignments were obtained by TROSY transfer from the βPGM_{WT}:MgF_3:G6P complex. The weighted aggregate ^{1}H^{15}N chemical shift changes (**Figure 3-7**) are small in magnitude (< 0.1 ppm) indicating that the behaviour of the complex is typical of a TSA complex. Slightly larger chemical shift changes are found in the vicinity of N77, N118, R49 and within the hinge, indicating subtle conformational changes in these regions. Some residues around the mutation site could not be identified confidently by transfer assignment due to large chemical shift changes.

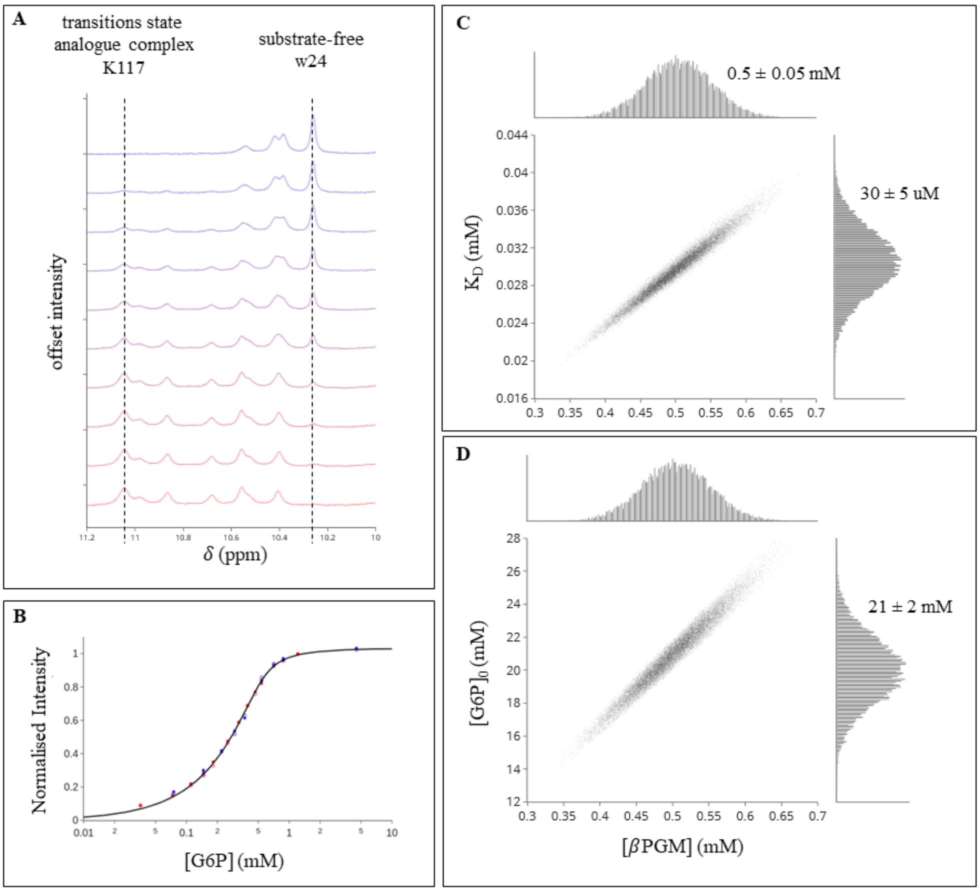


Figure 3-6 – βPGM$_{WT}$ MgF$_3$ G6P titration. (A) ^{1}H NMR spectra titration series showing increasing intensity of the amide K117 resonance of the βPGM$_{WT}$:MgF$_3$:G6P complex; and the decreasing intensity of the W24 indole resonance of the βPGM$_{WT}$ substrate free complex. Progression of the titration is show from top (blue) to bottom (red) (B) Normalised intensities plotted against G6P concentration on a linear-log scale for each of two replicates, combining data from the K117 and W24 resonances. Least-squares linear regression of the concentration-response data to a quadratic 1:1 binding model using Monte Carlo sampling yields correlations between [βPGM] and (C) K_D (D) [G6P]$_0$. The uncertainties in [βPGM] and [G6P]$_0$ determine the uncertainty in K_D.

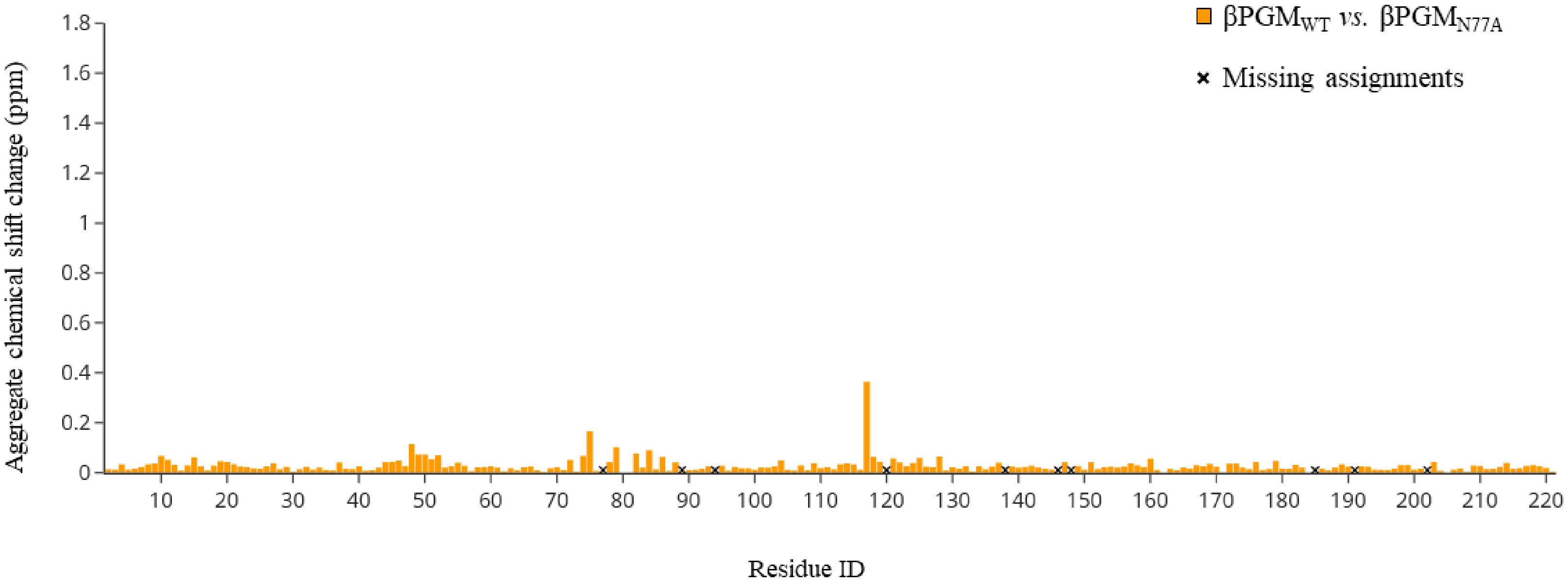


Figure 3-7 – Chemical shift changes between *β*PGM$_{WT}$ and *β*PGM$_{N77A}$ TSA complexes. Aggregate backbone ^{15}N-^{1}H chemical shift changes are plotted against residue ID. Modest chemical shift changes are observed throughout the majority of backbone atoms, indicating a canonical TSA conformation is established in solution. Larger perturbations are observed in residues around the mutation site, in the 50s helix which packs against the mutation site and around N118 whose side chain hydrogen bonds with that of N77. Missing assignments, including A77, are highlighted with a black cross.

3.2.3 G1P turnover kinetics

The effect of the N77A mutation on steady state kinetics were measured using a Glucose-6-phosphate dehydrogenase (G6PDH) coupled assay to monitor the conversion of βG1P to G6P. βG16BP was used as a phosphorylating agent to generate fast linear kinetics, free of an allomorphic lag phase [69]. Initial rates were measured across a range of βG1P (0-1800 μM) and βG16BP (0-70 μM) concentrations. Historically for βPGM and for other mutase enzymes, a Ping-Pong Bi Bi steady-state model has been used to derive kinetic parameters. The King-Altman diagram which underpins the derivation of this model is given along-side the equivalent mechanism utilised by βPGM (**Figure 3-8**). Ping-Pong Bi Bi describes a sequential mechanism with two substrates and two products, where the first product is released before the second substrate binds. This has close parallels with a mutase mechanism, which has one substrate and one product with an intermediate which is released and rebinds. The mutase mechanism describes the transfer of a functional group (e.g. phosphate) from one position to another within the same molecule. However, the Ping-Pong Bi Bi more accurately describes a transferase mechanism, where a functional group is transferred from one molecule to another. This has consequences for the derivation of steady state velocity equation. During the derivation it is assumed that the concentration of the first product is zero, as is usually the case when designing stead-state kinetics experiments. However, βG16BP acts as both the first product and the second substrate. Therefore, the concentration of the first product cannot be assumed to be zero (**Section 6.4**). Previous attempts to fit steady-state kinetics data from βPGM have failed at higher concentrations of βG16BP, which were attributed to multimeric interactions with Mg^{2+} ions in the reaction buffer and a substantial back reaction giving rise to higher steady-state concentration of βG1P [69]. This also results in higher levels of βG1P inhibition than predicted from the model. In the absence of an improved model, minimising the βG16BP concentration will improve the accuracy of the fitted parameters. However, a sufficiently high βG16BP concentration is required to define the maximum velocity with reasonable uncertainty. Initial rates were measured using a range of βG16BP concentrations, which offer a compromise between these two factors.

Fitting initial rates acquired at low βG16BP concentration implicitly requires extrapolation to higher concentrations. This will give rise to greater uncertainties in the parameter estimates. Fitting was used with Monte-Carlo sampling to accurately propagate experimental uncertainties though to the fitted parameters. The values for WT are k_{cat} = 164 ± 25 s^{-1}, $K_{\beta g1p}$ = 50 ± 13.8 µM, $K_{\beta G16BP}$ = 2.4 ± 0.5 µM, K_i = 442 ± 98 µM. The values for N77A are k_{cat} = 104 ± 7 s^{-1}, $K_{\beta G1P}$ = 654 ± 60 µM, $K_{\beta G16BP}$ = 19.4 ± 0.5 µM, K_i = 867 ± 72 µM **Figure 3-9**. The N77A mutation results in an approx. 10-fold reduction in K_m for βG1P and βG16BP, a negligible effect on k_{cat} and a 2-fold effect on K_i. The contribution of N77 sidechain to the stability of each complex can be estimated from the fold-change in the corresponding Michaelis constant ($\Delta G_{\beta G1P}$ = 1.5 kcal/mol, $\Delta G_{\beta G16BP}$ = 1.2 kcal/mol).

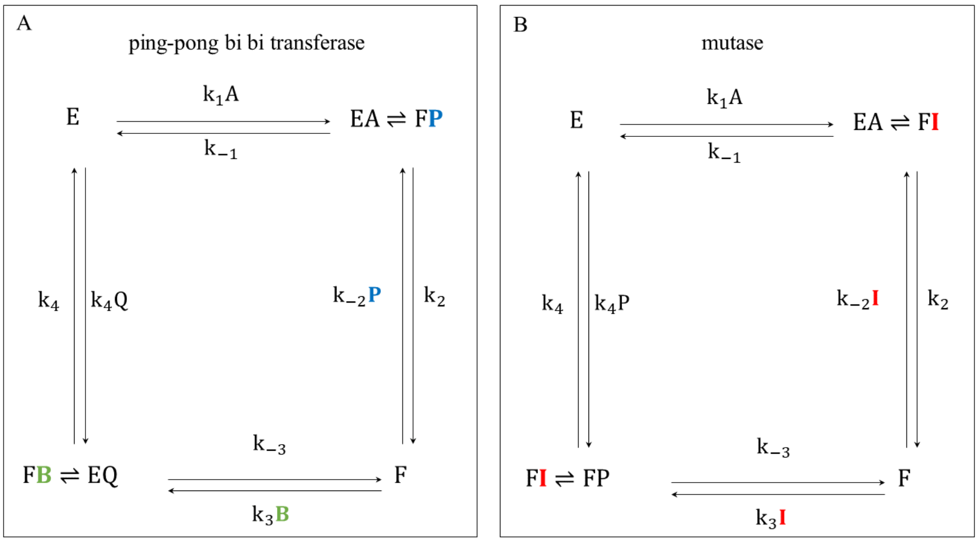


Figure 3-8 – King-Altman diagrams. (A) A ping-pong bi bi mechanism utilised by transferase enzymes. The product of the first step (blue) is released followed by binding of the second substrate (green). (B) The mechanism utilised by a mutase enzyme. The product of the first step is an intermediate (red) which rebinds to the enzyme in an alternate orientation for the second step of the reaction.

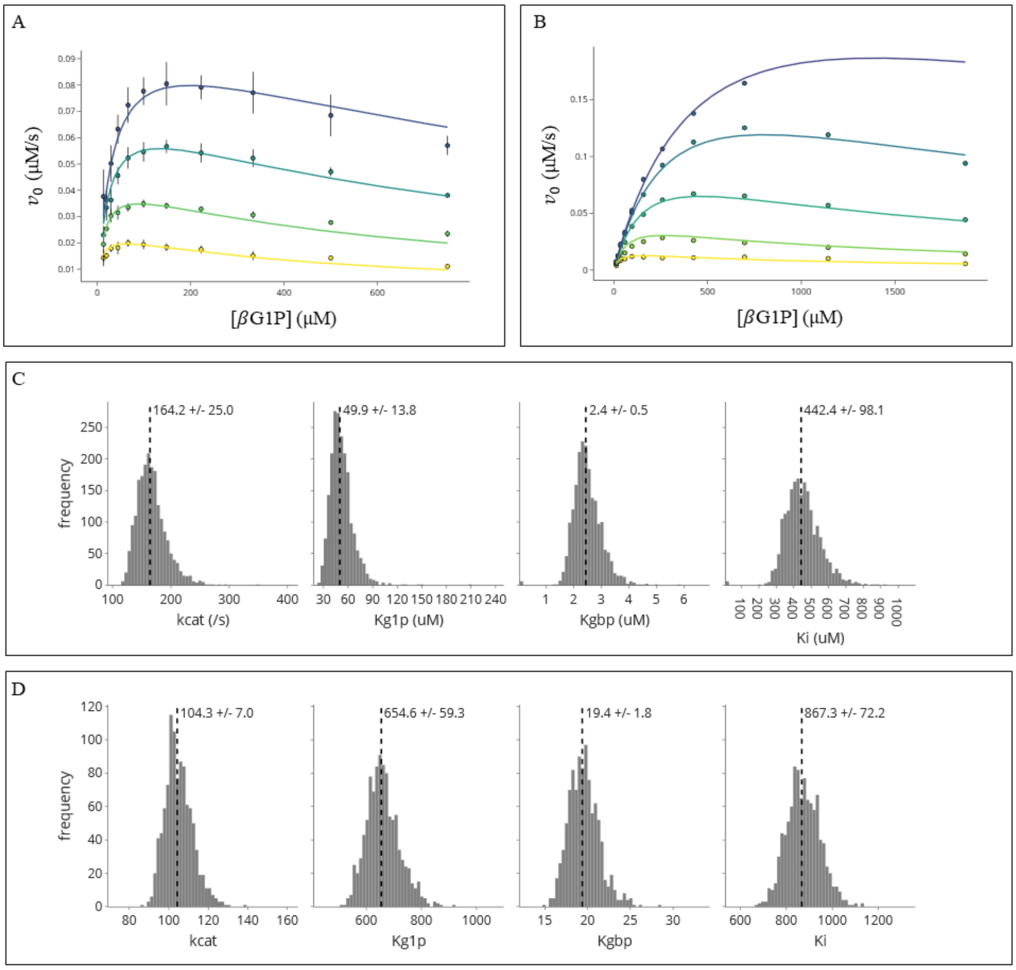


Figure 3-9 – Steady-state kinetics data. The initial rate of reaction (v_0) plotted against βG1P concentration at multiple βG16BP concentrations, with lower concentrations coloured yellow and higher concentrations coloured blue, for each of two variants (A) βPGM_{WT} (B) βPGM_{N77A}. The ping-pong Bi Bi parameter distributions derived from least-squares non-linear regression with Monte Carlo error propagation, for (C) βPGM_{WT} and (D) βPGM_{N77A}. The affinity for βG1P and βG16BP are reduced 10-fold in the N77A variant, while k_{cat} is unaffected.

3.2.4 BeF3, G6P titration

The effect of the N77A mutation on the stability of the βPGM:BeF_3:G6P GSA complex was determined by serial titration of G6P against either 0.5 mM βPGM_{N77A} or 0.5 mM βPGM_{WT} in 50 mM K+ HEPES pH 7.2 + 10 mM $MgCl_2$, + 5 mM BeF_3 + 20 mM NaF + 2 mM NaN_3. The titration was monitored by ^{1}H-^{15}N BEST-TROSY NMR. The product complexes are relatively low affinity, and thus the bound and unbound species exist in fast chemical-exchange, resulting in incremental chemical shift changes. The start and endpoints of the titration were independently assigned by comparison with the resonance assignments of $^{2}H^{13}C^{15}N$ labelled βPGM_{WT}:BeF_3 and βPGM_{WT}:BeF_3:G6P complexes, respectively. Trajectories showing chemical shift changes during the titration for each variant are shown in **Figure 3-10A**. The aggregate chemical shift changes were fitted to a 1:1 binding model (**Figure 3-10B**).

$$\Delta\delta = \Delta\delta_{\max}\frac{([\mathrm{G6P}] + [\mathrm{E}]_0 + \mathrm{K_D}) - \sqrt{([\mathrm{G6P}] + [\mathrm{E}]_0 + \mathrm{K_D})^2 - 4[\mathrm{E}]_0[\mathrm{G6P}]}}{2[\mathrm{E}]_0}$$

The mean K_D derived from hinge residues is reduced by an order of magnitude for N77A compared to WT ($K_{D,WT}$ = 3.2 ± 0.2 mM, $K_{D,N77A}$ = 24.9 ± 4.4 mM). The corresponding free energy contribution is ΔG_{G6P} = 1.2 kcal/mol. The Δδmax derived from βPGM_{N77A} dataset was calculated as a percentage of the corresponding values from βPGM_{WT} (ΔΔδmax). The values for hinge residues which display a linear trajectory between open, NAC and TSA complexes, and thus are specifically sensitive to domain orientation, have a ΔΔδmax = 82 %. Therefore, the stability of the ground state product complex has been destabilised sufficiently to introduce a significant population of an open complex. The chemical shift for the remaining residues may be interpreted using a two-state model (open/closed) at a ratio 1:4. Therefore, if the ΔΔδmax is significantly higher or lower than 82% this would indicate that the theoretical chemical shift of the closed complex in isolation is altered significantly. Making this assumption, significant structural changes through the 70s helix can be detected (**Table 3-1).**

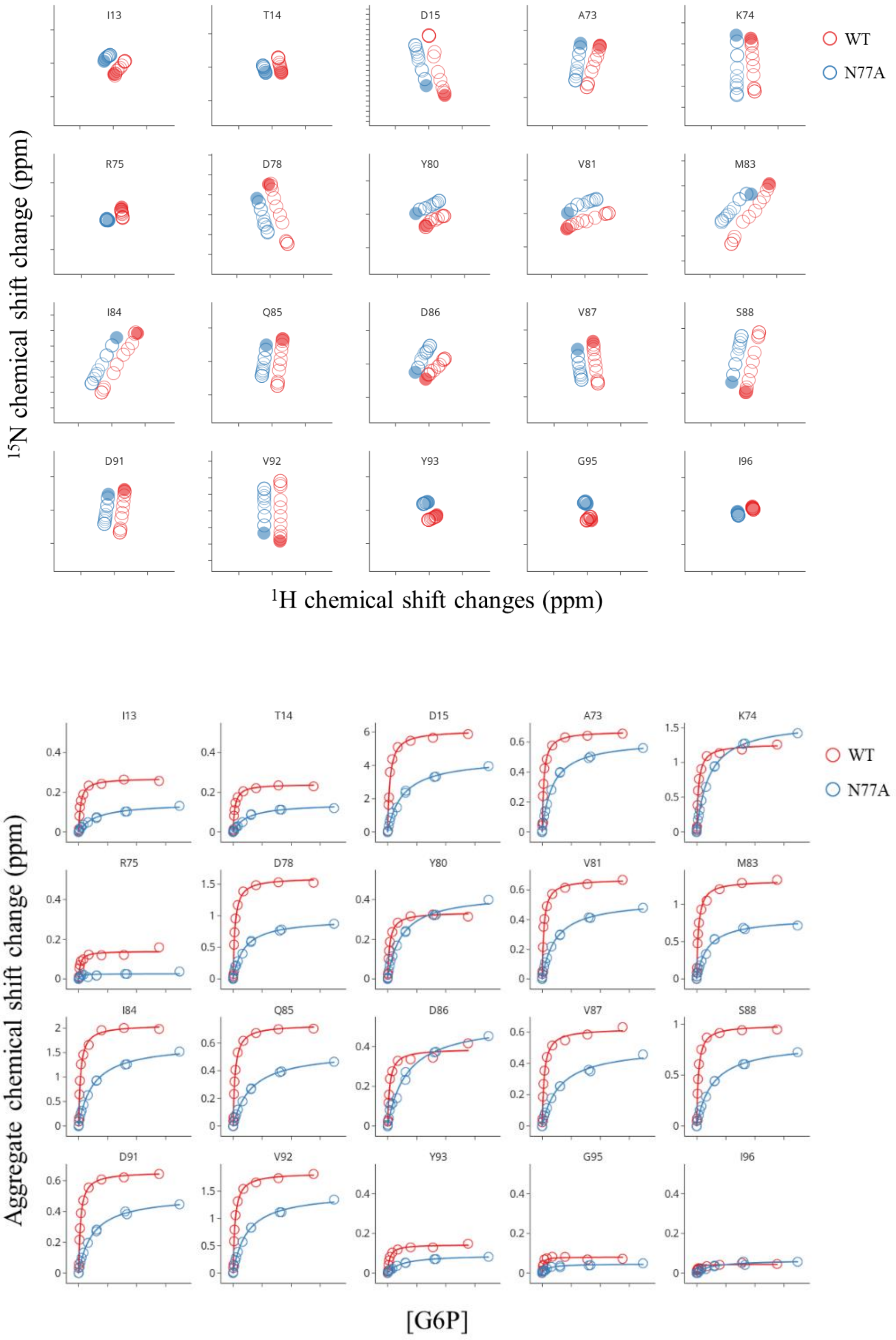


Figure 3-10 – G6P titration into a βPGM:BeF$_3$ complex. (TOP) The ^{15}N-^{1}H TROSY trajectories from the titration of G6P into either a βPGM$_{WT}$ (red) or a βPGM$_{N77A}$ (blue) trifluoroberyllate complexes. Chemical shifts are centred on the midpoint of the trajectory. Data from each variant are perpendicularly-offset from one another by ± 0.1 ppm. The titration endpoints are represented as closed circles. (BOTTOM) The aggregate chemical shift changes are plotted against G6P concentration (0-180 mM). Locally fit 1:1 binding models are show as lines.

residue	K_D (mM) WT	K_D (mM) N77A	$\Delta\delta_{max}$ (ppm) WT	$\Delta\delta_{max}$ (ppm) N77A	$\Delta\Delta\delta_{max}$ (%)
I13	3.3	29.7	0.3	0.1	54.2
T14	3.1	23.9	0.2	0.1	59.7
D15	3.9	26.4	6.1	4.5	73.1
A73	2.8	18.0	0.7	0.6	91.4
K74	3.2	21.4	1.3	1.6	126.3
R75	2.7	-	0.1	0	< 20
D78	3.1	18.9	1.6	1.0	59.7
Y80	3.4	24.1	0.3	0.4	128.4
V81	2.9	23.1	0.7	0.5	79.3
M83	3.1	18.7	1.3	0.8	62.1
I84	3.2	25.0	2.1	1.7	81.6
Q85	3.3	31.1	0.7	0.5	74.3
D86	3.4	33.1	0.4	0.5	135.7
V87	3.5	30.3	0.6	0.5	81.7
S88	2.9	27.0	1.0	0.8	82.9
D91	3.2	24.4	0.7	0.5	77.0
V92	3.3	24.1	1.8	1.5	80.6
Y93	3.4	24.4	0.1	0.1	64.3

Table 3-1 – Parameters derived from titrations of G6P into βPGM:BeF$_3$ complexes. The dissociation constant (K_D) and endpoint chemical shifts (Δδmax) derived from G6P titrations into βPGM$_{WT}$ and βPGM$_{N77A}$ trifluoroberyllate complex, for each of a selection of spin systems monitored by ^{1}H^{15}N-TROSY-NMR. The endpoint chemical shift derived from the titration against βPGM$_{N77A}$ is expressed as a percentage of the corresponding βPGM$_{WT}$ value. The values derived for key hinge resides S88 and I84 are used to estimate the %age closure under saturation (blue). Many residues have a significantly smaller (red) or larger (green) endpoint aggregate shift, than would be predicted purely from a ~20% population of an open conformer.

3.2.5 α-Galactose-1-phosphate titration

The βPGM:BeF$_3$:G6P complex is a step 2 product complex. The effect of the N77A mutation on the stability of the step-1 ground state complex was determined by serial titration of the substrate analogue α-Galactose-1-phosphate (αGal1P) against either 0.5 mM βPGM$_{N77A}$ or 0.5 mM βPGM$_{WT}$ in 50 mM K+ HEPES pH 7.2, 10 mM MgCl2 and 2 mM NaN$_3$. The titration was monitored by ^{1}H-^{15}N BEST-TROSY NMR, which showed incremental chemical shift changes between bound and unbound complexes in fast chemical exchange. The titration was monitored using the well resolved W24 indole resonance. The aggregate chemical shift change is plotted against αGal1P concentration (**Figure 3-11)**. Dissociation constants (K$_D$) and endpoint chemical shifts (Δδmax) values were obtained by fitting the data to a 1:1 binding model. The dissociation constants were 1.1 mM and 11.4 mM for βPGM$_{WT}$ and βPGM$_{N77A}$, respectively. The free energy contribution is $\Delta G_{\alpha Gal1P}$ = 1.4 kcal/mol. The change in endpoint chemical shift for N77A vs. WT is given by ΔΔδmax = 45 %.

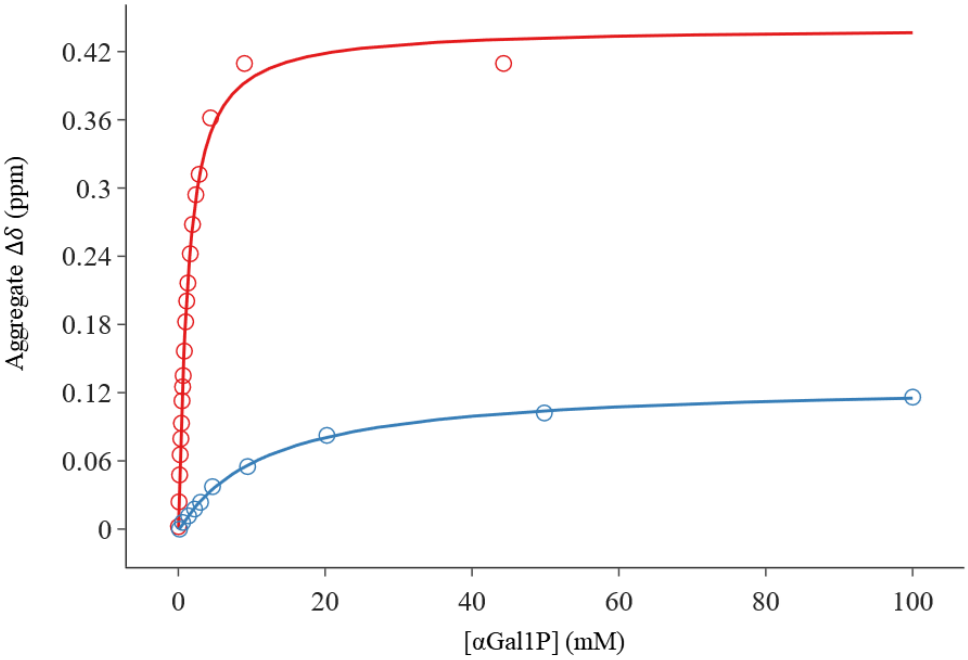


Figure 3-11 – αGal1P titration into substrate-free βPGM. The ^{15}N-^{1}H aggregate chemical shifts change (Δδ) from the titration of αGal1P into either a βPGM$_{WT}$ (red) or a βPGM$_{N77A}$ (blue) substrate free complexes. The aggregate ^{1}H^{15}N-chemical shift changes are plotted against G6P concentration (0-180 mM). The 1:1 binding models, fit by least-squares non-linear regression, are show as lines.

3.2.6 N77A TSA crystal structures

βPGM$_{N77A}$:MgF$_3$:G6P and βPGM$_{N77A}$:AlF$_4$:G6P TSA complexes were crystallised in the orthorhombic space group $P2_12_12_1$, under standard conditions [17,64,69,139], and their structures solved by molecular replacement to 1.5Å. (Section 7.8). These complexes align to the corresponding βPGM$_{WT}$:MgF$_3$:G6P and βPGM$_{WT}$:AlF$_4$:G6P TSA complexes with a 0.28 Å and 0.25 Å Cα-RMSD, respectively. The C_β atom of A77 occupies a similar position to that of N77, packing against the R49 sidechain guanidino group. The missing carboxamide group in A77 is replaced by a water molecule which hydrogen bonds to the carboxamide of N118. The backbone at A77 is freed from the restraints imposed by side chain hydrogen bonding to N118. Consequently, the backbone hydrogen bond between A77-O and V81-N is slightly longer ($\Delta d \simeq 0.2$ Å) (**Figure 3-12, Figure 7-6**).

The domain orientational Euler angles of all but three βPGM TSA complexes deposited to the PDB are within 0.5° of 2WF5. The three outliers, 1Z4O$_A$, 6I03$_A$ and 5OLY$_A$, are compromised in domain orientation due to coordination of αGal1P, 5’-fluorotryptophan labelling at W24 and a D10N mutation, respectively. The crystal structures of the N77A TSA complexes also fall outside of the main distribution (pitch = 0.5 °, yaw = 1.2 °, θ = 1.4 °) (**Figure 7-5**). The per-atom root-mean-square-fluctuation (RMSF) (Section 5.2.9) was calculated, providing a baseline profile of structural variability within the canonical TSA structures. The RMSF was then calculated by comparison of the βPGM$_{N77A}$ TSA structures with each of the canonical TSA structures. The final baseline-subtracted RMSF profile for the βPGM$_{N77A}$ TSA structures is shown in **Figure 3-13A**, illustrating the collective motions within the cap domain and in the hinge. A significant RMSF is also observed in the side chains of I56, W24, H20, K76, Y80, I84 and V87, illustrating the connection between the hinge and the internal hydrophobics of the cap domain (**Figure 3-13B**). The distal site architecture remains unaffected. The distal phosphate is coordinated directly by R49 and N118, and indirectly to K76 via a water molecule, as in the equivalent WT complexes. The plasticity in the sidechain rotamers of R49 allow it to maintain an optimal interaction with the distal phosphate. Thus, the phosphodianion binding energy is fully utilised but is not efficiently coupled to optimal TSA conformation.

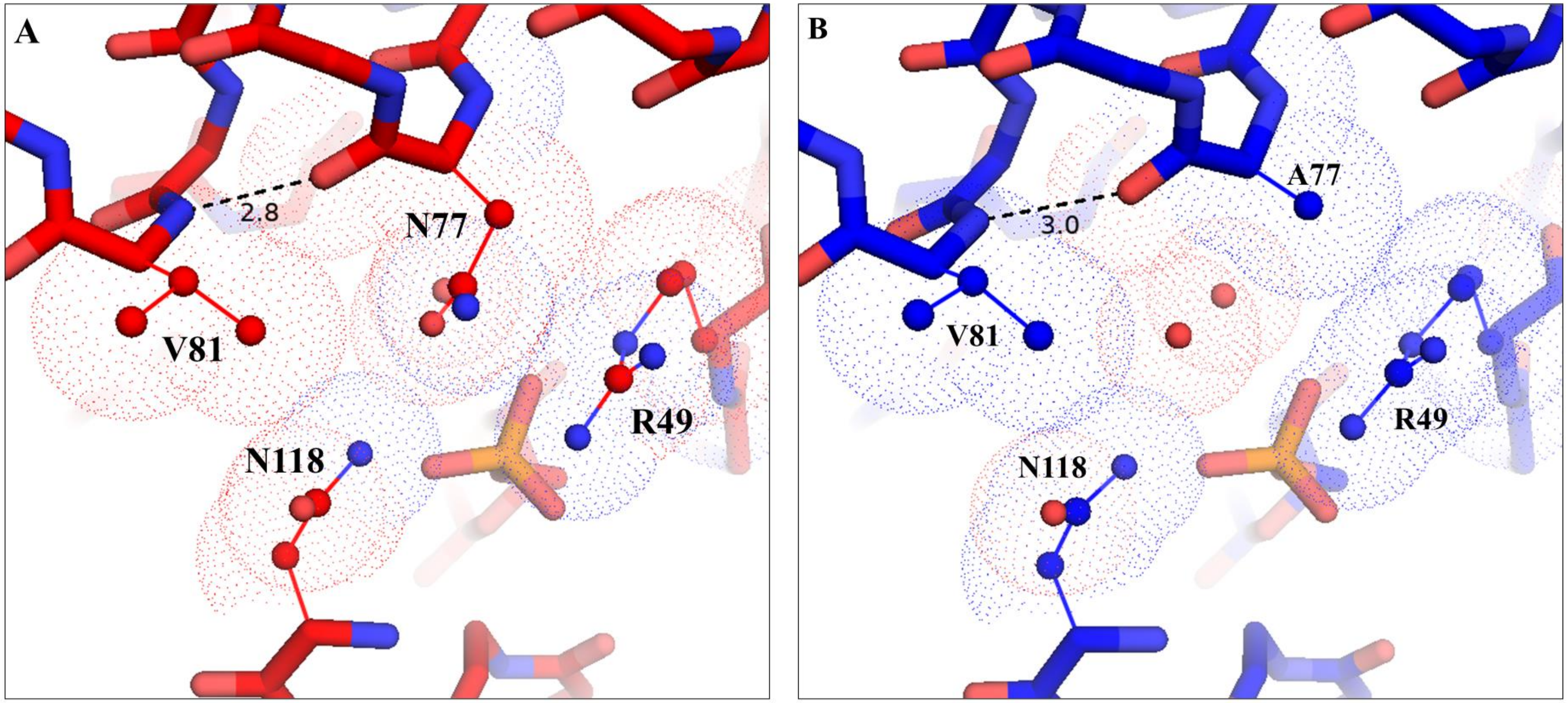


Figure 3-12 – Conformation of mutation site within βPGM$_{N77A}$:MgF$_3$:G6P TSA complex. Crystal structures comparing the local structure of the mutation site in (A) the βPGM$_{WT}$:MgF$_3$:G6P TSA complex (B) βPGM$_{N77A}$:MgF$_3$:G6P TSA complex. The side chain carboxamide of N77 is replaced by a water molecule in the mutant complex. Thus, A77 is free from the restraints imposed by the hydrogen bond with N118 and the hydrogen bond between A77-O and V81-N increases by 0.2 Å. Coordination of the distal phosphate by N118 and R49 is unaffected by the mutation.

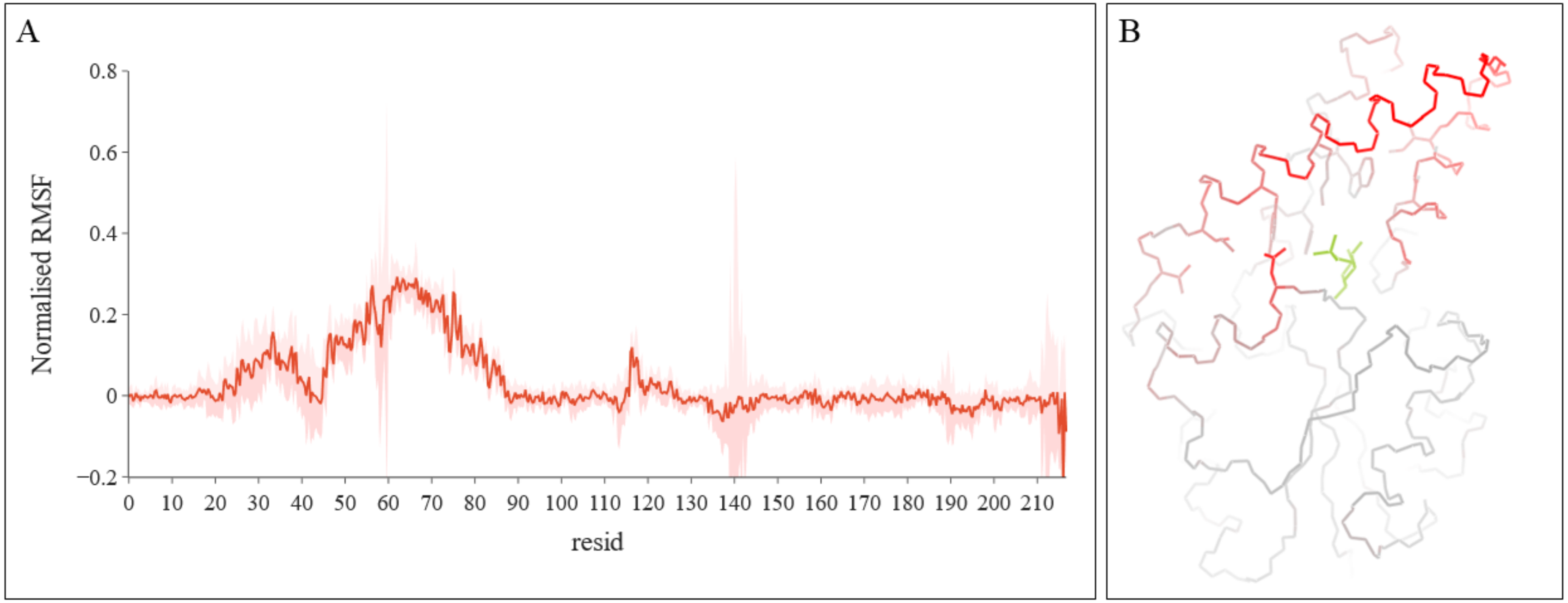


Figure 3-13 – RMSF analysis of the βPGM$_{N77A}$:MgF$_3$:G6P TSA complex. (A) The per-residue RMSF of the βPGM$_{N77A}$:MgF$_3$:G6P TSA complex normalised against the baseline RMSF of the canonical TSA structures. The uncertainty (red area) was calculated using the standard deviation of the pairwise backbone-RMSD calculations. (B) βPGM$_{N77A}$:MgF$_3$G6P structure coloured according to the normalised RMSF, illustrating the subtle reorientation of the cap domain including the side chains of residues V87, I84, Y80, K76, H20, W24 and I56.

3.3 Discussion

3.3.1 Energy landscape of βPGM$_{N77A}$

βPGM is amenable to a variety of experimental methods which allow the consequences of removal of the N77 sidechain carboxamide to be evaluated at multiple stages along the reaction coordinate (**Table 3-2**).

Parameter	**Fold-change (X_{WT} / X_{N77A})**	**ΔG (kcal / mol)**
K_m (βG1P)	13	1.5
K_D (αGal1P)	11	1.4
K_m (βG16BP)	8	1.2
K_D (G6P:BeF_3)	8	1.2
k_{cat}	< 2	0.3
K_D (G6P:MgF_3)	30	2.0

Table 3-2 – Free energy changes for different βPGM$_{N77A}$ and βPGM$_{WT}$ complexes. The fold change and free energy perturbation is calculated for each parameter derived from NMR titrations or steady state, Michaelis Menten kinetics. The free energy is defined by RTln(X_{WT}/X_{N77A}) where R is the gas constant (0.001987 kcal.K^{-1}.mol^{-1}) T is the temperature (298 K) and X_{WT}/X_{N77A} is the fold-change for a given experimental parameter for βPGM$_{WT}$ and βPGM$_{N77A}$.

The contribution to βG16BP binding (corresponding to the step-2 Michaelis complex, with 6'phosphate occupying the distal site) can be estimated from the Michaelis constant for βG16BP measured for βPGM$_{WT}$ and βPGM$_{N77A}$ ($\Delta\Delta G_{\beta G16BP}$ = 1.2 kcal/mol). This is in good agreement with the contribution made to the βPGM:BeF_3:G6P product complex, which has an equivalent sugar orientation ($\Delta\Delta G_{BeF3:G6P}$ = 1.2 kcal/mol). Given the close agreement between N77's contribution to the stability of the Michaelis complex and to the βPGM:BeF_3:G6P complex, it is likely that the Michaelis complex is dominated by NAC I. From here, the enzyme undergoes brief dynamical excursions to the high energy transition state conformation via a rearrangement of the hinge and a 17 ° rotation of the cap domain [139].

I84 and S88 chemical shifts serve as direct reporters of domain orientation in βPGM. The chemical shifts (extrapolated to 100% saturation) show that the βPGM$_{N77A}$:BeF_3:G6P complex exists as a 1:4 mixture of open and closed conformers. Thus, the 1.2 kcal destabilisation of NAC I was sufficient to introduce a measurable population of an open conformation. Given an 8-fold stabilisation measured using the apparent K_D, the ratio of open and closed conformers in WT can, as a first approximation, be estimated at 32:1 (~97%). This shows that the enzyme stabilises the closed, ground state complex sufficiently, but not excessively. Any additional stabilisation could inhibit product release. The closed conformer still dominates the βPGM$_{N77A}$ Michaelis complex and domain closure itself does not make a significant contribution to the activation energy barrier – as it does in the R49A and R49K distal site mutants [139] – which is consistent with a small change in k_{cat} (0.3 kcal/mol) and a small additional stabilisation of the βPGM:MgF_3:G6P TSA complex (2.0 kcal/mol) over the βPGM:BeF_3:G6P NAC I complex (1.2 kcal/mol).

The contribution made by N77 sidechain carboxamide to βG1P binding (corresponding to the step-1 Michaelis complex, with 1'-phosphate occupying the distal site) can be estimated from the change in the Michaelis constant measured for βG1P between βPGM$_{WT}$ and βPGM$_{N77A}$ ($\Delta\Delta G_{\beta G1P}$ = 1.5 kcal/mol). αGal1P binds with the same sugar orientation and shows a similar sensitivity to the N77A mutation ($\Delta\Delta G_{\alpha Gal1P}$ = 1.4 kcal/mol). N77A, thus appears to confer additional stabilisation to the step 1 complexes, where the 1'-phosphate is bound in the distal site and the 6'phosphate is bound in the proximal site. In this configuration, the pyranose ring makes direct contacts with cap domain residues (L44, S52, W24 and K76), whereas water-mediated contacts with the cap domain are made in the step-2 complex [65]. The plasticity in the water-mediated, cap domain interactions with the substrate must reduce the sensitivity of the step-2 complexes to the N77A mutation.

3.3.2 The coupling mechanism

Several pieces of evidence point to a relationship between the V81-N77 backbone hydrogen bond, and the N77-N118 carboxamide pair. Analysis of the βPGM crystal structures has shown a concerted relationship between domain closure; side chain hydrogen bonding between N77 and N118; and shortening of the V81-N77 hydrogen bond. This is corroborated by the analysis of the library of $^{1}H^{15}N$ backbone chemical shifts. Furthermore, truncation of N77 sidechain results in a long hydrogen bond between V81 and N77 within the fully closed TSA complexes. Within the canonical TSA complexes, the sidechain torsions (χ_1 and χ_2) place N77-O_δ within 3Å of N118-N_δ. No further optimisation can be obtained via these torsions. Thus, the only remaining degree of freedom which allows the N77-N118 hydrogen bond to form, is via compression of the V81-N77 hydrogen bond. Thus, the coupling between these two hydrogen bonds is strictly maintained by placing the interacting carboxamides just out of reach. The corollary of this is that an N77Q mutant would also decouple these two hydrogen bonds owing to the plasticity inferred by an additional degree of freedom in the sidechain. The N77 and N118 relationship is the only direct hydrogen bond between the two domains and is well conserved in the vast majority of βPGM sequences (**Figure 7-9**, **Section 5.2.14**). The N77/N118 pair represents 95.4 % of the analysed sequences with a small number of N77-S118 pairs, which would likely perform an analogous role. A Q77/N118 pairing is present in less than 0.1% of the sequences.

Such a mutation might normally be tolerated as ASN and GLN have similar physiochemical properties. However, geometry/distance restraints appear to be a functionally important aspect of this interaction.

The analysis of the chemical shifts suggests a concerted shift in hydrogen bonding through the 70s helix, whereby V81 and N77 make longer hydrogen bonds in the closed complex, while N79, M83 and I84 make shorter hydrogen bonds. Thus, there is antagonism within the helix, which is biased by the N77-N118 carboxamide pair. The crystal structures are not fully representative of this model, however, antagonistic behaviour between the V81/N77 hydrogen bond and the I84/Y80 hydrogen bond is present. The Y80/V81 peptide connects N77 to I84. If N77 makes a short hydrogen bond with V81, Y80-O pulls away from I84-N and no longer influences the backbone torsions of I84.

The analysis of the βPGM$_{N77A}$ TSA crystal structures suggest the involvement of an additional mechanism. The kink in the helix at N77 causes a semi-rigid reorientation of the cap domain, as indicated from the Euler Angles and the RMSF calculations. This rigid body effect transmits through the internal hydrophobics of the cap domain (H20, W24, Y80, K76, I56) to the sidechain of I84. This mechanism is consistent with the additional sensitivity of the step-1 complexes to the N77A mutation, owing to the direct interactions between the pyranose sugar and cap domain via W24 and K76. The plasticity in the water-mediated interactions found in the step-2 complexes may dampen the transmission of this effect through the cap domain.

N77 makes a planar stacking interaction with R49. Thus, R49 sidechain guanidino group might serve as a buttress against N77 which works synergistically with the carboxamide hydrogen bonding interaction to stabilise the conformation of the 70s helix. However, the N77-N118 hydrogen bond is unperturbed by the R49A mutation (PDB: 6HDM). We therefore suggest an alternative mechanism, in which the role of R49 is to stabilise the conformation of N118 sidechain via the distal phosphate, which is required to maintain coupling with the hinge via N77 and V81. Whilst N118 appears to be preorganised to bind the distal phosphate within the open structures, its position is stabilised by contacts with a neighbouring molecule in the crystal lattice.

Coupling via N77 can only account for part of the total stabilisation conferred by R49 to the closed complex. The minimal requirement for domain closure is binding of a sugar phosphate with the phosphate bound in the distal site. Proximal site occupancy is not required for full domain closure – evidenced by the substrate-inhibited βPGM$_{D170N}$:βG1P complex and the βPGM$_{WT}$:αGal1P complex [60,139]. Therefore, the sugar likely plays a role in coupling R49's interaction with domain closure. The pyranose ring of the sugar-phosphate packs against H20 from the cap domain and makes hydrogen bonds with V47-O and G46-N. Hydrogen bonds with

W24, K76, L44 and S52, are made directly in the step-1 complexes and indirectly, via water molecules, in the step-2 complexes [65]. A single hydrogen bond to the core domain is made via D10 sidechain. It can be seen from **Figure 3-2** that the orientation of the distal phosphate is influenced by the tethered sugar, thus R49 may play a role in stabilising the orientation of the sugar such that it can make contacts with the cap domain which position 1’OH to donate a hydrogen bond either to D10, or in the case of the NAC I complex, to the BeF_3 moiety in the proximal site.

3.4 Conclusions

The phosphodianion-driven enzyme activation framework states that the intrinsic phosphodianion binding energy is coupled to domain closure. NMR chemical shifts and x-ray crystal structures implicate the N77-N118 carboxamide pair in this coupling mechanism. Free energy contributions of this interaction have been measured kinetically; using a βPGM:BeF_3:G6P GSA complex; using a βPGM:MgF_3:G6P TSA complex; and a βPGM:αGal1P complex. The structural basis of this effect was inferred from crystal structures of the WT and N77A TSA complexes. The N77-N118 carboxamide pair is tightly coupled to the hydrogen bonding in the 70s helix. Stabilisation of the V81-N77 backbone amide hydrogen bond releases I84 from its hydrogen bond with Y80 allowing torsional changes which give rise to hinge bending. These structures also provide evidence of a second pathway, transmitted as a rigid body effect, through the internal hydrophobics of the cap domain. R49 thus confers additional stability by maintaining a competent distal site which organises the sidechain of N118 to accept a hydrogen bond from N77. To our knowledge, no previous studies have provided such a detailed characterisation of hinge bending and domain closure, which is made possible by the collation and curation of a large library of NMR data and x-ray crystal structures.

4 Future perspectives

The role of Y19 in facilitating product release could be investigated by characterising a Y19A mutant using steady state kinetics and NMR titrations, similar to those described in section 3.2. The mutation could also be made *in silico* and monitored by MD simulation to observe the effect on domain orientation. Residual dipolar couplings could be measured to provide additional information on domain orientation for validation and refinement of MD simulation data. RDCs are more sensitive to subtle changes in domain orientation which may not give a significant change in the radius of gyration.

Prediction of protein structure from the primary sequence has recently become possible using deep learning based on multiple sequence alignments (e.g. AlphaFold2 [159], RosettaFold2 [160]) or protein language models (ESMFold [161]). Several structure predictions using the pgmB sequence from *Lactococcus lactis* (Uniprot codes: P71447, A0A2A9IDT2, A0A552Z207, A0A552XK19, A0A2X0R470) can be found in the AlphaFold Protein Structure Database. Each of the predictions align closely with the fully closed, transition state analogue (PDB 2WF5) rather than the ligand-free open conformer (PDB 2WHE). The problem of inter-domain accuracy of structure predictions has been reported for other systems [162–164]. This issue is acknowledged by AlphaFold, which reports a pairwise alignment error (PAE) matrix quantifying the uncertainty in the predicted pairwise interatomic distances. For βPGM structure predictions, there is a lower uncertainty for distances within each domain and a higher uncertainty for distances between each domain. The deep neural network used by AlphaFold2 is trained on x-ray crystal structures. Our experimentally validated MD simulations of βPGM, as well as previous studies of T4 lysozyme and phosphoglycerate kinase suggest that, in general, crystal structures may favour more compact interdomain relationships. This bias may be learned by deep neural networks. A larger survey of experimental and predicted protein structures would be useful in determining the severity and generality of this issue.

Some simulations implemented a metadynamics bias to accelerate sampling of the K145-P146 peptide bond. Whilst the sampling of this torsion angle was successfully increased, the dynamics of residues in its vicinity were substantially increased and did not converge. Recent MD studies of carbohydrate polymers demonstrated high-quality conformational sampling using replica exchange methods (REST2-RECT) [165]. It is feasible that this method would be equally successful when applied to the conformational dynamics of short peptides or short loops within larger proteins. A thoroughly sampled dynamic ensemble of the P146 loop may provide a valuable insight into the mechanism behind allomorphic regulation in βPGM (Section 1.2.8).

A full assignment of βPGM_{N77A} complexes using $^{2}H^{13}C^{15}N$-labelled material could be useful to validate the transfer assignment and to find missing assignments. At present, the structural consequences of the N77A mutation were only structurally accessible via the TSA complexes. A D10N N77A double mutant could be made and characterised by NMR and x-ray crystallography – in particular, the high affinity βG16BP complex would likely be accessible by crystallography where the βPGM_{N77A}:BeF_3:G6P was not. These structures may provide further insights into the compromised stability of the βPGM_{N77A} closed, ground state complexes.

The N77A has a modest effect on the stability of the closed ground state complexes, in comparison to the R49A and R49K mutants. Consequently, the closed ground state complexes are compromised but are still amenable to NMR and crystallography. This offers a unique opportunity to study the relationship between domain closure and other dynamical process within the enzyme including the isomerisation of P146 and recruitment of K145. It may be informative to expand the analysis of the βPGM_{N77A}:BeF_3:G6P titration trajectories to other residues in the protein.

The evidence of a direct, through-domain communication pathway between the substrate binding site and the hinge in βPGM may be considered a specific implementation of a general mechanism for substrate/ligand-induced domain reorientation in proteins. Structural tightening/coalescence of protein domains in response to ligand binding may propagate towards interdomain hinge regions in multidomain proteins, which influences the interdomain relationship. The involvement of this mechanism might be relevant to many multi-domain proteins, which can be tested through a combined use of NMR and site-directed mutagenesis.

5 Materials & Methods

All materials reagents were purchased from Sigma-Aldrich, Roche, Alfa Aesar, VWD, Fisher Scientific, unless otherwise stated.

βG1P was enzymatically synthesised from maltose using maltose phosphorylase (EC 2.4.1.8). 7.2 g maltose was dissolved in 0.5 M phosphate buffer and adjusted to pH 7. 40 Units of maltose phosphorylase was added to the maltose solution. The mixture was incubated at 18 °C overnight. Maltose phosphorylase was removed from the sample using a centrifugal filtration unit (Sartorius, Vivaspin 5 kDa cut-off). 500 mg $CaCl_2$ was added resulting in a thick white precipitate ($CaPO_4$) formed which was removed by centrifugation at 4500 RPM for 20 mins (Thermo Scientific Heraeus Labofuge 400 R). The concentration of βG1P was verified by comparison with a known concentration of G6P using ^1H NMR spectroscopy.

βG16BP was enzymatically synthesised [166]. A reaction mixture containing 200 mM K^+ HEPES + 20 µM βPGM_{D170N} + 100 mM $MgCl_2$ + 20 mM βG1P + 40 mM AcP + 2 mM NaN_3, was incubated for 265 min, after which the reaction was quenched by heat denaturation, at 90 °C for 10 min. Denatured protein was removed by centrifugation at 4,500 RPM (Thermo Scientific Heraeus Labofuge 400 R). The solution was separated on a IR120 (H^+) ion-exchange resin, pre-washed in milliQ-H_2O. The resulting solution was neutralised with 0.2 M Barium Hydroxide on ice, resulting in selective phosphorylation of βG16BP. βG1P and βG16BP were retained in solution. The precipitate was extracted by centrifugation at 4,500 RPM (Thermo Scientific Heraeus Labofuge 400 R). The precipitate was diluted in 1 L cold milliQ-H_2O and passed through a R120 (Na +) ion-exchange resin.

5.1 Buffers Solutions & Media

5.1.1 Luria Bertani (LB) Agar

- 10 g/L Tryptone
- 5 g/L Yeast Extract
- 10 g/L NaCl
- 15 g/L Bacto-Agar

5.1.2 LB Liquid Media

- 10 g/L Tryptone
- 5 g/L yeast extract
- 10 g/L NaCl

5.1.3 M9 Minimal media

- 6 g/L Na_2HPO_4
- 3 g/L KH_2PO_4
- 0.5 g/L NaCl

The pH was adjusted before autoclave sterilization.

The following sterile reagents were added to 1L media…

- 1 mL 1 M $MgSO_4$
- 100 uL 1 M $CaCl_2$
- 650 uL Trace Elements
- 100 uL 10 mg/mL Thiamine
- 2 mL 50 % (w/v) $(^{15}NH_4)SO_4$
- x mL Glucose

5.1.4 SOB media

- 20 g/L tryptone
- 5 g/L yeast extract
- 0.5 g/L NaCl

5.1.5 SOC media

Add the following to 50 mL SOB

- 500 uL 1 M MgCl2
- 500 uL 1 M MgSO4
- 2 mL 20% (w/v) glucose, filter-sterilised

5.1.6 Trace elements

- 550 mg CaCl2.2H2O
- 220 mg ZnSO4.7H2O
- 140 mg MnSO4.H2O
- 45 mg CoCl2.6H2O
- 40 mg CuSO4.5H2O
- 40 mg H3BO3
- 26 mg Na2MoO4.2H2O
- 26 mg KI

The pH was adjusted with acetic acid to pH 8.0 before adding:

- 500 mg EDTA

The pH was re-adjusted with acetic acid to pH 8.0 before adding:

- 375 mg FeSO4.7H2O

The solution was then made up to 100 ml and sterilised by autoclave.

5.1.7 TFB1 buffer

- 100 mM RbCl
- 50 mM $MnCl_2$
- 30 mM Kac
- 10 mM $CaCl_2$
- 15% glycerol

The pH was adjusted to 5.8 with NaOH before the solution was filter-sterilised.

5.1.8 TFB2 buffer

- 10 mM MOPS
- 10 mM $RbCl_2$
- 75 mM $CaCl_2$
- 15 % Glycerol

The pH was adjusted to 5.8 with NaOH before the solution was filter-sterilised.

5.1.9 Standard HEPES Buffer

- 50 mM HEPES
- 5 mM MgCl2
- 2 mM NaN3

The pH was adjusted to 7.2 before the solution was filter-sterilised.

5.1.10 βPGM K-HEPES High Salt Buffer

- 50 mM HEPES
- 100 mM MgCl2
- 2 mM NaN3

The pH was adjusted to 7.2 before the solution was filter-sterilised.

5.1.11 Chromatography Buffer A

- 50 mM K+ HEPES pH 7.2
- 2 mM NaN3
- 1 mM EDTA

5.1.12 Chromatography Buffer B

- 50 mM K+ HEPES pH 7.2
- 2 mM NaN3
- 1 mM EDTA
- 1 M NaCl

5.1.13 Lysis Buffer

- 50 mM K+ HEPES pH 7.2
- 2 mM NaN3
- 1 mM EDTA
- Protease Inhibitor cocktail

5.1.14 4X SDS-PAGE Stacking Gel Buffer

- 0.5 M Tris
- 0.4% (w/v) SDS

The pH was adjusted to 6.8 with HCl before the solution was filter-sterilised.

5.1.15 4X SDS-PAGE Resolving Gel Buffer

- 1.5 M Tris
- 0.4% (w/v) SDS

The pH was adjusted to 8.8 with HCl before the solution was filter-sterilised.

5.1.16 SDS-PAGE Stacking Gel (4.5% BisAcrylamide)

- 2.5 ml 4X SDS-PAGE Stacking Gel Buffer
- 1.125 ml 40% (w/v) BisAcrylamide (37.5:1)

This solution was then diluted to 10 ml before adding:

- 110 µl 10% (w/v) Ammonium Persulphate (APS) (filter-sterilised)
- 11 µl Tetramethylethylenediamine (TEMED)

The solution was thoroughly mixed before pouring into the gel apparatus to set.

5.1.17 SDS-PAGE Resolving Gel (18% BisAcrylamide)

- 2.5 ml 4X SDS-PAGE Resolving Gel Buffer
- 4.5 ml 40% (w/v) BisAcrylamide (37.5:1)

This solution was then diluted to 10 ml with milliQ-H_2O before adding:

- 100 µl 10% (w/v) Ammonium Persulphate (APS) (filter-sterilised)
- 10 µl Tetramethylethylenediamine (TEMED)

The solution was thoroughly mixed before pouring into the gel apparatus to set.

5.1.18 SDS-PAGE Running Buffer

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

The pH was adjusted to 8.3 with HCl before the solution was filter-sterilised.

5.1.19 2X SDS-PAGE Loading Buffer

- 100 mM Tris
- 200 mM DTT
- 4% (w/v) SDS
- 0.2% (w/v) Bromophenol Blue
- 20% (v/v) Glycerol

The pH was adjusted to 6.8 with HCl before the solution was filter-sterilised.

5.1.20 Gel stain and destain

- 10% (v/v) Acetic Acid
- 45% (v/v) Methanol
- 45% (v/v) Milli-Q Water
- 0.25% (w/v) Coomassie Brilliant Blue R250 (stain only)

5.2 Methods

5.2.1 Crystal structure processing

51 *β*PGM structures from *lactococcus lactis* were downloaded from the protein data bank. Each of the chains were extracted using pdb-tools software [167] as well as fixing formatting errors and removing duplicate atoms. OpenMM's **`pdbfixer`** python module [105] was used to remove non-protein atoms and replace non-standard residues (e.g. aspartyl phosphate, 5'-fluoro-tryptophan). The final collection of 64 βPGM crystal structures are summarised in **Table 5-1**.

1o08_A	tsa	6hdj_A	tsa	6h8v_B	open	6hdi_B	open
6hdg_A	tsa	6hdk_A	tsa	6h8v_A	open	6hdh_B	open
4c4s_A	tsa	5ok2_A	tsa	6h8u_A	open	6h91_B	open
1o03_A	tsa	1z4o_A	tsa	6h8w_A	open	6qzg_B	open
4c4r_A	tsa	5oly_A	tsa	6h90_A	open	6h93_B	open
2wf5_A	tsa	5oly_G	tsa	5ojz_A	open	6h8x_B	open
6i03_A	tsa	2wf9_A	nac	5olw_A	open	6h8y_A	open
2wf8_A	tsa	1z4n_B	nac	5olw_B	open	6h93_A	open
2wf7_A	tsa	1z4o_B	nac	1zol_A	open	6ydm_B	open
2wf6_A	tsa	5o6p_A	nac	6ydk_A	open	6hdi_A	open
4c4t_A	tsa	1z4n_A	nac	3fm9_A	open	6hdh_A	open
5olx_A	tsa	5o6r_A	nac	2whe_A	open	6h8x_A	open
3zi4_A	tsa	5ok1_A	nac	2wfa_A	open	6qzg_A	open
6hdm_A	tsa	5ok0_A	nac	6h94_A	open	6h91_A	open
6hdl_A	tsa	6ydl_A	open	6h8z_A	open	6hdf_B	open
6ydj_A	tsa	1lvh_A	open	6ydm_A	open	6hdf_A	open

Table 5-1 – The βPGM PDB dataset. The Protein Data Bank ID code and the chain identifier for each structure. The domain orientation categorised by hierarchical clustering is also stated for each structure.

5.2.2 Crystal contact analysis

Pymol's `symexp` function was used to generate a PDB file containing the neighbouring asymmetric units. Crystal contact maps were generated using python scripts. A distance matrix was calculated between the coordinates from the central copy and the surrounding copies of the asymmetric unit. The number of contacts within 3.5 Å were counted cumulatively across all of the open βPGM crystal structures.

5.2.3 Molecular Dynamics acquisition

All calculations were performed using openMM [105], using either Amber's ff14sb paired with the TIP3P water model [111] or ff15ipq and the SPC/E_b water model [129]. Some simulations included a further electronic continuum correction (ECC), implemented as 75% scaled charges for monoatomic ions ($Mg^{1.5+}$, $Na^{0.75+}$, $Cl^{0.75-}$) with modified LJ parameters [168] summarised in **Table 5-2**. Nonbonded interactions were truncated using a switching function.

$$y = 1 - 6x^5 + 15x^4 - 10x^3 \quad x = \frac{r - r_{sw}}{r_{co} - r_{sw}}$$

$$r_{sw} = 0.75\text{nm} \quad r_{co} = 0.9\text{nm}$$

Periodic boundary conditions were applied using a cubic unit cell. Electrostatics were calculated using the Particle Mesh Ewald (PME) method (5th order B-splins, error-tolerance = 0.0005). The SETTLE algorithm was used to constrain bond lengths and angles in water molecules. The SHAKE algorithm was used to constrain protein bonds containing hydrogen atoms. 4 g/mol hydrogen mass repartitioning was applied. Langevin integration was used with 4fs timestep at 298 K and 0.1 ps^{-1} thermal coupling. Starting coordinates were derived from the crystal structure of either the βPGM$_{WT}$:MgF_3:G6P complex (PDB: 2WF5) or the βPGM$_{WT}$ substrate-free complex (PDB:2WHE). Hydrogen atoms were added using Amber's LEaP module (pH 7.2). The starting structure was placed in the centre of a pre-equilibrated water box with a minimum distance of 2 nm from the periodic boundary. Overlapping water molecules were removed. Na^+ and Cl^- ions were added to neutralize the system. The final ionic strength was 160 mM. The L-BFGS algorithm was used for local minimization of the potential energy. The solvent was equilibrated under NPT conditions followed by NVT equilibration of the whole system. A further 5μs data were collected. Coordinates were saved at 100 ps intervals for analysis. For some simulations, a metadynamics bias was applied (implemented in openMM) to the ζ-angle of K145 (defined by atoms: K145-C_α, K145-O, P146-C_δ and P146-C_α) and to the ψ angle of P146 (defined by atoms: P146-N, P146-C_α, P146-C and A147-N). A bias factor of 15, the height was set to 0.8 kJ/mol and the deposition frequency was 500 ns^{-1}. The openMM metadynamics module was modified to report the collective variables (ζ, ψ) for each deposited periodic gaussian kernel.

Ion	Full Charge Force Field			ECC Force Field				
	q (e)	σ (Å)	ε (kJ/mol)	q (e)	σ (Å)	ε (kJ/mol)	$\frac{r_{min}}{2}$ (Å)	ε (kcal/mol)
Na	+1	2.3500	0.5439	+0.75	2.1150	0.5443	1.187004	0.130091
Mg	+2	1.8900	3.6610	+1.5	1.3600	3.6610	0.763274	0.875001
Cl	–1	4.4000	0.4184	–0.75	4.1000	0.4928	2.301047	0.117782

Table 5-2 – **Electronic Continuum Correction Parameters for monoatomic ions.** The Lennard-Jones parameters for a full-charge and charge-scaled forcefield for each of three monoatomic ions (Mg^{2+}, Na^{+} and Cl^{-}) are stated. For clarity and ease of comparison with literature values, the energy (ϵ) is expressed in kcal/mol and kJ/mol and the atomic radius is expressed as either σ or $R_{min}/2$.

$$u = E\left(\left(\frac{r_{min}}{r}\right)^{12} - 2\left(\frac{r_{min}}{r}\right)^{6}\right)$$

$$u = 4E\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right)$$

$$r_{min} = 2^{\frac{1}{6}} \cdot \sigma$$

5.2.4 MD Analysis

Euler Angles

Euler Angles are calculated for a structure X compared to a reference structure (PDB: 2WF5). The reference is aligned to the principal axes of its cap domain c_α-atoms. The core domain of structure X is aligned to the core domain of the reference structure. The cap domain of structure X is then aligned to the cap domain of the reference structure. The rotation matrix associated with the latter alignment is used to calculate intrinsic Euler angles, using the XYZ convention (aka. nautical angles). Alignments were implemented using **MDAnalysis** [169,170], calculation of Euler angles from the rotation matrix was done using **Scipy**, organisation and storage of data tables was done using **Numpy** and **Pandas**; data visualisation was done using **Plotly**.

```
import numpy as np
import pandas as pd
import MDAnalysis as mda
from MDAnalysis.analysis import align
from scipy.spatial.transform import Rotation as R

def calculateEulerAngles(ref, topl, traj, outfile):

    # read reference pdb file
    ref = mda.Universe(ref)

    # define the cap domain CA atoms
    refcap = ref.select_atoms("(resid 15-87) and (name CA)")

    # translate/rotate cap domain onto its principal axes
    ref.atoms.translate(-refcap.center_of_mass())
    ref.atoms.rotate(refcap.principal_axes())

    # create mda trajectory object for input file
    sim = mda.Universe(topl, traj)

    # setup trajectory alignemnt
    print("calculating core domain alignment...")
    coreAlign = align.AlignTraj(
        sim,
        ref,
        select="(resid 1:14 88:218) and (name CA)",
        filename="./tempCoreAligned.dcd",
        match_atoms=True,
        verbose=True,
    )

    # run alignment
    coreAlign.run()

    # read aligned trajectory
    sim = mda.Universe(topl, "tempCoreAligned.dcd")

    # select cap domain atoms
    tscap = sim.select_atoms("(resid 15-87) and (name CA)")

    # initialise a data list
    data = []

    # iterate through each frame in the trajectory
    print("calculating cap alignment & euler angles...")
    for ts in tqdm(sim.trajectory):

        # calculate the rotation matrix
        rot, rmsd = align.rotation_matrix(
            tscap.atoms.positions,
            refcap.atoms.positions
        )

        # calculate euler angles from the rotation matrix
        yaw, pitch, roll = R.from_matrix(rot.T).as_euler('XYZ', degrees=True)
        # calculalte the axis-angle from the rotation matrix
        theta = np.degrees(np.arccos((np.trace(rot) - 1) / 2))

        # add data to the data list
        data.append([yaw, -pitch, roll, theta])
```

```
    # create a dataframe
    df = pd.DataFrame(data)

    # define the column names
    df.columns = ["yaw", "pitch", "roll", "theta"]

    # define the frame id
    df["structure"] = "frame" + df.index.astype(str)

    # write the results to a file in the simulation directory
    with open(outfile, "w") as f:

        table_string = df.to_string(index=False, float_format="%.7f", col_space=15)

        f.write(table_string)
```

Dihedral Angles

An `AtomGroup` is defined with four atoms in order which define the torsion to be calculated. For example, the `phi_selection` method selects atoms C_i, N_i, $C\alpha_{i+1}$ and C_{i+1} for any given residue. The `Dihedral` function calculates the torsions for each of the given `AtomGroups`.

```
# atom selection functions
def phi(residue): return residue.phi_selection()
def psi(residue): return residue.psi_selection()
def omega(residue): return residue.omega_selection()

def get_dihedrals(residues, dihedral, outfile):

    # initialise empty lists
    atom_groups = []
    residue_identifiers = []

    # loop through each residue
    for residue in residues:

        # if the dihedral angle is applicable to that residue...
        if dihedral(residue) is not None:

            # ... append residue identifier to list
            residue_identifiers.append(f'{residue.resname}-{residue.resid}')

            # ... append atom group to list
            atom_groups.append(dihedral(residue))

    # calculate the dihedral angles
    dihedrals = Dihedral(atom_groups, verbose=True).run()
    angles = dihedrals.results.angles

    # convert to a dataframe
    table = pd.DataFrame(angles)
    table.columns = residue_identifiers

    # write dataframe to a text file
    f = open(outfile, "w")
    f.write(table.to_string(index=None, float_format="%.4f"))
```

Interatomic distances

```
# create MDAnalysis simulation object
u = mda.Universe(topologyfile, trajectoryfile)

# select relevent atoms
a1 = u.select_atoms(atomSelectionString1)
a2 = u.select_atoms(atomSelectionString2)

# initialise an empty list
distances = []

# loop through each frame in the trajectory
for ts in tqdm(u.trajectory):

    # calculate the distance between selected atoms
    d = np.linalg.norm(np.squeeze(a1.positions) - np.squeeze(a2.positions))

    # append distance to list
    distances.append(d)

# create a pandas dataframe
df = pd.DataFrame(distances)
df.columns = ["distance"]

# write dataframe to file
with open(args.outfile, "w") as f:
    f.write(df.to_string(index=None))
```

Radius of gyration

```
# create MDAnalysis simulation object
u = mda.Universe(args.topl, args.traj)

# select relevent atoms
s = u.select_atoms(args.sele)

# initialise an empty list
RgData = []

# loop through each frame in the trajectory
for ts in tqdm(u.trajectory):

    # calculate radius of gyration and store in dictionary
    RgData.append({"frame": ts.frame, "rg": s.radius_of_gyration()})

# create pandas Dataframe
df = pd.DataFrame(RgData)

# write dataframe to file
print(f'writing data to file: {args.outfile}')
with open(args.outfile, "w") as f:

    contents = df.to_string(index=None)
    f.write(contents)
    f.write("")
```

Metadynamics Bias Kernel Density Estimation

Reconstruction of the metadynamics bias over the course of the MD simulation was calculated using a Kernal Density Estimation procedure with a periodic gaussian kernel. This is the same kernel used by the `OpenMM` Metadynamics module, rather than the Von Mises kernel implemented in `PLUMED`. The `gaussian_periodic_1Dkernel` function translates the input x-values to x_0 and creates a hard-coded periodic boundary. The modified x-variable is then input into a standard gaussian kernel function. The `gaussian_periodic_2Dkernel`, simply takes the outer (pairwise) product of two 1D kernels. The `gaussian_periodic_kde` calculates the cumulative sum of periodic gaussian kernels centred on a series of input dihedral angles, with specified width.

```
def gaussian_periodic_1Dkernel(x0, x, sigma):
    """ define a periodic gaussian kernel """

    # calculate circular x-variable
    x_shift = np.abs(x - x0)
    x_wrapped = np.min(np.array([x_shift, np.abs(x_shift - 2 * np.pi)]), axis=0)
    x_wrapped[-1] = x_wrapped[0]

    # calculate gaussian kernel using circular x-variable
    kernel = np.exp(-(x_wrapped ** 2) / (2 * (sigma ** 2)))

    return kernel


def gaussian_periodic_2Dkernel(x0, x, y0, y, sigx, sigy, height):
    """ calculate 2D periodic gaussian kernel """

    kernel_x = gaussian_periodic_1Dkernel(x0, x, sigx)
    kernel_y = gaussian_periodic_1Dkernel(y0, y, sigy)
    kernel_xy = height * np.multiply.outer(kernel_x, kernel_y)

    return kernel_xy


def gaussian_periodic_kde(data, x, y, sigx, sigy):
    """ calculate kernel density etimation """

    # initialise an array of null values
    kernels = np.zeros((len(x), len(y)))

    # calculate cummulative sum of metadynamic bias
    for row in data:
        x0, y0, height = row
        kernels += gaussian_periodic_2Dkernel(x0, x, y0, y, sigx, sigy, height)

    return kernels
```

Dihedral Categorisation

A series of functions are defined below, which are used convert continuous Ramachandran angle variables into categorical variables which provide a conformation signature which can be used to evaluate conformation sampling and simulation convergence. The hard coded **`CATEGORY_DEFINITIONS`** define the boundaries of different discrete regions of Ramachandran space, selected based on population-density criteria. An atypical periodic boundary is used such that none of the defined categories is split by a periodic boundary.

```
CATEGORY_DEFINITIONS = {
    # Proline
    "PRO": [
        [220, 360, -100, 20],   # category 1
        [220, 360, 20, 108],    # category 2
        [220, 360, 108, 220],   # category 3
    ],

    # Glycine
    "GLY": [
        [15, 125, 80, 250],     # category 1
        [125, 235, 80, 250],    # category 2
        [235, 350, 80, 250],    # category 3
        [15, 180, -110, 80],    # category 4
        [180, 360, -110, 80],   # category 5
    ],

    # All Other Residues
    "AOR": [
        [170, 360, -90, 38],    # category 1
        [255, 310, 38, 95],     # category 2
        [175, 255, 70, 190],    # category 3
        [255, 340, 95, 210],    # catrgory 4
        [20, 190, 180, 250],    # catrgory 5
        [15, 130, -50, 80],     # catrgory 6
    ]
}

def _wrap_ramachandran_angles(phi: np.array, psi: np.array) -> (np.ndarray,
np.ndarray):
    """ translates the periodic boundaries of
    the ramachandran plot from their default
    values...

        phi: [-180, 180], psi: [-180, 180]

    to a new frame where each population is not
    split by a periodic boundray...

        phi: [0, 360], psi: [-110, 250] """

    _phi = phi % 360
    _psi = ((psi + 110) % 360) - 110

    return _phi, _psi
```

```python
def read_and_filter(phi_file: str,
                    psi_file: str,
                    omg_file: str,
                    selection: str) -> pd.DataFrame:
    """ reads a whitespace delimitted phi/psi dihedral angle
    ascii table. Returns a numpy.ndarray for phi and psi respectively
    as well as a list of residues representing the columns in
    each of the phi/psi arrays """

    # read phi/psi tables (whitespace delimited, ascii file)
    phi = pd.read_table(phi_file, delim_whitespace=True)
    psi = pd.read_table(psi_file, delim_whitespace=True)
    omg = pd.read_table(omg_file, delim_whitespace=True)

    # only keep residues for which phi, psi and omega angles are defined
    residues = [
        residue for residue in phi.columns
        if (residue in psi.columns) & (residue in omg.columns)
    ]

    if selection is not None:

        # filter user-selected residues
        user_selected_resids = _parse_selection_string(selection)
        residues = [
            residue for residue in residues
            if int(residue.split("-")[-1]) in user_selected_resids
        ]

    # apply filter
    phi = phi.filter(residues)
    psi = psi.filter(residues)
    omg = omg.filter(residues)

    return phi, psi, omg


def _parse_selection_string(selection: str) -> List[int]:
    """ takes a comma separated sequence of residue ids
    and converts it to a list of integers. Parses "x-y" as
    a range of integers from x to y."""

    resids = []

    for sel in selection.split(","):

        if "-" in sel:

            lower = int(sel.split("-")[0])
            upper = int(sel.split("-")[1])
            resids += [x for x in range(lower, upper + 1, 1)]

        else:

            resids.append(int(sel))

    return resids
```

```
def categorise_peptide_conformers(phi: pd.DataFrame,
                                  psi: pd.DataFrame,
                                  omg: pd.DataFrame) -> (list, np.ndarray):
    """ categorises peptide conformations by organising
    phi/psi dihedral angles into discrete ramachandran
    categories (see global variable: CATEGORY_DEFINITIONS) """

    # check dihedral angles correspond to the same residues (in order)
    assert list(phi.columns) == list(psi.columns)
    assert list(phi.columns) == list(omg.columns)

    # convert column values to a list
    residues = phi.columns

    # convert pandas dataframe to numpy array
    phi = phi.to_numpy()
    psi = psi.to_numpy()
    omg = omg.to_numpy()

    # initialize an array of NaN (not a number) values
    peptide_categories = np.full(phi.shape, np.nan)

    # categorise proline residues
    proline_filter = [
        i for i, x in enumerate(residues)
        if ("PRO" in x)
    ]
    pro_cat = _categories_dihedrals(
        phi[:, proline_filter],
        psi[:, proline_filter],
        omg[:, proline_filter],
        CATEGORY_DEFINITIONS["PRO"]
    )
    peptide_categories[:, proline_filter] = pro_cat

    # categorise glycine residues
    glycine_filter = [
        i for i, x in enumerate(residues)
        if ("GLY" in x)
    ]
    gly_cat = _categories_dihedrals(
        phi[:, glycine_filter],
        psi[:, glycine_filter],
        omg[:, glycine_filter],
        CATEGORY_DEFINITIONS["GLY"]
    )
    peptide_categories[:, glycine_filter] = gly_cat

    # categorise all other residues
    aor_filter = [
        i for i, x in enumerate(residues)
        if ("PRO" not in x) & ("GLY" not in x)
    ]
    aor_cat = _categories_dihedrals(
        phi[:, aor_filter],
        psi[:, aor_filter],
        omg[:, aor_filter],
        CATEGORY_DEFINITIONS["AOR"]
    )
    peptide_categories[:, aor_filter] = aor_cat

    # convert to a pandas dataframe
    peptide_categories = pd.DataFrame(peptide_categories)
    peptide_categories.columns = residues

    return peptide_categories
```

```python
def _categories_dihedrals(phi: np.ndarray,
                          psi: np.ndarray,
                          omg: np.ndarray,
                          ramachandran_categories: dict) -> np.ndarray:
    """ places ramachandran angles into discrete
    categories using the 'CATEGORY_DEFINITIONS'
    global variable. New categories a """

    phi, psi = _wrap_ramachandran_angles(phi, psi)

    # initialize an array of NaN values
    categories = np.full(phi.shape, np.nan)

    for i, boundary in enumerate(ramachandran_categories):

        x1, x2, y1, y2 = boundary

        # identify values which meet the criteria ...
        boolean_filter = (x1 <= phi) & (phi <= x2) & (y1 <= psi) & (psi <= y2)

        # ... assign a category ID to these values
        categories[boolean_filter] = i + 1

    # recategorise cis peptides
    is_cis = (-30 < omg) & (omg < 30)
    categories[is_cis] = categories[is_cis] + len(ramachandran_categories)

    # set 'transitional' angles to NaN
    unclassified = ((-150 < omg) & (omg < -30)) | ((30 < omg) & (omg < 150))
    categories[unclassified] = np.nan

    return categories
```

Dihedral angle tables are parsed, and the data are separated into different residues groups. The Ramachandran distributions for Proline, Glycine and all other residues are treated separately, i.e. they have different category definitions. A data point is defined for each residue in each frame of the simulation. Each data point which falls within the bounds of a specified category are assigned a number. The categories do not cover the whole of Ramachandran space, data points which fall within the transition state space are assigned to an N/A category.

Each frame contains a data point for each residue. The peptide backbone conformation is represented by a series of integers corresponding to a specific Ramachandran category. If two frames have a similar conformation they will have the same backbone conformational code.

If, during the simulation, the structure of the peptide does not change then the code will not change. The conformational coverage over a simulation can be determined by counting the cumulative number of unique conformational codes. If the simulation starts exploring new conformations the unique category count will increase. If the simulation starts to explore conformation space which has already been sampled, then the cumulative unique category count will not increase. For a conformationally convergent simulation, the cumulative unique category count will plateau. This is necessary but not sufficient for demonstrating convergence.

```
def convert_peptide_categories(categories: np.ndarray) -> np.ndarray:
    """ counts the number of unique peptide
    categories at each time step """

    # extract the unique peptide categories
    unique_categories = np.unique(categories, axis=0)

    # initialize an array of NaN values
    category_ids = np.full(categories.shape[0], np.nan)

    # assign a single-digit dihedral category ID to each peptide conformation
    for i, cat in enumerate(unique_categories):

       # select all equivalent rows
        boolean_selection = np.all(categories == cat, axis=1)

        # set value for row
        category_ids[boolean_selection] = i + 1

    return category_ids


def count_unique_categories(categories: np.ndarray) -> np.ndarray:

    # initialize an array of NaN values
    unique_category_count = np.full(categories.shape[0], np.nan)

    # count the number of unique peptide categories at each time step
    for i in tqdm(range(len(unique_category_count))):

        unique_category_count[i] = len(np.unique(categories[:i]))

    return unique_category_count
```

5.2.5 Molecular Biology

UV-Vis spectroscopy

Thermofisher Nanodrop One UV-Vis Spectrophotometer was used for absorbance measurements (λ = 600 nm, cuvette mode) to monitor the growth of bacterial culture, using sterile liquid media (LB or M9) as a blank. Samples were added to plastic cuvettes with a 1 cm path length. Buffer-blanked, absorbance measurements (λ = 280 nm, nanodrop mode) were used to measure total protein concentration. A280 readings were converted to molar concentration using the Beer-Lambert equation (ϵ = 19.94 $mM^{-1}cm^{-1}$) . Mass concentration was calculated assuming a molecular weight of 24.21 kDa. Similarly, DNA concentrations were measured using absorbance (λ = 260 nm, ϵ = 0.02 $uL{\cdot}ng^{-1}cm^{-1}$, nanodrop mode). A260/A280 and A260/A230 ratios were used to assess protein and RNA contamination. Ratios greater than 1.8 were generally considered to be suitable for downstream processing.

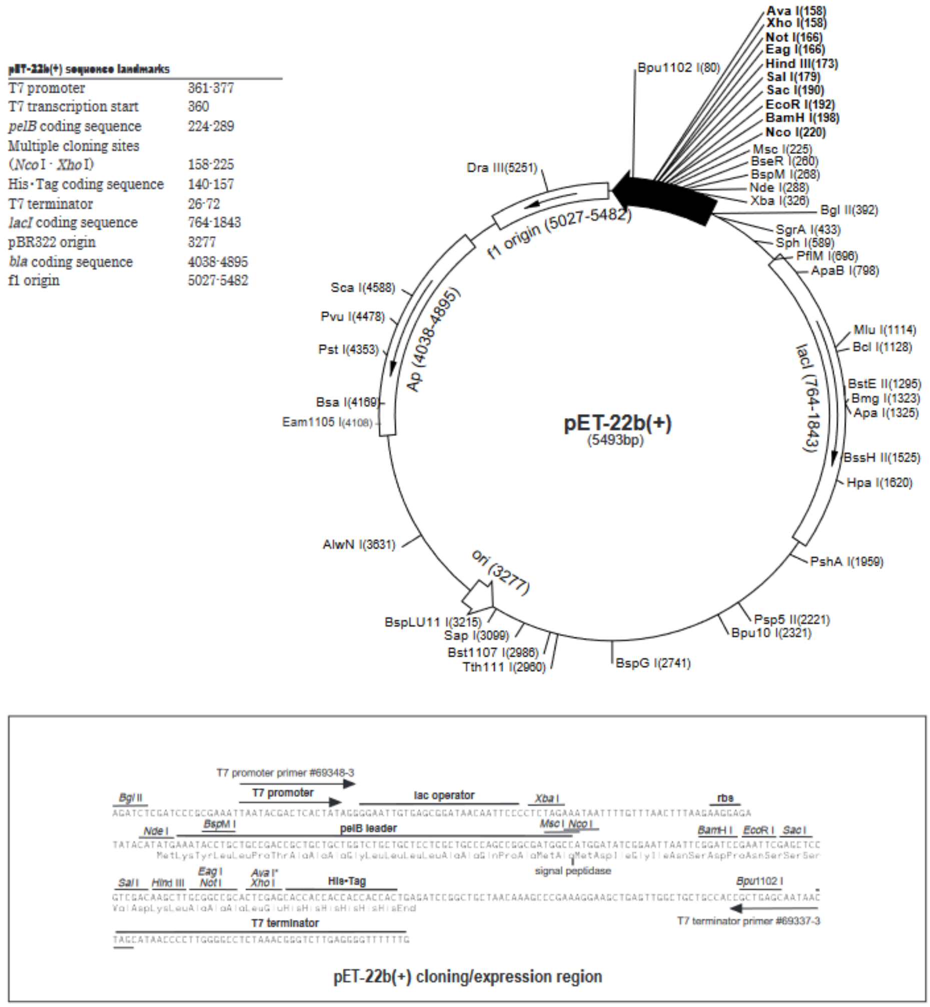
SDS-PAGE

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) was used to monitor protein content during protein expression and purification. Bio-Rad Mini-Protean Tetra cell apparatus was used to cast acrylamide gels and to carry out electrophoresis. 18% BisAcrylamide Resolving Gel solutions were prepared and immediately added to a secured Bio-Rad gel casting cassette. A few drops of isopropanol were added to level the acrylamide solution. The acrylamide solution was allowed to polymerise for < 20 min. Residual isopropanol was poured off. Stacking gel solution was prepared and layered on top of the resolving gel. A comb was inserted into the top of the cassette and the stacking gel solution was allowed to set in < 20 min. Gels were either used immediately or stored in wet tissue roll at 4°C for up to 2 weeks.

Cast gels were loaded into the BioRad companion electrode module and suspended in SDS running buffer. 15 uL sample was mixed with 5 uL SDS loading buffer and heated to ~95 °C for 3min. The samples were loaded into wells of the gel. Gels were run for 5 mins at 50 V and 45-60 mins at 180 V. Gels were stained with InstantBlue™ Coomassie Stain.

Site-directed mutagenesis

The *pgmB* gene from *Lactococcus lactis* was cloned into pET-22b(+) plasmid cloned into the multiple cloning site (NdeI and XhoI) [61] (**Figure 5-1**). The pET-22b(+) plasmid containing the βPGM_{WT} gene was purified from recombinant *E. coli* cells using the Monarch ® plasmid miniprep kit (T1010L) following the manufacturer's instructions. A 5 mL starter culture of transformed XL1-Blue cells were grown overnight at 36 °C (<16hr incubation, as per manufacturer instructions). Autoclaved milliQ-H_2O was used in place of the elution buffer. Agilent's Quickchange II site-directed mutagenesis kit was used to introduce the N77A mutation into βPGM_{WT} plasmid following the manufacturer's instructions. The reaction mixture was comprised of 20 ng plasmid, 125 ng of each primer, dNTP mixture, PfuUltra High Fidelity DNA polymerase in Monarch's reaction buffer diluted in sterile milliQ-H_2O to a final volume of 50 uL. SDM was carried out using Techne Progene FPR0G050 Thermocycler, with 30 s at 95 °C followed by 16 cycles of 30 s at 95 °C, 60 s at 55 °C and 6 min at 68 °C, followed by Dpn1 digestion for 1hr at 37 °C. Digest products were transformed into XL1-Blue super-competent cells using Agilent's recommended heat shock protocol (30 min on ice; 45s at 42 °C; 2 min on ice). Mutations were confirmed by DNA sequencing (GENEWIZ, Azenta Life Sciences), using the T7 promotor/terminator sequences for forward/reverse sequencing, respectively.



pET-22b(+) sequence landmarks	
T7 promoter	361-377
T7 transcription start	360
pelB coding sequence	224-289
Multiple cloning sites (*Nco* I - *Xho* I)	158-225
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	764-1843
pBR322 origin	3277
bla coding sequence	4038-4895
f1 origin	5027-5482

Figure 5-1 – The pET-22b(+) vector. (TOP) the pET-22(b) plasmid map. The location of the multiple cloning site (MCS) is indicated with a black arrow. (BOTTOM) A detailed diagram showing the location of functional elements and restriction sites within the MCS.

N77A site-directed mutagenesis sequences

> pgmB (WT)

atgtttaaagcagtattgtttgatttagatggtgtaattacagataccgcagagtatcattttagagct

tggaaagctttggctgaagaaattggcattaatggtgttgaccgccaatttaatgagcaattaaaaggg

gtctcacgagaagactcgcttcagaaaattctagatttagctgataaaaaagtatcagctgaggaattt

aaagaacttgctaagagaaaa**aat**gataactatgtgaaaatgattcaggatgtgtcgccagccgatgtc

tatcctggaattttacaattactcaaagatttacgttcaaataaaatcaaaattgctttagcgtcggct

tctaagaatggtccattttttattagagagaatgaatttaactggatattttgatgcaattgctgatccg

gctgaagttgcagcatcaaaaccagcaccagatattttttattgcagcagcacatgcagtgggtgttgcc

ccctctgaatcaattgggttagaggattctcaagctggaattcaagccatcaaagattcaggggcttta

ccaattggtgtagggcgcccagaagatttgggagatgatatcgtcattgtgcctgatacttcacactat

acattagaattttttgaaagaagtttggcttcaaaagcaaaaataa

N77 CODON

Primer Name	**Primer Sequence (5' to 3')**
a229g_a230c_	cctgaatcattttcacatagttatc**agc**ttttctcttagcaagttctttaaattcctca
	tgaggaatttaaagaacttgctaagagaaaagctgataactatgtgaaaatgattcagg

Primer Name	**Length (nt.)**	**Tm**	**Duplex Energy at 68 °C**	**Energy Cost of Mismatches**
a229g_a230c_	59	78.48 ℃	-49.09 kcal/mole	2.96%
	59	78.48 ℃	-50.68 kcal/mole	1.17%

> pgmB N77A

atgtttaaagcagtattgtttgatttagatggtgtaattacagataccgcagagtatcattttagagct

tggaaagctttggctgaagaaattggcattaatggtgttgaccgccaatttaatgagcaattaaaaggg

gtctcacgagaagactcgcttcagaaaattctagatttagctgataaaaaagtatcagc**tgaggaattt**

aaagaacttgctaagagaaaagct**gataactatgtgaaaatgattcagg**atgtgtcgccagccgatgtc

tatcctggaattttacaattactcaaagatttacgttcaaataaaatcaaaattgctttagcgtcggct

tctaagaatggtccattttttattagagagaatgaatttaactggatattttgatgcaattgctgatccg

gctgaagttgcagcatcaaaaccagcaccagatattttttattgcagcagcacatgcagtgggtgttgcc

ccctctgaatcaattgggttagaggattctcaagctggaattcaagccatcaaagattcaggggcttta

ccaattggtgtagggcgcccagaagatttgggagatgatatcgtcattgtgcctgatacttcacactat

acattagaattttttgaaagaagtttggcttcaaaagcaaaaataa

Preparation of chemically competent BL21(DE3) expression host

A 100 mL LB liquid culture of BL21(DE3) *E. coli* was incubated at 37 °C until the culture OD_{600} reached ~0.5. Cells were collected by centrifugation for 5 mins at 4,500 RPM, 4 °C (Thermo Scientific Heraeus Labofuge 400 R). Cells were gently resuspended in 30 mL ice-cold TFB1 buffer and incubated on ice for 90 min. Cells were collected by for 5 mins at 4,500 RPM, 4 °C (Thermo Scientific Heraeus Labofuge 400 R) and gently resuspended in ice-cold TFB2 buffer. 200 uL aliquots were flash frozen and stored at -80 °C.

Transformation of competent BL21(DE3)

50 uL chemically competent BL21(DE3) were mixed with 5 ng plasmid DNA in a pre-chilled 15 mL round-bottom, polypropylene Falcon tube and incubated on ice for 30 min. The cells were heat-shocked at 42 °C for 30s and immediately placed back on ice for 2 min. 5 mL SOC was pre-heated to 37 °C and added to the heat-shocked cells and incubated for 1 hr at 37 °C. The culture was used to inoculate solid LB agar plates containing 100 µg/mL Ampicillin, which were incubated overnight (~16 hr) at 37 °C.

Expression

A 5 mL LB liquid culture containing 100 ug/ml Ampicillin was inoculated with BL21(DE3) *E. coli* cells containing the pgmB encoding the N77A mutation and grown overnight at 37 °C. The culture was centrifuged 4,500 RPM, 5 min, 4 °C (Thermo Scientific Heraeus Labofuge 400 R) and re-suspended in 1 L M9 minimal media containing 100 ug/mL ampicillin and distribute between two sterile baffled conical flasks. The cultures were incubated at 37 °C and induced with 0.5 mM IPTG when the culture reaches OD_{600} 0.6. The cultures were then incubated overnight at 25 °C.

Cell Lysis & purification

Cells were harvested by centrifugation at 8,000 RPM using a Beckman Coulter Avanti centrifuge, with a F10BCI-6x500y rotor (FIBERLite, Piramoon technologies); resuspended in ice cold Lysis Buffer; and subject to six rounds of sonication (20 s pulse, 20 % amplitude, 60 s rest). The lysate was centrifuged at 20,000 RPM for 35 min at 4 °C, using a Beckman Coulter Avanti centrifuge with a Beckman JA-20 rotor. The supernatant was filtered using a 0.22 µm syringe filter. The soluble fraction was loaded onto a DEAE-sepharose Fast-Flow Anion Exchange column, washed pre-equilibrated with chromatography buffer A, using an ÄKTA start purification system. Unbound proteins were eluted from the column with chromatography buffer A. Bound proteins were eluted with a 0-50% chromatography buffer B gradient over 300 mL. Fractions containing βPGM were identified by SDS-PAGE, combined and concentrated using a Vivaspin centrifugal concentrator (Satorius) with a 10 kDa molecular weight cutoff, following the manufacturer instructions. The concentrated sample was loaded onto a Hiload 26/600

Superdex 75 size-exclusion column, pre-equilibrated with 3 column volumes of chromatography buffer B, using an ÄKTA Prime purification system. Fractions were checked for purity by SDS-PAGE, combined and exchanged into βPGM buffer and concentrated to ~ 1 mM using a Vivaspin centrifugal concentrator (Sartorius). Aliquts were stored at -20 °C.

5.2.6 Small Angle X-ray Scattering

Sample Preparation, Data Acquisition and Analysis

βPGM was exchanged into different buffers to create different complexes.

- βPGM$_{WT}$ was exchanged into 50 mM K$^+$ HEPES pH 7.2 + 5 mM $MgCl_2$ + 2 mM NaN_3 to generate a substrate-free magnesium loaded complex, populating both *cis* and *trans* P146 isomers in solution.
- βPGM$_{WT}$ was exchanged into 50 mM K$^+$ HEPES pH 7.2 + 100 mM $MgCl_2$ + 2 mM NaN_3 to generate a substrate-free magnesium loaded complex, which populates the *cis* P146 isomer.
- βPGM$_{P146A}$ was exchanged into 50 mM K$^+$ HEPES pH 7.2 + 5 mM $MgCl_2$ + 2 mM NaN_3 to generate a substrate-free magnesium loaded complex, which populates a *trans* A146 isomer.
- βPGM$_{D10N}$ was incubated with 5 mM MgCl2, 10 mM AcP and 10 mM G6P at room temperature overnight, to generate a trapped βG16BP complex. The complex was exchanged into 50 mM K$^+$ HEPES pH 7.2 + 5 mM $MgCl_2$ + 2 mM NaN_3.
- βPGM$_{D10N}$ was incubated with 5 mM MgCl2, 10 mM AcP and 10 mM G6P at room temperature, overnight, to generate a trapped βG16BP complex. The complex was exchanged into 50 mM K$^+$ HEPES pH 7.2 + 2 mM NaN_3, to remove magnesium from the complex.
- βPGM$_{WT}$ was exchanged into 50 mM K$^+$ HEPES pH 7.2 + 5 mM $MgCl_2$ + 10 mM G6P + 20 mM NaF + 2 mM NaN_3 to generate a canonical transition state analogue complex.

Small angle X-ray scattering data were collected at the Diamond Light Source, beamline B21 which operates in a 3.6 meter camera length configuration at 13 keV and wavelength 0.9524Å. Data was collected using an EigerX 4M (Dectris) detector. Data collection was coupled to an upstream Agilent 120 HPLC system, equilibrated with matched sample buffer.

Radially averaged data was integrated and buffer-subtracted using Chromixs from the ATSAS software suite [171]. Processed data from the βPGM$_{WT}$ substrate-free sample were submitted to the WAXSIS sever [131] with each of 64 processed crystal structures and a sample of 45 frames from the ff14sb-tip3p-1 simulation. The electron density of the solvent was set to 339 $e{\cdot}nm^{-3}$ with ‘thorough’ convergence criteria. The same SAXS dataset and crystal structures were analysed using Pepsi-SAXS software package [140], which uses an implicit solvation model. A parameter-sweep of $\delta\rho$ and r_0 was performed to determine which values gave best agreement with the explicit solvation model implemented in the WAXSIS sever. These parameters ($\delta\rho$ = +5 %, r_0 = -1 %) were then used to process the remaining datasets paired with all 64 processed crystal structures and 1000 frames from the ff14sb-tip3p-1 simulation.

Pepsi-SAXS Python Wrapper

```
import os              # operating system interface
import shutil          # high-level file interface
import glob            # unix style file patten expansion
import subprocess      # run external processes
import numpy as np     # numerical operations

def run_pepsisaxs(pepsipath, saxsdata, outdir, indir, opt):

    # make output directory
    try:
        os.mkdir(outdir)
    except FileExistsError:
        shutil.rmtree(outdir)
        os.mkdir(outdir)

    # define a list of command line arguments
    args = [
        pepsipath,     # path to pepsi-saxs program exe file
        None,          # placeholder for pdbfile path
        saxsdata,      # input saxsdata.dat
        None           # placeholder for output file
    ]

    # append optional command line arguments
    args += opt

    # iterate through input pdb files and run pepsi-saxs program
    for pdbpath in glob.glob(indir):

        # define input file.pdb
        args[1] = pdbpath

        # define output file name
        filename = pdbpath.split("/")[-1].split(".")[0]
        outfile = outdir + filename + ".fit"
        args[3] = f'-o {outfile}'

        # run program
        subprocess.run(args)
```

A python wrapper was created to run the Pepsi-SAXS program [140] for a batch of PDB files. On execution, the wrapper creates an output directory and populates an argument list to provide to the Pepsi-SAXS executable. This list is updated with each of the input batch of PDB files before execution of the Pepsi-SAXS program.

Pepsi-SAXS Parameter Sweep

```
# loop through values for δρ
for dro in np.arange(2, 8, 0.5):
    # loop through values for ,r0.
    for d0 in np.arange(0.95, 1.05, 0.01):
        # run PepsiSAXS python wrapper
        run_pepsisaxs(
            outdir=f"./dro-{int(dro*10)}_d0-{int(d0 * 100)}/",
            indir="path/to/pdbfiles/*.pdb",
            pepsipath="/path/to/pepsi-saxs.exe",
            saxsdata="/path/to/saxsdata.dat",
            opt=[
                f'--dro {dro}',             # solvation layer
                f'--r0_min_factor {d0}',    # r0 parameter lower restraint
                f'--r0_max_factor {d0}',    # r0 parameter upper restraint
                '--cst',                    # fit a constant baseline subtraction
            ]
        )
```

Pepsi-SAXS fitting of the data acquired for the βPGM$_{WT}$ substrate-free magnesium loaded complex (populating both *cis* and *trans* P146 isomers in solution), against all of the crystal structures and each of 1000 frames of the ff15ipq-speceb-ecc-2whe simulation, was executed using an array of fixed values for δρ (**`dro`**) and r$_0$ (**`d0`**). The PepsiSAXS python wrapper was executed in a loop over a range of δρ and r$_0$ values. The values for δρ and r$_0$ which resulted in χ^2 vs. R$_g$ profile which best matched that generated using calculations made by the WAXSIS server [131], were taken forward to the analysis of all SAXS datasets.

Pepsi-SAXS Global Rollout

```
# define inputs
pdbfiles = glob.glob("/path/to/pdbfiles/*.pdb")
saxsdata = glob.glob("/path/to/saxsdata/*.dat")

# loop through data for each sample
for saxsdata in saxdatasets:

    # run pepsisaxs python wrapper
    run_pepsisaxs(
        pepsipath="/path/to/pepsi-saxs.exe",
        saxsdata=saxsdata,
        outdir=saxsdata.split("/")[-1].split(".")[0],
        indir=pdbfiles,
        opt=[
            '--dro 5',
            '--r0_min_factor 0.99',
            '--r0_max_factor 0.99',
            '--cst',
        ]
    )
```

Pepsi-SAXS fitting of each acquired dataset against each of the crystal structures and each of 1000 frames from the ff15ipq-speceb-ecc-2whe simulation, was executed using a fixed, optimised value for δρ = 5 % and r$_0$ = 0.99 (- 1%).

5.2.7 Analytical Ultracentrifugation

Sample preparation

Two samples were prepared for analytical ultracentrifugation (AUC). A magnesium-loaded substrate-free βPGM$_{WT}$ complex was prepared by exchanging βPGM$_{WT}$ into 50 mM K+ HEPES pH 7.2 + 5 mM $MgCl_2$ + 2 mM NaN_3. A βPGM$_{WT}$:MgF_3:G6P TSA complex was prepared by exchanging βPGM$_{WT}$ into 50 mM K+ HEPES pH 7.2 + 5 mM $MgCl_2$ + 5 mM G6P + 20 mM NaF + 2 mM NaN_3.

Data Acquisition

AUC sedimentation velocity experiments were run using a Beckman Optima analytical ultracentrifuge with an An-50Ti rotor, at 50,000 RPM. Samples with their corresponding buffers were loaded into each side of standard two-sector epon centerpieces. The reference data is subtracted from the sample data during data acquisition. Data were recorded using the absorbance optical detection system at 280 nm, at three different concentrations (0.25 mg/mL, 0.5 mg/mL and 1.0 mg/mL). Buffer densities and viscosities were measured using a DMA 5000M densometer equipped with a Lovis 2000ME viscometer.

Data Analysis

Radial Absorbance scans were analysed using SVEDBERG software [172]. Scans 2-250 were selected and the data between 0.02-7.00 cm from the meniscus were fit using the Behlke whole boundary model [134], with a time-invariant noise correction. Starting values were estimated from c(s) distributions calculated using SEDFIT [173]. The data derived from the closed βPGM$_{WT}$:MgF_3:G6P TSA complex was fit to a single species model. The data derived from the open βPGM$_{WT}$ substrate free complex was fit to a 3 species model. The c(s) analysis indicated the presence of at least 3 species for each dataset. The dominant population is the folded monomer. A small amount of higher molecular weight, aggregated material was also present. The H-statistics for were not significantly improved by including more than 3 species, whilst parameter covariances were introduced by inclusion of more than 3 species.

5.2.8 NMR Relaxation

Sample preparation, data acquisition and processing were carried out by Dr Angus Robertson. A βPGM$_{WT}$ substrate-free sample was prepared containing 50 mM K$^+$-HEPES pH 7.2 + 1 mM ^{15}N^{2}H-labelled βPGM$_{WT}$ + 5 mM MgCl$_2$ + 2 mM NaN$_3$. A βPGM$_{WT}$:AlF$_4$:G6P TSA was prepared by addition of a further 10 mM G6P + 10 mM NaF. ^{15}N R1, R1ρ and ^{1}H-{^{15}N} measurements were carried out at 25 °C at each of two field strengths (600 MHz, 800 MHz), with a temperature-compensated, TROSY-readout pulse scheme [137]. Pseudo-randomised decay durations were used for R1 experiments: 0, 400, 1200, 80, 1760, 240, 800, 2400 and 640ms and R1ρ experiments: 20, 110, 1, 200, 60, 90, 160, 40 and 20ms. The spin lock-field was set to 1.4 kHz. Peak intensities were fit to time-dependent exponential equations to derive R1 and R1ρ. R2 values were calculated using the following equation.

$$R_2 = \frac{R_{1\rho}}{\sin^2(\theta)} - \frac{R_1}{\tan^2(\theta)}$$

Where, $\theta = \frac{\omega_1}{\Omega}$, ω$_1$ is the spin lock field strength and Ω is the offset from the ^{15}N carrier frequency. Data were analysed using a local τ$_m$ diffusion model and a prolate diffusion model using the RELAX software suite [138,174]. The prolate diffusion model was fit using either an open crystal structure (PDB: 6YDL) or the average coordinates from the ff15ipq-spceb-ecc-2whe simulation.

5.2.9 X-ray Crystallography

Crystallisation

Crystals of $\beta PGM_{N77A}:MgF_3:G6P$ TSA complex were obtained by preparing solutions containing 0.7 mM βPGM_{N77A}, 5 mM $MgCl_2$, 20 mM NaF and 5 mM G6P, which were mixed 1:1 with a solution containing 100 mM Tris-HCl pH 7.5, 200 mM Sodium Acetate and 28 % PEG4000. Crystals of $\beta PGM_{N77A}:AlF_4:G6P$ TSA complex were obtained by preparing solutions containing 0.6 mM βPGM_{N77A}, 5 mM $MgCl_2$, 5 mM $AlCl_3$, 20 mM NaF and 5 mM G6P, which were mixed 1:1 with a precipitant solution containing 100 mM Tris-HCl pH 7.5, 200 mM Sodium Acetate and 22 % PEG4000. Crystals were grown by vapour diffusion using 2 uL hanging drops, suspended from a silanised glass slide, over a reservoir containing 700 uL of precipitant soliton. TSA complexes grew as thin plates over 1 – 2 weeks. Crystals were collected using a mounted LithoLoop (Molecular Dimensions Ltd.) in a drop of precipitant soliton containing 25% ethylene glycol.

Data Acquisition & processing

Crystals were sent to the Diamond Light Source (DLS), Oxfordshire, United Kingdom for data collection at 100 K on beamline ID14-2. Data processing was done using the xia2 pipeline [175]. CC-half values were used to define the resolution cut-off. Structures were solved by molecular replacement using MolRep [176], using 2WF5 and 2WF6 for the $\beta PGM_{N77A}:MgF_3:G6P$ and $\beta PGM_{N77A}:AlF_4:G6P$ data, respectively. Model building and refinement was done using COOT and REFMAC5 [177]. Ligands were added during the final stages of refinement. COOT and MolProbity [178] were used for validation of the geometry of the final structures. Figures were created using PyMOL [179].

Linear Determinant Analysis

Linear determinant analysis (LDA) was implemented in Python using the MDAnalysis (version 2.3.0) [169,170], Scikit-Learn (version 1.2.1) [180] and Scipy (version 1.5.3) [181] libraries. The interatomic Euclidean distances between all pairs of backbone atoms were calculated, for each of the βPGM crystal structures. Intra-residue distances (excluding O_i and N_i), and distances calculated between atoms which are NOT within the same helical turn (between O_i and N_{i+4}) were discarded. The interatomic distance (feature) vector was calculated for each structure (observation). The distances were normalized by subtracting the mean and scaling the variance to unity. Linear discriminants were optimized for the interatomic distance array using the Open, NAC and TSA clusters defined by backbone RMSD hierarchical clustering.

```
# -----------------
# Libraries
# -----------------

import numpy as np
import MDAnalysis as mda
from scipy.spatial.distance import pdist, squareform
from sklearn.preprocessing import StandardScaler
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis

# -----------------
# Feature vector
# -----------------

def N77_to_M83_distance_vector(pdbfile):
    """ calculate the interatomic distance vector
    from a complete bpgm pdb file """

    # read pdbfile using MDAnalysis library
    u = mda.Universe(pdbfile)

    # select atoms for the analysis
    atoms = u.select_atoms(f"resid 77-83 and backbone")

    # calculate the interatomic distance vector
    distance_vector = pdist(atoms.positions, metric="euclidean")

    # convert distance vector to interatomic distance matrix
    distance_matrix = squareform(distance_vector)

    # create an array of indices corresponding to those of the distance matrix
    i, j = np.meshgrid(range(len(distance_matrix)), range(len(distance_matrix)))

    # create a boolean filter to select specific distances from the distance matrix
    dmat_mask = (
        # keep atoms which are within the same helical turn (N[i] - O[i+4])
        (np.abs(i - j) <= 16) & \
        # discard intra-residue distances (except the N[i] - O[i] distance)
        (np.abs(i - j) >= 3) & \
        # discard the diangonal
        (np.abs(i - j) != 0)
    )

    # convert the square boolean filter to a vector format
    dvec_mask = squareform(dmat_mask)

    # apply boolean filter
    filtered_distance_vector = distance_vector[dvec_mask]

    return filtered_distance_vector

# calculate the interatomic distances for each input pdbfile
data_list = [N77_to_M83_distance_vector(pdbfile) for pdbfile in pdbfiles]

# store the data in a table format
data_array = np.stack(data_list, axis=0)

# Normalise the data for Linear Discriminant Analysis
data_scaled = StandardScaler().fit_transform(data_table)

# initialise the LDA solver
lda = LinearDiscriminantAnalysis()

# fit linear discriminants to the scaled data, using categories [open, nac, tsa]
lda.fit(data_scaled, user_defined_categories)

# project observations onto the linear discriminants
linear_discriminants = lda.transform(data_scaled)
```

Root-Mean-Square-Fluctuation Analysis

The RMSD was calculated between every pair of TSA crystal structures, by alignment to the core-domain and calculating the interatomic distances between equivalent backbone atoms. The data were split into two groups 1) the βPGM$_{N77A}$ TSA crystal structures and 2) all other βPGM TSA structures (excluding: 1Z4O$_A$, 6I03$_A$ and 5OLY$_A$, which have unusual domain orientation). The mean and standard deviation of the per-residue RMSD was calculated for each group. The mean RMSD for the canonical βPGM TSA structures was subtracted from that of the βPGM$_{N77A}$ TSA crystal structures. The residual RMSD represents a statically significant perturbation in structure. The residual error was calculated as the sum-of-squares of the standard deviation calculated within each group.

5.2.10 G6P titration in the presence of MgF$_3$ TSA

Apparent dissociation constants (K$_D$) for G6P were measured for βPGM$_{N77A}$ were determined by serial titration of a solution of 20 mM G6P in standard HEPES buffer, into a solution containing 0.5 mM βPGM$_{N77A}$ + 5 mM MgCl2 + 10 mM NaF + 1 mM Trimethylsilyl propionate (TSP) + 10 % D$_2$O. The titration was monitored by ^{1}H NMR Spectroscopy using a Bruker 800 MHz Avance III spectrometer equipped with a 5-mm TCI cryoprobe and *z* axis gradients, at 298 K. The dilution-corrected intensity of W24 indole resonance was fit to a standard 1:1 binding model for tight binding.

$$I = I_0 \frac{([G6P] + [E]_0 + K_D) - \sqrt{([G6P] + [E]_0 + K_D)^2 - 4\,[E]_0[G6P]}}{2[E]_0}$$

$$\text{where,} \quad [G6P] = \text{dilution} \times [G6P]_0$$

Monte-Carlo sampling of [E]$_0$ ± 5 % and 5 % pipetting error was combined with free fitting of K$_D$, I$_0$ and [G6P]$_0$ to estimate the error associated with fitting a value for K$_D$ close to the tight-binding limit. Note, fitting the titrant concentration [G6P]$_0$, using the dilution-factor as independent variable, is only possible where the ratio of [E]$_0$ and [G6P]$_0$ is constrained by data obtained close to the tight binding limit. Least-squares, non-linear regression and Monte-Carlo sampling were implemented using the Scipy and Numpy python libraries, respectively.

5.2.11 αGal1P titration

The dissociation constants (K$_D$) for αGal1P and end-point chemical shifts βPGM$_{WT}$ and βPGM$_{N77A}$ were determined by serial titration of a solution of 20 mM G6P in standard HEPES buffer, into a solution containing 0.5 mM βPGM$_{N77A}$ + 10 mM MgCl$_2$ + 1 mM Trimethylsilyl propionate (TSP) and 10 % D$_2$O. The titration was monitored by ^{1}H^{15}N BEST-TROSY at 298 K, using a Bruker 800 MHz Avance III spectrometer equipped with a 5-mm TCI cryoprobe and z-axis gradients. Binding occurred under a fast-exchange regime, meaning the chemical shifts could be monitored incrementally throughout the titration. Aggregate chemical shifts were calculated for of the indole resonance of W24 and were fit to a 1:1 binding model.

5.2.12 Steady State Kinetics

Steady state β-Glucose-1-phosphate (βG1P) turn-over kinetics for $\beta\mathrm{PGM}_{\mathrm{WT}}$ and $\beta\mathrm{PGM}_{\mathrm{N77A}}$ were acquired using Hidex Sense microplate reader. The production of G6P was measured indirectly using a glucose-6-phosphate dehydrogenase (G6PDH) coupled assay. Oxidation of G6P is coupled to reduction of NAD^+, which is monitored by absorbance (λ = 340 nm, ϵ = 6220 M^{-1} cm^{-1}). Reactions were initiated by the addition of βG1P and β-Glucose-1,6-bisphosphate (βG16BP) to solutions containing 5 mM MgCl_2 5 Units/mL G6PDH, 1 mM NAD^+ and 4 nM $\beta\mathrm{PGM}_{\mathrm{WT}}$ or $\beta\mathrm{PGM}_{\mathrm{N77A}}$ in standard kinetics buffer (200 mM K^+ HEPES pH 7.2 + 5 mM MgCl_2 + 2 mM). A variable range of βG1P and βG16BP concentrations were prepared by serial dilution ($\beta\mathrm{PGM}_{\mathrm{WT}}$: [βG1P] 0-700 μM, [βG16BP] 0-16 μM; $\beta\mathrm{PGM}_{\mathrm{N77A}}$: [βG1P] 0-1800 μM, [βG16BP] 0-22 μM). The data were fit to a steady state equation derived for a bi bi ping-pong reaction mechanism, with substrate inhibition [182]. Errors were estimated using Monte Carlo sampling with a 5% error for the concentrations of βG1P, βG16BP and E_0. The error in v_0 was derived from linear least-squares fit of the time-course data.

$$v_0 = \frac{k_{cat}E_0[\beta G1P][\beta G16BP]}{[\beta G1P][\beta G16BP] + K_{\beta G1P}[\beta G16BP] + K_{\beta G16BP}[\beta G1P]\left(1 + \frac{[\beta G1P]}{K_i}\right)}$$

Non-linear least squares regression was done using a the `Scipy` python library. The `scipy.optimize.curve_fit` function takes a function with a single input variable (x) and an unspecified number of parameters to be optimised. The code below acts as an interface to allow fitting of multivariate data with `scipy.optimize.curve_fit`.

```
class Fit2D:

    def __init__(self):

        pass

    def load_data(self, x, y, z, c, x_var_name, y_var_name, z_var_name):
        """ Store input data as Fit2D class attributes
               - x and y must be 1 dimensional np.arrays
               - z must be 2 dimensional
               - x and y must correspond to rows and columns of z """

        self.x = x
        self.y = y
        self.z = z
        self.c = c
        self.x_var_name = x_var_name
        self.y_var_name = y_var_name
        self.z_var_name = z_var_name

    def load_model(self, func):
        """ a user-defined python function which takes two
            input arrays (x, y), a list of constants (c),
            a list of params (parms) and returns an output
            array (z). x, y and z have the shape and the
            same number of dimensions """

        self.func = func

    def _get_scipy_compatible_2dfunction(self, y, c):
        """ convert a function with multiple inputs (x,y,c,parms)
        to a function with a single input variable and parameters
        (x, *parms) """

        def function(x, *parms):

            return self.func(x, y, c, list(parms))

        return function

    def run_fit(self, p0, bounds):

        # convert x, y and z into a 1D array of same length
        x, y, c = _broadcast_data(self.x, self.y, self.c)

        # convert z-array into a 1D array
        z = self.z.flatten()

        # filter N/A values
        x, y, z, c = filter_missing_values(x, y, z, c)

        # convert the input 2D function to a pseudo-1D function
        func = self._get_scipy_compatible_2dfunction(y, c)

        # run non-linear least squares curve fitting sub-routine
        popt, pcov = scipy.optimize.curve_fit(
            f=func,
            xdata=x,
            ydata=z,
            p0=p0,
            bounds=bounds,
        )

        # return optimised parameters
        self.popt = popt

        # return parameter covariances
        self.pcov = pcov
```

```
    def run_monte_carlo(self, x_err, y_err, z_err, c_err, p0, bounds, N):

        # generate random sampling of monte-carlo variables
        x_mc = np.random.normal(self.x, x_err, (N, len(self.x)))
        y_mc = np.random.normal(self.y, y_err, (N, len(self.y)))
        z_mc = np.random.normal(self.z, z_err, (N, len(self.x), len(self.y)))
        c_mc = np.random.normal(self.c, c_err, (N, len(self.c)))

        self.popt_mc = []

        # loop through monte carlo iterations
        for i in tqdm(range(N)):

            # convert x, y and z into a 1D array of same length
            x, y, c = _broadcast_data(x_mc[i], y_mc[i], c_mc[i])

            # convert z-array into a 1D array
            z = z_mc[i].flatten()

            # filter N/A values
            x, y, z, c = filter_missing_values(x, y, z, c)

            # convert the input 2D function to a pseudo-1D function
            func = self._get_scipy_compatible_2dfunction(y, c)

            # run non-linear least squares curve fitting sub-routine
            popt, pcov = scipy.optimize.curve_fit(
                f=func,
                xdata=x,
                ydata=z,
                p0=p0,
                bounds=bounds,
            )

            # store data in list
            self.popt_mc.append(popt)

        self.popt_mc = np.stack(self.popt_mc, axis=0)

def _broadcast_data(x, y, c):

    # generate a 2d meshgrid of x and y data
    y, x = np.meshgrid(y, x)

    # arange the grid to a 1d array
    x = x.flatten()
    y = y.flatten()

    # broadcast c to match x & y
    c = np.repeat(c[:, np.newaxis], len(x), axis=1)

    return x, y, c


def filter_missing_values(x, y, z, c):

    # create a filter
    filt = ~np.isnan(z)

    # filter x, y z and c arrays
    return x[filt], y[filt], z[filt], c[:, filt]
```

5.2.13 G6P titration in the presence of BeF_3 GSA

Apparent dissociation constants (K_D) for G6P and end-point chemical shifts βPGM_{WT} and βPGM_{N77A} were determined by serial titration of a solution of 20 mM G6P in standard HEPES buffer, into a solution containing 0.5 mM βPGM_{N77A} + 10 mM $MgCl_2$ + 5 mM $BeCl_2$ + 15 mM NaF + 1 mM Trimethylsilyl propionate (TSP) and 10 % D_2O. The titration was monitored by $^1H^{15}N$ BEST-TROSY at 298 K, using a Bruker 800 MHz Avance III spectrometer equipped with a 5-mm TCI cryoprobe and *z* axis gradients. Resonances for $\beta PGM_{N77A}:BeF_3$ were assigned by comparison with the chemical shifts from $\beta PGM_{WT}:BeF_3$ (BMRB: 17851). Binding occurred under a fast-exchange regime, meaning the chemical shifts could be monitored incrementally throughout the titration. Assignments were validated by comparison of the end-point chemical shifts with those of the $\beta PGM_{WT}:BeF_3:G6P$ complex (BMRB: 7234). Dissociation constants were fit to a 1:1 binding model. End-point chemical shifts were derived by linear extrapolation of chemical shift trajectories to infinite concentration of G6P.

5.2.14 Multiple Sequence Alignment

A diverse set of βPGM sequences (Section 7.12 - Uniprot Accession Codes) were subject to multiple sequence alignment using the Clustal Omega web server [183]. The alignment conservation score was calculated using Jalview [184,185]. The selected sequences were sufficiently diverse to detect residue conservation whilst maintaining sufficient structural similarity with βPGM from *Lactococcus lactis*.

6 Appendix

6.1 Thermodynamics

When studying conformational dynamics of biomolecules, we would like to understand the relative stability of different conformations, how they interchange and ultimately how they relate to function. Conformations of biomolecules can be described as continuous distribution in a high dimensional space. This conformational space has a dimension for each bond length, angle, and torsion within the biomolecule (aka. internal/BAT coordinates) as well as for every possible configuration of the solvent.

It is often helpful to reduce the dimensionality of this landscape allowing one to conceptualise just a few important variables, such as a domain orientation, a secondary structure metric or a few selected torsions/distances. More sophisticated dimensionality reduction techniques have been developed to extract more abstract variables in an unsupervised manner (e.g. Principal Component Analysis).

Each point within the conformational space represents a *microstate* of the system. Each microstate within the conformational landscape has an associated potential/internal energy (E). The probability of sampling a particular microstate (P) in thermodynamic equilibrium is defined by the Boltzmann equation,

$$\mathrm{P} = \frac{\mathrm{e}^{-\mathrm{E}/\mathrm{k_B T}}}{\mathrm{Z}}$$

where E is the potential energy, $\mathrm{k_B}$ is the Boltzmann constant, T is the temperature above absolute zero and Z is the partition function. The partition function takes a sum over all microstates.

$$Z = \sum_i e^{-\frac{E_i}{k_B T}}$$

Certain regions within the conformational landscape are stabilized by favourable interactions (local energy minima), which often prompts discretisation of the continuous distribution into conformational *macrostates* (e.g. open/closed, apo/bound, helix/strand/loop, cis/trans etc.). Macrostates are separated by less favourable conformations which act as barriers to conformational change. Entropy is a measure of the number of microstates belonging to a macrostate.

$$\frac{\mathrm{N_A}}{\mathrm{N}} = \mathrm{e}^{\frac{\mathrm{S_A}}{\mathrm{k_B}}}$$

… where S_A is the entropy of macrostate A and N_A is the number of constituent microstates of A and N is the total number of microstates. If a macrostate is comprised of more microstates, it will be sampled more frequently and will appear more stable. The potential energy and entropy together define the Helmholtz free energy (F).

$$e^{-\frac{E_A}{k_B T}} \cdot e^{\frac{S_A}{k_B}} = e^{-\frac{E_A - TS_A}{k_B T}} = e^{\frac{-F_A}{k_B T}}$$

$$F = E - TS$$

Gibbs Free energy also accounts for pressure-volume work,

$$H = E + PV$$

$$G = H - TS$$

where H is the enthalpy, P is the pressure, V is the volume and G is the Gibbs free energy. Free energy changes describe the ratio of the probabilities/frequencies of different states.

$$\frac{P_A}{P_B} = \frac{e^{-G_A/k_B T}}{e^{-G_B/k_B T}} = e^{\Delta G/k_B T}$$

Therefore…

$$\Delta G = k_B T \cdot \ln\left(\frac{P_A}{P_B}\right)$$

The ratio of probabilities corresponds to the ratio of the populations of A and B, i.e. the equilibrium constant (K_{EQ}).

$$A \rightleftharpoons B$$

$$k_1[A] = k_2[B]$$

$$\frac{k_2}{k_1} = \frac{[A]}{[B]} = \frac{P_A}{P_B} = K_{EQ}$$

$$\Delta G = k_B T \cdot \ln(K_{EQ})$$

6.2 The Michaelis-Menton model

A basic model for an enzyme-catalysed reaction is given by…

$$E + S \rightleftharpoons ES \ \rightleftharpoons E + P$$

Initial rates assumption: Reaction rates are measured at an early timepoint during the reaction such that the accumulation of product and the rate of the reverse reaction are negligible.

Rapid equilibration assumption: The substrate-binding equilibrium is much faster than turnover. This means that the initial rate is directly proportional to the concentration of ES at equilibrium.

$$v_0 = k_{cat}[ES]_{EQ}$$

Quasi-steady-state assumption: The concentration of ES is effectively constant during the period at which the initial rates are measured.

$$\text{Rate of ES formation} = k_1[E][S]$$

$$\text{Rate of ES dissociation} = (k_{-1} + k_2)[ES]$$

$$k_1[E][S] = (k_{-1} + k_2)[ES]$$

$$\frac{k_{-1} + k_2}{k_1} = \frac{[E][S]}{[ES]} = K_m$$

It is assumed that the enzyme and substrate exist in either of two states: associated or dissociated.

$$[E]_{TOT} = [E] + [ES]$$

$$[S]_{TOT} = [S] + [ES]$$

The concentration of substrate is assumed to be in large excess of the total enzyme.

$$[S]_{TOT} \gg [E]_{TOT}$$

$$[S]_{TOT} \approx [S]$$

Under these assumptions, an expression can be derived for the concentration of ES.

$$[ES]K_m = ([E]_{TOT} - [ES])[S]$$

$$[ES]K_m = [E]_{TOT}[S] - [ES][S]$$

$$[ES](S + K_m) = [E]_{TOT}[S]$$

$$[ES] = \frac{[E]_{TOT}[S]}{[S] + K_m}$$

The initial reaction velocity is directly proportional to the concentration of ES.

$$v_0 = k_{cat}[ES] = \frac{k_{cat}[E]_{TOT}[S]}{[S] + K_m}$$

6.3 Two-state 1:1 binding model

A simple 1:1 two state binding model between A and B is stated below.

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB}$$

$$\mathrm{k_1[A][B]} = \mathrm{k_2[AB]}$$

$$\mathrm{K_D} = \frac{\mathrm{k_2}}{\mathrm{k_1}} = \frac{\mathrm{[A][B]}}{\mathrm{[AB]}}$$

$$\Delta \mathrm{G} = -\mathrm{k_B T} \cdot \ln(\mathrm{K_D})$$

The model can be expressed in terms of known experimental concentrations $[\mathrm{A}]_{\mathrm{TOT}}$ and $[\mathrm{B}]_{\mathrm{TOT}}$.

$$\mathrm{K_D} = \frac{([\mathrm{A}]_{\mathrm{TOT}} - [\mathrm{AB}])([\mathrm{B}]_{\mathrm{TOT}} - [\mathrm{AB}])}{[\mathrm{AB}]}$$

$$\mathrm{K_D} = \frac{[\mathrm{AB}]^2 - ([\mathrm{A}]_{\mathrm{TOT}} + [\mathrm{B}]_{\mathrm{TOT}}) \cdot [\mathrm{AB}] + [\mathrm{A}]_{\mathrm{TOT}} \cdot [\mathrm{B}]_{\mathrm{TOT}})}{[\mathrm{AB}]}$$

$$0 = [\mathrm{AB}]^2 - [\mathrm{A}]_{\mathrm{TOT}}[\mathrm{AB}] - [\mathrm{B}]_{\mathrm{TOT}}[\mathrm{AB}] - \mathrm{K_D}[\mathrm{AB}] + [\mathrm{A}]_{\mathrm{TOT}}[\mathrm{B}]_{\mathrm{TOT}}$$

$$\mathrm{x} = -\mathrm{b} \pm \mathrm{b2}{-}4ac\mathrm{x} = \frac{-\mathrm{b}\pm\sqrt{\mathrm{b}^2-4\mathrm{ac}}}{2\mathrm{a}}$$

$$\mathrm{a} = 1$$

$$\mathrm{b} = -([\mathrm{A}]_{\mathrm{TOT}} + [\mathrm{B}]_{\mathrm{TOT}} + \mathrm{K_D})$$

$$\mathrm{c} = [\mathrm{A}]_{\mathrm{TOT}}[\mathrm{B}]_{\mathrm{TOT}}$$

$$\frac{[\mathrm{AB}]}{[\mathrm{B}]_{\mathrm{TOT}}} = \frac{[\mathrm{A}]_{\mathrm{TOT}} + [\mathrm{B}]_{\mathrm{TOT}} + \mathrm{K_D} \pm \sqrt{([\mathrm{A}]_{\mathrm{TOT}} + [\mathrm{B}]_{\mathrm{TOT}} + \mathrm{K_D})^2 - 4[\mathrm{A}]_{\mathrm{TOT}}[\mathrm{B}]_{\mathrm{TOT}}}}{2 \cdot [\mathrm{B}]_{\mathrm{TOT}}}$$

6.4 The King-Altman Model: Bi Bi Ping-Pong Reactions

The primary output of the King Altman method for a Bi Bi Ping-Pong reaction is given below.

$$\mathrm{v_0} = \frac{\mathrm{c_1AB + c_2PQ}}{\mathrm{c_aA + c_bB + c_pP + c_qQ + c_{ab}AB + c_{ap}AP + c_{pq}PQ + c_{bq}BQ}}$$

In the derivation of the steady state velocity equation, one can assume that the two products P and Q have negligible concentration during the early stages of the reaction. Thus, the equation simplifies to…

$$\mathrm{v_0} = \frac{\mathrm{c_1AB}}{\mathrm{c_aA + c_bB + c_{ab}AB}}$$

$$\mathrm{c_{ap}AP} = 0$$

$$\mathrm{c_pP} = 0$$

$$\mathrm{c_{ab}} \rightarrow 1$$

However, for a mutase enzyme, the first product is also the intermediate which has a non-zero concentration. Thus, many of the terms which were discarded for the bi bi ping-pong mechanism, must be retained for the derivation of a steady state velocity equation of a mutase enzyme.

$$\mathrm{P = B = I}$$

$$\mathrm{Q} \rightarrow \mathrm{P}$$

$$\mathrm{v_0} = \frac{\mathrm{c_1AI + c_2IP}}{\mathrm{c_aA + c_bI + c_pI + c_qP + c_{ab}AI + c_{ap}AI + c_{pq}IP + c_{bq}IP}}$$

$$\mathrm{P} = 0$$

$$\mathrm{v_0} = \frac{\mathrm{c_1AI}}{\mathrm{c_aA + (c_b + c_p)I + (1 + c_{ap})AI}}$$

$$\mathrm{c_{ap}AI} \neq 0$$

$$\mathrm{c_pI} \neq 0$$

$$\mathrm{c_{ab}} \rightarrow 1$$

6.5 Linear Discriminant Analysis

A multivariate gaussian distribution is defined by a mean and variance for each variable (μ_i, σ_i) and a parameter defining the correlation between each pair of variables (ρ_{ij}). The general covariance matrix (Σ) for a distribution with 3-variables is stated below.

$$\Sigma = \begin{matrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_1\sigma_2 & \sigma_3^2 \end{matrix}$$

Linear discriminant analysis (LDA) is a classification method. Each of k classes in an N dimensional dataset are modelled using a multivariate gaussian distribution. All classes are assumed to have an identical covariance matrix. New data points can be classified using Bayes' Theorem. This is equivalent to finding the class which minimises the Mahalanobis distance, i.e. the distance from the class mean, measured in standard deviations.

Dimensionality reduction is implicitly associated with this method. Consider two gaussian distributions in 3D space, whose covariances have been normalised such that each distribution appears spherical. A single dimension can be defined by drawing a line connecting the centres of each distribution. Classification of this data, expressed in a single dimension, is just as effective as classification of the 3D dataset. Similarly, a 2D plane is defined by a dataset with 3 classes, and a 3D hyperplane is defined by a dataset with 4 classes. In general, an ND dataset with k classes can be effectively classified in a subspace with k-1 dimensions. This subspace which maximises the separability between the classes within the dataset. This is achieved by maximising the between-class variance and minimising the within-class variance. LDA is similar to principal component analysis (PCA), which instead finds linear combinations that maximises the total variance within a dataset.

7 Supplementary Data

7.1 Euler Angles by simulation

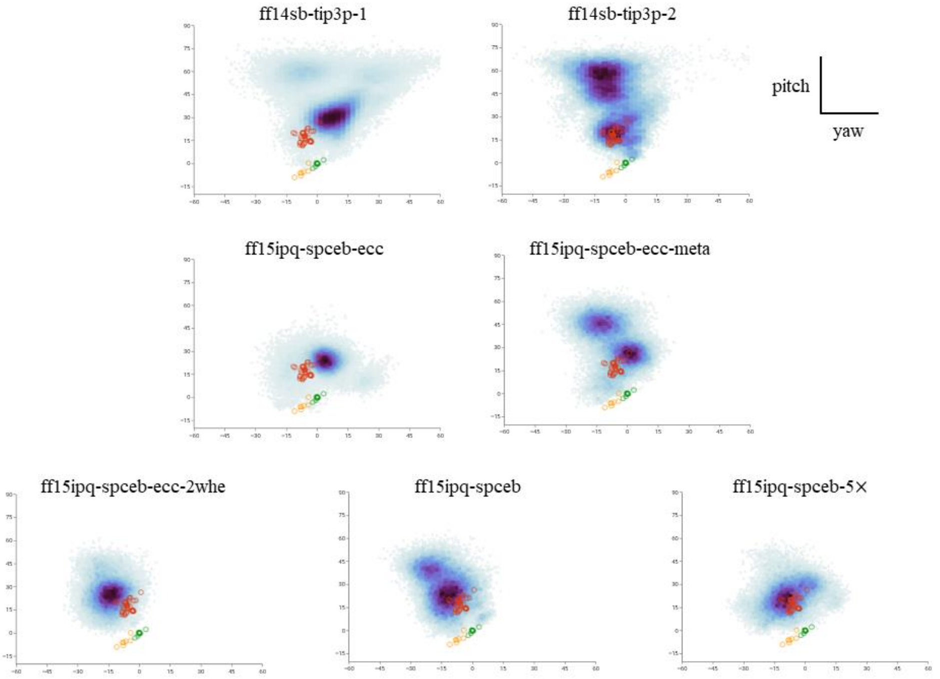


Figure 7-1 – Cap domain Euler angles from βPGM MD simulations. The yaw and pitch angles calculated from each of MD simulations specified in **Figure 2-3** are shown separately. Angles calculated for each of the crystal structures are also shown in each plot as open circles. The TSA, NAC and Open structures are coloured green, yellow, and red, respectively. The simulation data are coloured by sampling density (light blue – dark purple). All subplots have the same x-scale (-60 to 60 °) and y-scale (-15 to 90°)

7.2 Y19 sidechain trajectory

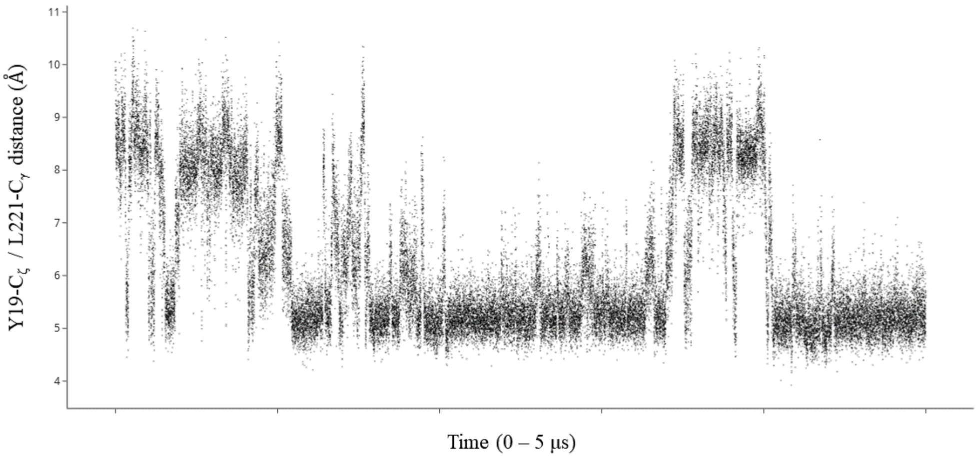


Figure 7-2 – Y19 occupancy of hinge cavity. Trajectory showing the distance between Y19-C$_\zeta$ and L221-C$_\gamma$ during the ff15ipq-speceb-ecc-2whe Molecular Dynamics simulation. This distance acts as a reporter of occupancy of the cavity found within the hinge region in βPGM, by the sidechain of Y19. The simulation quickly tends towards a conformation which places Y19 sidechain into the cavity.

7.3 P146 loop convergence

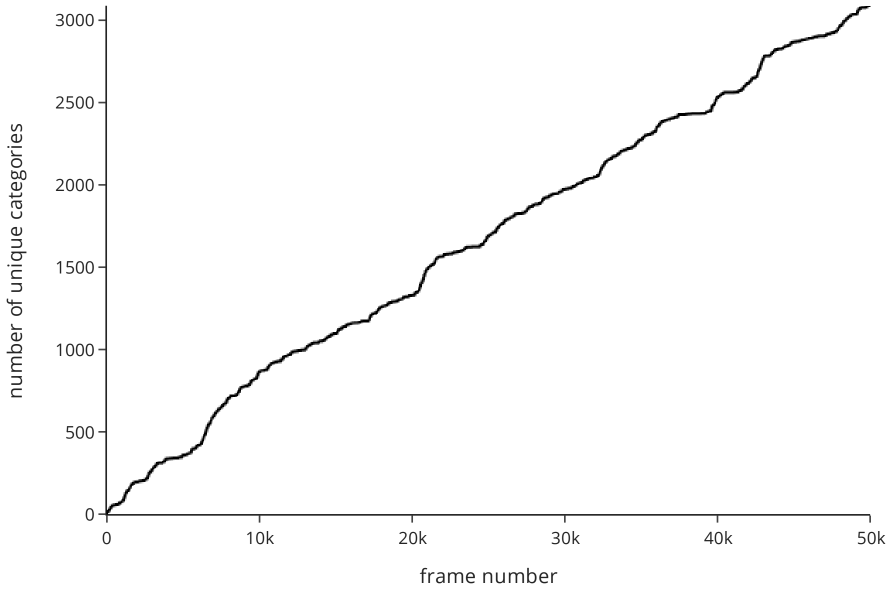


Figure 7-3 – Conformational convergence in the 140s loop under a metadynamics bias. Each frame of the ff15ipq-spceb-ecc-meta simulation was categorised according to the backbone torsion angles of residues P138-P148. The cumulative structural category count is plotted throughout the trajectory.

7.4 Hierarchical clustering of βPGM crystal structures

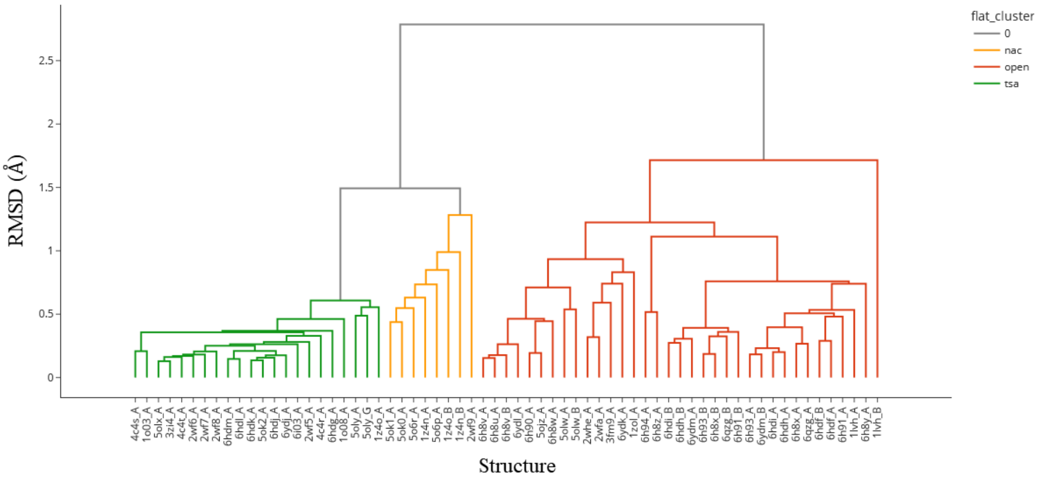


Figure 7-4 – Hierarchical clustering of βPGM crystal structures. Pairwise, average-linkage, agglomerative clustering, was calculated using the backbone-RMSD as a distance metric (MaxCluster, Alex Herbert, the Structural Bioinformatics Group, Imperial College, London, http://www.sbg.bio.ic.ac.uk/~maxcluster/). The results are shown as a dendrogram. Three flat clusters can be defined, corresponding to the TSA structures (green), NACs (yellow) and open structures (red).

7.5 Crystal Structure Euler Angles

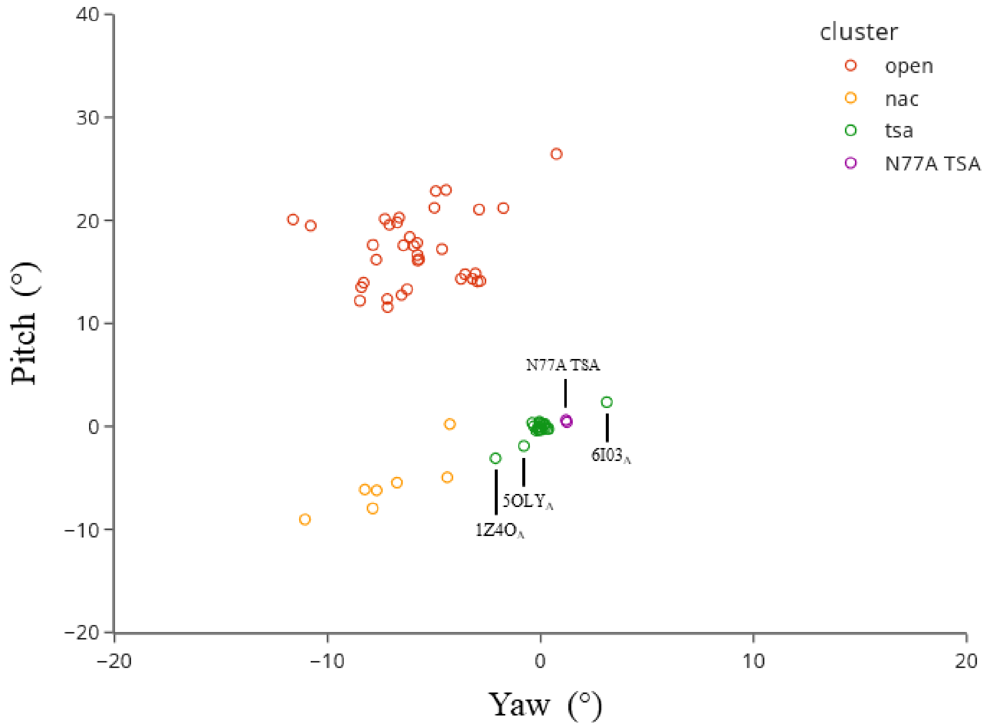


Figure 7-5 – βPGM$_{N77A}$ TSA crystal structures have distorted cap domain Euler Angles. A scatter plot of the Pitch and Yaw Euler angles describing the orientation of the cap domain for each of the βPGM crystal structures. The open structures are coloured red, the NACs are coloured yellow, the TSAs are coloured green and the βPGM$_{N77A}$:MgF$_3$:G6P and βPGM$_{N77A}$:AlF$_4$:G6P TSA complexes are coloured purple. There are five outliers amongst the TSA complexes: **1.** a βPGM$_{WT}$:α-Gal1P complex (1Z4O$_A$) **2.** βPGM$_{5FW}$:MgF$_3$:G6P (5OLY$_A$) **3.** βPGM$_{D10N}$:AlF$_4$:G6P (6I03$_A$) **4.** βPGM$_{N77A}$:AlF$_4$:G6P **5.** βPGM$_{N77A}$:MgF$_3$:G6P.

7.6 Hydrogen bond lengths in βPGM TSA crystal structures

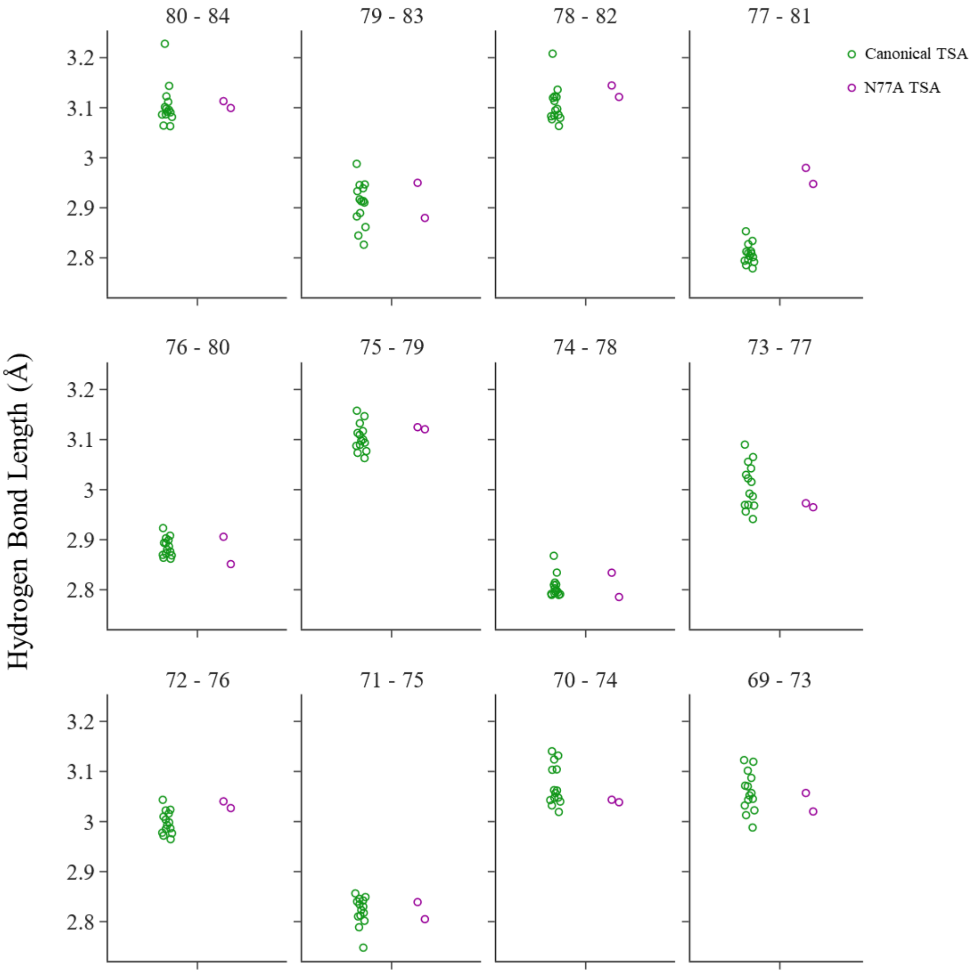


Figure 7-6 – βPGM$_{\text{N77A}}$ TSA hydrogen bonding in the 70s Helix. Hydrogen bond lengths in the 70s helix in βPGM TSA crystal structures. Each subplot represents a backbone helical hydrogen bond between residue i and i+4 (indicated above each subplot). Hydrogen bond lengths in the canonical TSA crystal structures are coloured green; the N77A TSA structures are coloured purple. The hydrogen bond between V81-N and N77-O is elongated in the N77A TSA structures.

7.7 Structure of the 70s helix

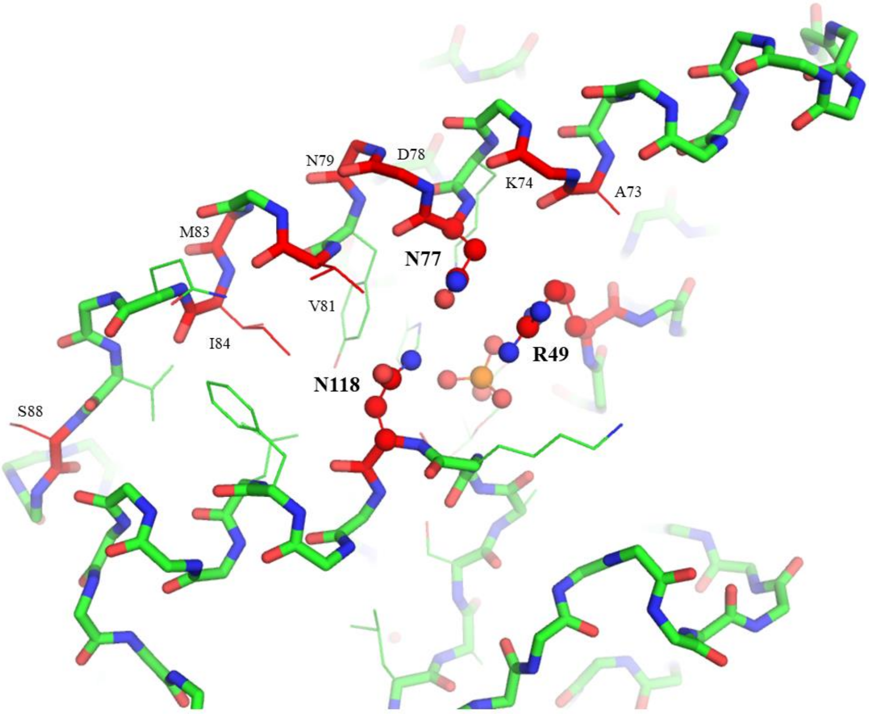


Figure 7-7 – Residues in α4 which are sensitive to closure. The crystal structure of βPGM$_{WT}$ substrate free complex (2WHE). Backbone atoms are represented as sticks and sidechain atoms are represented as lines. Residues backbone ^{15}N-^{1}H chemical shifts which are sensitive to domain closure (S88, I84, M83, V81, N79, N77, K84, A83) are highlighted in red and annotated.

7.8 Crystal Structure Processing and Refinement Statistics

Data collection and data processing statistics		
Complex	**$\beta PGM_{N77A}:MgF_3:G6P$**	**$\beta PGM_{N77A}:AlF_4:G6P$**
Beamline, Facility	i03, DLS	i03, DLS
Space Group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
(a, b, c)	37.49, 54.22, 105.19	37.11, 54.26, 105.41
(alpha, beta, gamma)	90.00, 90.00, 90.00	90.00 90.00 90.00
Resolution (Å)	1.50 - 1.45	1.49 - 1.44
Rmerge	0.632	0.538
Rpim	0.185	0.151
CC-half	0.921	0.939
I/σ (I)	1	2.2
Completeness (%)	100%	100%
Multiplicity	12.26	13.64
Total reflections	42545	59354
Unique Reflections	3471	4351
Molecular Replacement Model	PDB 2WF5	PDB 2WF5
Refinement Statistics		
R (%) / Rfree (%)	18.7 / 21.8	19.2 / 21.3
Number of Atoms		
Protein	1686	1686
Ligands	20	21
Metal Ions	2	2
Water	255	260
Protein Residues	219	219
RMS deviations		
Bonds (A)	0.01	0.01
Angles	1.6	1.5
Average B factors		
Main chains	16.2	15.1
Side chains	25.1	18.9
Ligands	12.9	13.7
Metal Ions	16.7	15.5
Water	25.9	28.2
Ramachandran Analysis		
Favoured / allowed (%)	98.2	97.8
Disallowed	0	0
MolProbity score (percentile)	100 (0.61)	100 (0.69)

** Statistics are given prior to the final refinement cycle*

7.9 βPGM NMR resonance assignments

Variant	Ligands	Species	Category
WT	-	A	OPEN
WT	-	B	OPEN
P146A	-	B	OPEN
D10N	-	A	OPEN
D10N	-	B	OPEN
D10N/P146A	-	B	OPEN
WT	BeF_3:G6P	-	NAC
D10N	βG16BP	-	NAC
D10N	βG16BP:Mg	-	NAC
D10N	F16BP	-	NAC
D10N/P146A	F16BP	-	NAC
D10N/P146A	F16BP:Mg	-	NAC
D10N/P146A	βG16BP	α	NAC
D10N/P146A	βG16BP	β	NAC
D10N/P146A	βG16BP	γ	NAC
D10N/P146A	AlF_4:G6P	-	NAC
K145A	MgF_2:G6P	-	NAC
WT	AlF_4:G6P	-	TSA
WT	MgF_3:G6P	-	TSA
K145A	MgF_3:G6P		TSA
P146A	MgF_3:G6P	-	TSA
D10N/P146A	AlF_4:G6P		TSA

Table 7-1 – Summary of βPGM NMR resonance assignments. The first column describes the enzyme variant (WT, P146A, D10N, K145A or a D10N/P146A double mutant. The second column describes the bound ligands: Trifluoromagnisate (MgF_3), Tetrafluoroaluminate (AlF_4), Glucose-6-phosphate (G6P), Glucose-1,6-bisphosphate (βG16BP) and Fructose-1,6-bisphosphate (F16BP). Some assignments contain multiple species. The open complexes may exist in slow exchange between two conformations (A and B) corresponding to isomerisation of P146. The $\beta PGM_{D10N/P146A}$: βG16BP complex is comprised of multiple species (α, β, γ). Each of the assignments can be classified as either open, NAC or TSA.

7.10 ShiftX2 predictions

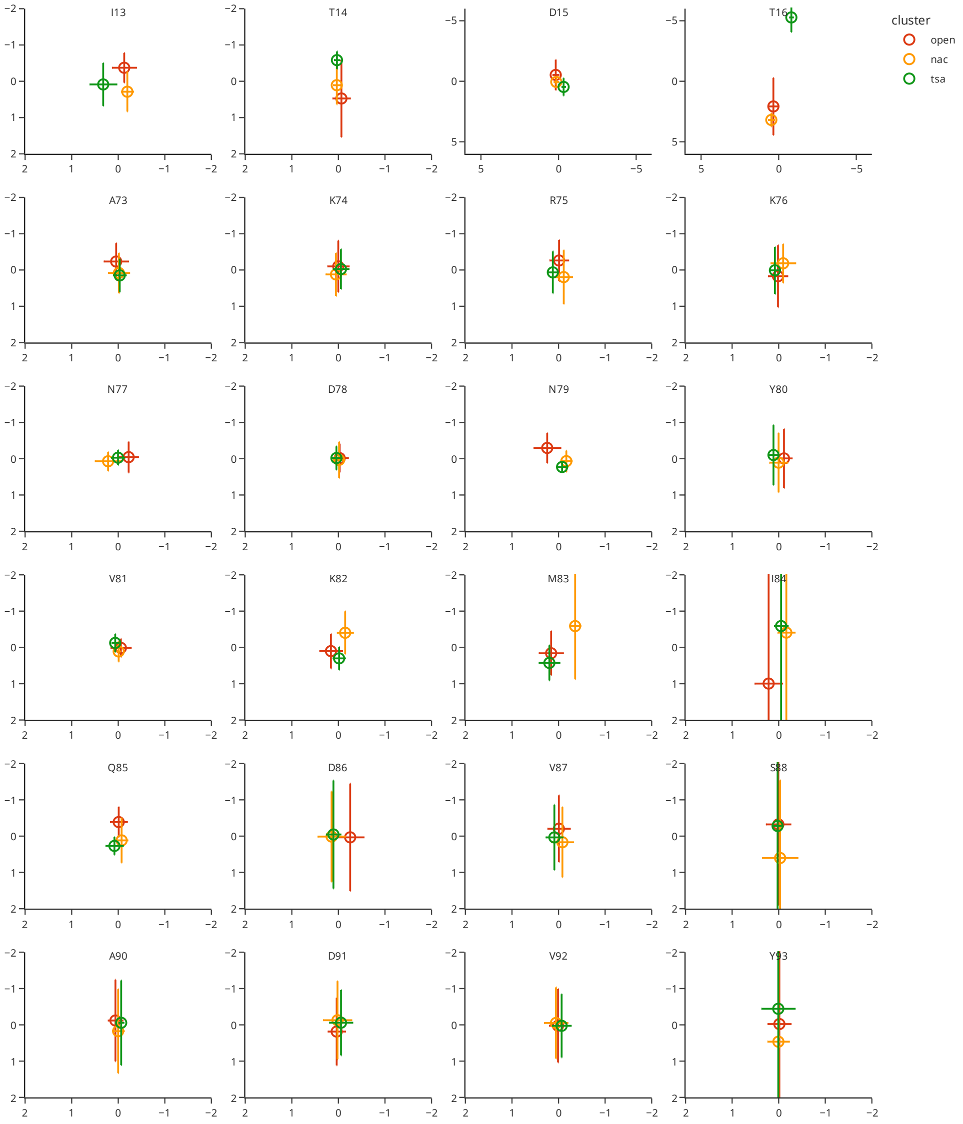


Figure 7-8 – ShiftX2 chemical shift predictions. The relative (centred), mean chemical shift predicted for the TSA (green) NAC (yellow) and Open (red) crystal structures, using ShiftX2. The error bars represent the standard deviation of each distribution.

7.11 Conserved Residues

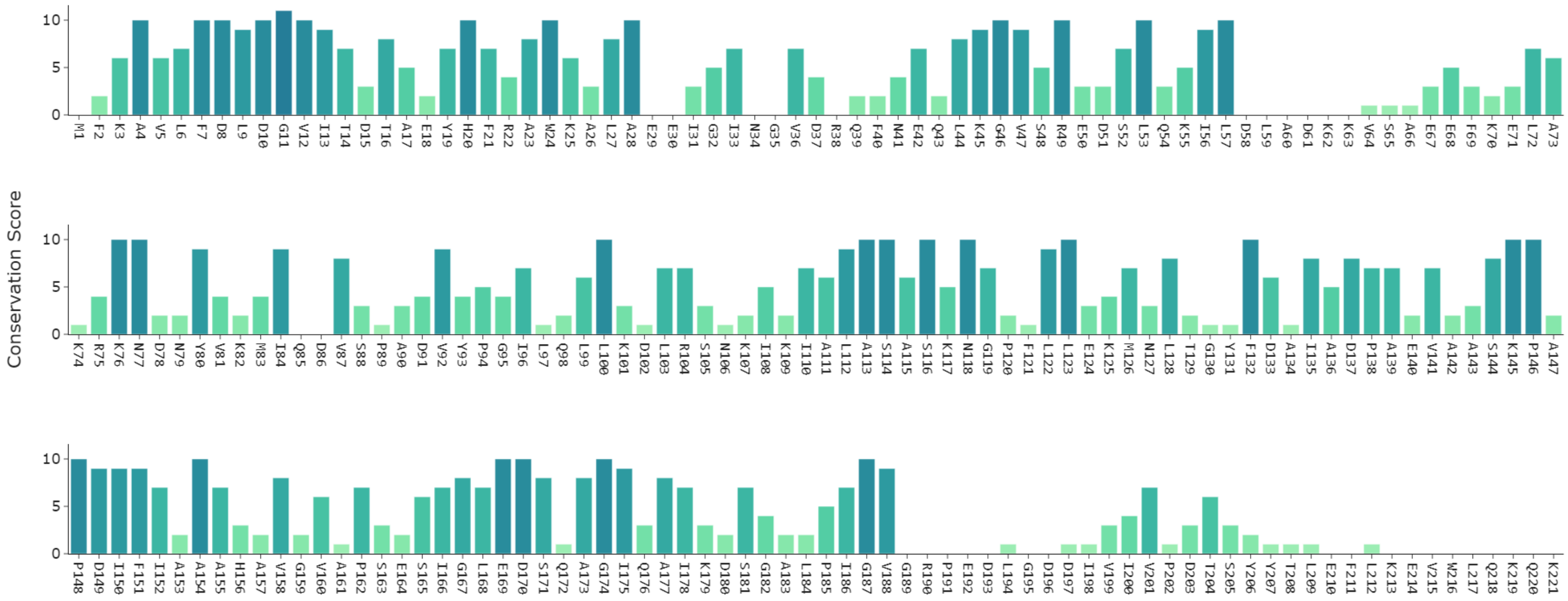


Figure 7-9 – βPGM Residue Conservation Scores. Conservation scores calculated using Clustal Omega for 2982 pgmB protein sequences (Section 7.12). Conservation scores account for residue identity and physiochemical properties. The N77 – N118 pairing is highly conserved amongst these sequences. Y19 has a moderate conservation score .

7.12 List of pgmB sequences

tr|A0A3G9K8E0|A0A3G9K8E0_9ACTN/1-234
tr|A0A0J5PC30|A0A0J5PC30_9LACO/1-200
tr|A0A4Z0Y7Q9|A0A4Z0Y7Q9_9FIRM/1-221
tr|A0A2Z6T6U8|A0A2Z6T6U8_9LACO/1-225
tr|A0A2Z6PRX2|A0A2Z6PRX2_9LACO/1-226
tr|A0A916QJY6|A0A916QJY6_9LACO/1-226
tr|A0A401IR42|A0A401IR42_9LACO/1-216
tr|A0A829ZLK7|A0A829ZLK7_9LACT/1-218
tr|A0A6F9XSA5|A0A6F9XSA5_9LACO/1-215
tr|A0A6F9Y6Z9|A0A6F9Y6Z9_9LACO/1-215
tr|A0A6F9XMN8|A0A6F9XMN8_9LACO/1-215
tr|A0A6F9YN37|A0A6F9YN37_9LACO/1-215
tr|A0A2X4N6F5|A0A2X4N6F5_9BACL/1-221
tr|A0A1C6F9Q9|A0A1C6F9Q9_9FIRM/1-219
tr|A0A1W6NXR0|A0A1W6NXR0_9RHOB/1-215
tr|A0A6N3B184|A0A6N3B184_9CLOT/1-208
tr|A0A2K8SEY3|A0A2K8SEY3_9MOLU/1-222
tr|A0A401IRQ1|A0A401IRQ1_9LACO/1-211
tr|A0A1A5VGR8|A0A1A5VGR8_PEDAC/1-204
tr|A0A1Y0VS70|A0A1Y0VS70_PEDPE/1-218
tr|A0A2I8AJS6|A0A2I8AJS6_9STRE/1-214
tr|A0A2R5HF31|A0A2R5HF31_9LACT/1-216
tr|A0A6F9XIH4|A0A6F9XIH4_9LACO/1-221
tr|A0A6F9XXM9|A0A6F9XXM9_9LACO/1-221
tr|A0A6F9XUD1|A0A6F9XUD1_9LACO/1-221
tr|A0A6F9Y619|A0A6F9Y619_9LACO/1-221
tr|A0A1J0LLQ6|A0A1J0LLQ6_9FLAO/1-219
tr|A0A383TZH7|A0A383TZH7_9FLAO/1-221
tr|A0A7L4ZF57|A0A7L4ZF57_9FLAO/1-222
tr|A0A1C5MFB7|A0A1C5MFB7_9FIRM/1-217
tr|A0A1C5SQN9|A0A1C5SQN9_9FIRM/1-217
tr|A0A239T8L7|A0A239T8L7_9FIRM/1-218
tr|A0A378NW23|A0A378NW23_9FIRM/1-195
tr|A0A6M1ZVR6|A0A6M1ZVR6_9BACT/1-227
tr|A0A6M2ADA1|A0A6M2ADA1_9BACT/1-227
tr|Q4JCF8|Q4JCF8_SULAC/1-216
tr|A0A822IUD2|A0A822IUD2_9EURY/1-217
tr|A0A822JFY1|A0A822JFY1_9EURY/1-210
tr|A0A388TMN1|A0A388TMN1_9BACT/1-216
tr|Q0W6Q6|Q0W6Q6_METAR/1-237
tr|Q0W893|Q0W893_METAR/1-238
tr|A0A136M128|A0A136M128_9BACT/1-202
tr|A0A1V5XTY8|A0A1V5XTY8_9BACT/1-204
tr|A0A3G8JJH4|A0A3G8JJH4_9ACTN/1-236
tr|A0A222TGK6|A0A222TGK6_GORRU/1-234
tr|A0A449GWL0|A0A449GWL0_9ACTN/1-234
tr|A0A949VV92|A0A949VV92_9BACT/1-215
tr|A0A5P9IWF0|A0A5P9IWF0_9GAMM/1-218
tr|A0A5Q0UHM2|A0A5Q0UHM2_9ARCH/1-215
tr|A0A1Y2MNG1|A0A1Y2MNG1_STRPT/1-239
tr|A0A1K2FW38|A0A1K2FW38_9ACTN/1-227
tr|A0A2P8A305|A0A2P8A305_9ACTN/1-236
tr|A0A518HAM3|A0A518HAM3_9BACT/1-219
tr|A0A142YGE0|A0A142YGE0_9PLAN/1-222
tr|A0A5B9W9K1|A0A5B9W9K1_9BACT/1-225
tr|A0A0P8A0W1|A0A0P8A0W1_9CYAN/1-238
tr|V5V6N0|V5V6N0_9CYAN/1-237
tr|A0A916G4Q8|A0A916G4Q8_9BACT/1-227
tr|A0A9E2TMW6|A0A9E2TMW6_UNCNT/1-235
tr|A0A379MPD2|A0A379MPD2_9BACT/1-210
tr|A0A173TDJ0|A0A173TDJ0_9FIRM/1-218
tr|A0A143X505|A0A143X505_9FIRM/1-217
tr|A0A174BZS0|A0A174BZS0_ANAHA/1-219
tr|A0A1C5YX23|A0A1C5YX23_9FIRM/1-219
tr|A0A6N2U7D8|A0A6N2U7D8_ANAHA/1-219
tr|A0A173TWM4|A0A173TWM4_ANAHA/1-219
tr|D4N1C3|D4N1C3_ANAHA/1-219
tr|A0A1V5L8V8|A0A1V5L8V8_9PROT/1-221
tr|A0A7Z3QE47|A0A7Z3QE47_LACPN/1-226
tr|H5T034|H5T034_LACLL/1-59
tr|A0A0P0GKK0|A0A0P0GKK0_9BACE/1-215
tr|A0A1C6I921|A0A1C6I921_9BACE/1-215
tr|A0A120A1H9|A0A120A1H9_BACSE/1-231
tr|A0A380YRK6|A0A380YRK6_9BACE/1-215
tr|A0A0K6BU29|A0A0K6BU29_BACFG/1-220
tr|A0A0P0EUD4|A0A0P0EUD4_BACT4/1-215
tr|A0A1C5TEM9|A0A1C5TEM9_9BACE/1-216
tr|A0A379EEA2|A0A379EEA2_9BACT/1-207
tr|A0A3S4X1L5|A0A3S4X1L5_9BACT/1-215
tr|A0A1B1NUT7|A0A1B1NUT7_9VIBR/1-202
tr|A0A1C7FF42|A0A1C7FF42_9VIBR/1-202
tr|A0A1E3WGE0|A0A1E3WGE0_9VIBR/1-202
tr|A0A5E7IEC4|A0A5E7IEC4_PSEFL/1-228
tr|A0A1V6D0Z1|A0A1V6D0Z1_9CHLR/1-219
tr|A0A6M1ZRH9|A0A6M1ZRH9_UNCCA/1-221
tr|A0A1V5QGY6|A0A1V5QGY6_9BACT/1-220
tr|A0A5E6MEG4|A0A5E6MEG4_9BACT/1-204
tr|A0A1Y5KW46|A0A1Y5KW46_PSEPU/1-227
tr|A0A1V5FIS5|A0A1V5FIS5_9BACT/1-234
tr|M7N2Y9|M7N2Y9_9BACT/1-186
tr|A0A379MSL9|A0A379MSL9_9BACT/1-221
tr|A0A0P8A1V3|A0A0P8A1V3_9BACT/1-215
tr|A0A0R3K0L2|A0A0R3K0L2_CALMK/1-211
tr|A0A369QL40|A0A369QL40_9BACT/1-229
tr|A0A257KBN2|A0A257KBN2_9FLAO/1-218
tr|A0A519KHJ6|A0A519KHJ6_FLASP/1-218
tr|A0A519K0K8|A0A519K0K8_FLASP/1-218
tr|A0A519MYP1|A0A519MYP1_FLASP/1-218
tr|A0A4Q6CM83|A0A4Q6CM83_9PROT/1-218
tr|A0A519NCU7|A0A519NCU7_FLASP/1-218
tr|A0A7Y8Y1I8|A0A7Y8Y1I8_9FLAO/1-218
tr|A0A4Q3W5N2|A0A4Q3W5N2_9BACT/1-205
tr|A0A972FYG8|A0A972FYG8_9FLAO/1-218
tr|A0A345HC31|A0A345HC31_9FLAO/1-218
tr|A0A2E4P6A0|A0A2E4P6A0_FLASP/1-218
tr|A0A255YUW3|A0A255YUW3_9FLAO/1-218
tr|A0A3S3TY70|A0A3S3TY70_9FLAO/1-218
tr|A0A519MN45|A0A519MN45_FLASP/1-218
tr|A0A519NLJ3|A0A519NLJ3_FLASP/1-218
tr|A0A964W6Y5|A0A964W6Y5_9FLAO/1-218
tr|A0A5B9FNA4|A0A5B9FNA4_9FLAO/1-218
tr|A0A6N8HG66|A0A6N8HG66_9FLAO/1-218
tr|A0A2S1QWN5|A0A2S1QWN5_9FLAO/1-218
tr|A0A552UVW3|A0A552UVW3_9FLAO/1-217

tr|A0A4Q3CFG8|A0A4Q3CFG8_9SPHI/1-203
tr|A0A6I4IKR3|A0A6I4IKR3_9FLAO/1-218
tr|A0A2E4PLB4|A0A2E4PLB4_FLASP/1-218
tr|A0A4V1N2L4|A0A4V1N2L4_9FLAO/1-217
tr|A0A4Q1L0B3|A0A4Q1L0B3_9FLAO/1-218
tr|A0A354D8U3|A0A354D8U3_FLASP/1-218
tr|A0A2S0RDG7|A0A2S0RDG7_9FLAO/1-218
tr|A0A257K4Q2|A0A257K4Q2_9FLAO/1-218
tr|A0A2S1SGE2|A0A2S1SGE2_9FLAO/1-218
tr|A0A3D1DAC7|A0A3D1DAC7_9FLAO/1-227
tr|A0A359DWD9|A0A359DWD9_9FLAO/1-218
tr|A0A3R8TMB3|A0A3R8TMB3_9FLAO/1-218
tr|A0A7H9DNV3|A0A7H9DNV3_9FLAO/1-218
tr|A0A348XZJ2|A0A348XZJ2_9FLAO/1-218
tr|A0A8J7K4K7|A0A8J7K4K7_9FLAO/1-218
tr|A0A944G502|A0A944G502_9FLAO/1-218
tr|A0A3L9MII2|A0A3L9MII2_9FLAO/1-218
tr|A0A5P2GHD3|A0A5P2GHD3_9FLAO/1-218
tr|A0A255ZL95|A0A255ZL95_9FLAO/1-221
tr|A0A924I7R1|A0A924I7R1_FLASP/1-217
tr|A0A924LVV4|A0A924LVV4_FLASP/1-216
tr|A0A7Y3R6Z8|A0A7Y3R6Z8_9FLAO/1-217
tr|A0A3S0EP15|A0A3S0EP15_9FLAO/1-216
tr|A0A352RC73|A0A352RC73_FLASP/1-216
tr|A0A9D9QWA7|A0A9D9QWA7_FLASP/1-218
tr|A0A2H3KXI3|A0A2H3KXI3_9FLAO/1-217
tr|A0A3D1N9Q6|A0A3D1N9Q6_FLASP/1-218
tr|A0A9D9QWS9|A0A9D9QWS9_FLASP/1-218
tr|A0A7J5AF38|A0A7J5AF38_9FLAO/1-219
tr|A0A3P3WID9|A0A3P3WID9_9FLAO/1-217
tr|A0A501Q6M8|A0A501Q6M8_9FLAO/1-218
tr|A0A924H734|A0A924H734_FLASP/1-218
tr|A0A2M8JEF5|A0A2M8JEF5_9FLAO/1-217
tr|A0A2W4X5X4|A0A2W4X5X4_9FLAO/1-217
tr|A0A3D4BFY5|A0A3D4BFY5_FLASP/1-217
tr|A0A2W7UGC0|A0A2W7UGC0_9FLAO/1-219
tr|A0A6M8S9L9|A0A6M8S9L9_9FLAO/1-219
tr|A0A553DRD2|A0A553DRD2_9FLAO/1-219
tr|A0A2U1JPZ5|A0A2U1JPZ5_9FLAO/1-219
tr|A0A3S2U5X7|A0A3S2U5X7_9FLAO/1-223
tr|A0A2U1JNR1|A0A2U1JNR1_9FLAO/1-219
tr|A0A3G2GL46|A0A3G2GL46_9FLAO/1-218
tr|A0A519M3W8|A0A519M3W8_FLASP/1-221
tr|A0A363N520|A0A363N520_9FLAO/1-219
tr|A0A7K1FVH4|A0A7K1FVH4_9FLAO/1-219
tr|A0A2U1QXF9|A0A2U1QXF9_9FLAO/1-219
tr|A0A521CDY5|A0A521CDY5_9FLAO/1-218
tr|A0A2S1YJ01|A0A2S1YJ01_9FLAO/1-218
tr|A0A7H8WAT2|A0A7H8WAT2_9FLAO/1-218
tr|A0A7K1WRL5|A0A7K1WRL5_9FLAO/1-217
tr|A0A2V4C2T5|A0A2V4C2T5_9FLAO/1-216
tr|A0A6I4NT91|A0A6I4NT91_9FLAO/1-218
tr|A0A6I3ML35|A0A6I3ML35_9FLAO/1-219
tr|A0A6I3LQ29|A0A6I3LQ29_9FLAO/1-217
tr|A0A5Q3QPW3|A0A5Q3QPW3_9FLAO/1-217
tr|A0A7L8UKS0|A0A7L8UKS0_9FLAO/1-217
tr|A0A4Q1FC80|A0A4Q1FC80_9FLAO/1-217
tr|A0A344LQS4|A0A344LQS4_9FLAO/1-217
tr|A0A4Q1FEV1|A0A4Q1FEV1_9FLAO/1-217
tr|A0A7L4TFM2|A0A7L4TFM2_9FLAO/1-217
tr|A0A941AXT6|A0A941AXT6_9FLAO/1-217
tr|A0A2R7LUK4|A0A2R7LUK4_9FLAO/1-218
tr|A0A1S8ZY36|A0A1S8ZY36_9FLAO/1-217
tr|A0A502E6Q5|A0A502E6Q5_9FLAO/1-218
tr|A0A7Y3X498|A0A7Y3X498_9FLAO/1-217
tr|A0A4Y7UJC8|A0A4Y7UJC8_9FLAO/1-218
tr|A0A434A9Y2|A0A434A9Y2_9FLAO/1-217
tr|A0A7K0F5E4|A0A7K0F5E4_9FLAO/1-217
tr|A0A553E3E1|A0A553E3E1_9FLAO/1-219
tr|A0A4R5CNJ7|A0A4R5CNJ7_9FLAO/1-218
tr|A0A6M2CB02|A0A6M2CB02_9FLAO/1-218
tr|A0A4R5CNW6|A0A4R5CNW6_9FLAO/1-219
tr|A0A930UF65|A0A930UF65_9FLAO/1-218
tr|A0A349PZS2|A0A349PZS2_FLASP/1-217
tr|A0A2S5AD69|A0A2S5AD69_9FLAO/1-217
tr|A0A553CKL0|A0A553CKL0_9FLAO/1-218
tr|A0A4P6YDS3|A0A4P6YDS3_9FLAO/1-217
tr|A0A7K3MTU5|A0A7K3MTU5_FLASP/1-217
tr|A0A519RRX8|A0A519RRX8_FLASP/1-217
tr|A0A923MYJ6|A0A923MYJ6_9FLAO/1-219
tr|A0A9E0HNY7|A0A9E0HNY7_FLASP/1-218
tr|A0A4R5FAI6|A0A4R5FAI6_9FLAO/1-218
tr|A0A4R5FRK9|A0A4R5FRK9_9FLAO/1-218
tr|A0A553BU95|A0A553BU95_9FLAO/1-218
tr|A0A4R5CL79|A0A4R5CL79_9FLAO/1-218
tr|A0A4V2YU97|A0A4V2YU97_9FLAO/1-217
tr|A0A482U4G1|A0A482U4G1_9FLAO/1-217
tr|A0A6G7CAX8|A0A6G7CAX8_9FLAO/1-217
tr|A0A366AZ64|A0A366AZ64_9FLAO/1-218
tr|A0A3S0PKF3|A0A3S0PKF3_9FLAO/1-218
tr|A0A432CDY8|A0A432CDY8_9FLAO/1-218
tr|A0A3M0G0S5|A0A3M0G0S5_9FLAO/1-218
tr|A0A2S1QA92|A0A2S1QA92_9FLAO/1-218
tr|A0A2E4HHN0|A0A2E4HHN0_9BACT/1-218
tr|A0A2E3F224|A0A2E3F224_9FLAO/1-219
tr|A0A2G1VMT0|A0A2G1VMT0_9FLAO/1-218
tr|A0A2E5GM99|A0A2E5GM99_9FLAO/1-218
tr|A0A2E8IDI5|A0A2E8IDI5_9FLAO/1-219
tr|A0A7G2T9Z7|A0A7G2T9Z7_9FLAO/1-219
tr|A0A7C1RZC7|A0A7C1RZC7_9FLAO/1-218
tr|A0A2E2SJW8|A0A2E2SJW8_9BACT/1-219
tr|A0A3D1E4G1|A0A3D1E4G1_9FLAO/1-217
tr|A0A7Y7CKP0|A0A7Y7CKP0_9FLAO/1-219
tr|A0A6M0CGD3|A0A6M0CGD3_9FLAO/1-218
tr|A0A831QQ99|A0A831QQ99_9FLAO/1-217
tr|A0A372GWZ2|A0A372GWZ2_9FLAO/1-220
tr|A0A849F4N9|A0A849F4N9_9FLAO/1-219
tr|A0A849FQ71|A0A849FQ71_9FLAO/1-219
tr|A0A3N4PKN7|A0A3N4PKN7_9FLAO/1-218
tr|A0A5B7TYC5|A0A5B7TYC5_9FLAO/1-219
tr|A0A1G7GZF2|A0A1G7GZF2_9FLAO/1-220
tr|A0A3Q9FKD2|A0A3Q9FKD2_9FLAO/1-218
tr|A0A5R9ASS4|A0A5R9ASS4_9FLAO/1-222
tr|A0A7C1QQG4|A0A7C1QQG4_9FLAO/1-223
tr|A0A3B9R3M8|A0A3B9R3M8_9FLAO/1-220
tr|A0A430K5W4|A0A430K5W4_9FLAO/1-221
tr|A0A848ZTH0|A0A848ZTH0_9FLAO/1-221
tr|A0A520Z5Z5|A0A520Z5Z5_9FLAO/1-218
tr|A0A7Y2ZU13|A0A7Y2ZU13_9FLAO/1-219
tr|A0A2G6QW26|A0A2G6QW26_9FLAO/1-218
tr|A0A6I2MG59|A0A6I2MG59_9FLAO/1-219
tr|A0A223V6G1|A0A223V6G1_9FLAO/1-216
tr|A0A426RN79|A0A426RN79_9FLAO/1-217

```
tr|A0A5B2TZ27|A0A5B2TZ27_9FLAO/1-216
tr|A0A5R8MBE2|A0A5R8MBE2_9FLAO/1-216
tr|A0A3G2LAD2|A0A3G2LAD2_9FLAO/1-217
tr|A0A2E5BIX5|A0A2E5BIX5_9FLAO/1-217
tr|A0A2M7NSW9|A0A2M7NSW9_9FLAO/1-221
tr|A0A9E5H2T7|A0A9E5H2T7_9BACT/1-221
tr|A0A2A5H247|A0A2A5H247_9FLAO/1-219
tr|A0A7H9AQN3|A0A7H9AQN3_9FLAO/1-224
tr|A0A5S3PWS1|A0A5S3PWS1_9FLAO/1-217
tr|A0A5B7SNQ5|A0A5B7SNQ5_9FLAO/1-217
tr|A0A941HLT6|A0A941HLT6_9FLAO/1-216
tr|A0A844NAW2|A0A844NAW2_9FLAO/1-218
tr|A0A7X2ZR99|A0A7X2ZR99_9FLAO/1-220
tr|G0LAH0|G0LAH0_ZOBGA/1-220
tr|A0A7Y7CQG6|A0A7Y7CQG6_9FLAO/1-219
tr|A0A2U2J9H7|A0A2U2J9H7_9FLAO/1-218
tr|A2TWV3|A2TWV3_9FLAO/1-217
tr|A0A5C6XVZ6|A0A5C6XVZ6_9FLAO/1-217
tr|A0A2K8XNY2|A0A2K8XNY2_9FLAO/1-217
tr|A0A5C6Y163|A0A5C6Y163_9FLAO/1-217
tr|A0A5S3N833|A0A5S3N833_9FLAO/1-218
tr|A0A975CMP6|A0A975CMP6_9FLAO/1-218
tr|A0A7G9LDN1|A0A7G9LDN1_9FLAO/1-218
tr|A0A7L8AFT8|A0A7L8AFT8_9FLAO/1-218
tr|A0A958TEZ7|A0A958TEZ7_9FLAO/1-219
tr|A0A7Y2ASS2|A0A7Y2ASS2_9FLAO/1-221
tr|A0A849EN50|A0A849EN50_9FLAO/1-218
tr|A0A6P0UGP4|A0A6P0UGP4_9FLAO/1-218
tr|A0A7Y3CAK3|A0A7Y3CAK3_9FLAO/1-218
tr|A0A946ZV19|A0A946ZV19_9FLAO/1-218
tr|A0A7Y2VP59|A0A7Y2VP59_9FLAO/1-218
tr|A0A7Y1THF6|A0A7Y1THF6_9FLAO/1-218
tr|A0A7Y2UYL1|A0A7Y2UYL1_9FLAO/1-218
tr|A0A411ECA1|A0A411ECA1_9FLAO/1-218
tr|A0A7Y2UKI2|A0A7Y2UKI2_9FLAO/1-218
tr|A0A7Y2MKB2|A0A7Y2MKB2_9FLAO/1-219
tr|A0A842IGM6|A0A842IGM6_9FLAO/1-218
tr|A0A947A2F6|A0A947A2F6_9FLAO/1-218
tr|A0A7Y2YAF2|A0A7Y2YAF2_9FLAO/1-218
tr|A0A849D504|A0A849D504_9FLAO/1-218
tr|A0A3B0CDE1|A0A3B0CDE1_9FLAO/1-224
```

```
tr|A0A7Y3F4D3|A0A7Y3F4D3_9FLAO/1-223
tr|A0A6L9EDN4|A0A6L9EDN4_9FLAO/1-224
tr|A0A524JQE4|A0A524JQE4_9FLAO/1-223
tr|A0A967E6L7|A0A967E6L7_9FLAO/1-218
tr|A0A349P1G1|A0A349P1G1_9FLAO/1-227
tr|A0A7Y1Z6E3|A0A7Y1Z6E3_9FLAO/1-218
tr|A0A2D5TNC4|A0A2D5TNC4_9FLAO/1-218
tr|A0A5C6ZQK9|A0A5C6ZQK9_9FLAO/1-229
tr|A0A9D9FVA1|A0A9D9FVA1_9FLAO/1-232
tr|A0A9E0WPI6|A0A9E0WPI6_9GAMM/1-232
tr|A0A958RIY9|A0A958RIY9_9FLAO/1-230
tr|A0A9E0SCI3|A0A9E0SCI3_9GAMM/1-221
tr|A0A2D5HU82|A0A2D5HU82_9FLAO/1-218
tr|A0A3D5J2U2|A0A3D5J2U2_9FLAO/1-218
tr|A0A2D9DME8|A0A2D9DME8_9FLAO/1-218
tr|A0A2R3Z9X0|A0A2R3Z9X0_9FLAO/1-222
tr|A0M3A5|A0M3A5_GRAFK/1-226
tr|A0A550I2J9|A0A550I2J9_9FLAO/1-227
tr|A0A7M3SX43|A0A7M3SX43_9FLAO/1-221
tr|A0A497CK87|A0A497CK87_9BACT/1-224
tr|A0A929GMN6|A0A929GMN6_9BACT/1-194
tr|A0A5N5ZH62|A0A5N5ZH62_9FLAO/1-218
tr|A0A506PCZ7|A0A506PCZ7_9FLAO/1-218
tr|A0A368MDK6|A0A368MDK6_9FLAO/1-218
tr|A0A3M7KJ87|A0A3M7KJ87_9FLAO/1-224
tr|A0A7Y2NRC0|A0A7Y2NRC0_9FLAO/1-218
tr|A0A7Y3M4Z9|A0A7Y3M4Z9_9FLAO/1-218
tr|A0A6L6U900|A0A6L6U900_9FLAO/1-218
tr|A0A7G8W822|A0A7G8W822_9FLAO/1-218
tr|A0A2D4Y9W8|A0A2D4Y9W8_9FLAO/1-218
tr|A0A958NLX7|A0A958NLX7_9FLAO/1-218
tr|A0A2E2ETQ8|A0A2E2ETQ8_9FLAO/1-218
tr|A0A2E6VAC3|A0A2E6VAC3_9FLAO/1-218
tr|A0A7K1GDB0|A0A7K1GDB0_9FLAO/1-218
tr|A0A842IMW9|A0A842IMW9_9FLAO/1-218
tr|A0A848UMY5|A0A848UMY5_9FLAO/1-218
tr|A0A2T4WKA9|A0A2T4WKA9_9BACT/1-218
tr|A0A7K0BGV4|A0A7K0BGV4_9FLAO/1-218
tr|A0A2D9Y344|A0A2D9Y344_9FLAO/1-224
tr|A0A5B1BIC9|A0A5B1BIC9_9FLAO/1-221
tr|A0A504J507|A0A504J507_9FLAO/1-219
```

```
tr|A0A3A9WFF6|A0A3A9WFF6_9FLAO/1-221
tr|A0A3A9YN55|A0A3A9YN55_9FLAO/1-221
tr|A0A941CCW8|A0A941CCW8_9FLAO/1-217
tr|A0A554VHW9|A0A554VHW9_9FLAO/1-221
tr|A0A3A9VG37|A0A3A9VG37_9FLAO/1-219
tr|A0A936ZQU5|A0A936ZQU5_9FLAO/1-220
tr|A0A7D7VMN9|A0A7D7VMN9_9FLAO/1-227
tr|A0A2M8AIG6|A0A2M8AIG6_9FLAO/1-226
tr|A0A966FLJ8|A0A966FLJ8_9FLAO/1-226
tr|A0A516GPV4|A0A516GPV4_9FLAO/1-218
tr|A0A2I7SEG2|A0A2I7SEG2_9FLAO/1-218
tr|A0A5Q0JRF0|A0A5Q0JRF0_9FLAO/1-226
tr|A0A3M8GBI1|A0A3M8GBI1_9FLAO/1-228
tr|A0A958T6Z4|A0A958T6Z4_9FLAO/1-218
tr|A0A958Y9D9|A0A958Y9D9_9FLAO/1-218
tr|A0A7Y1TQH4|A0A7Y1TQH4_9FLAO/1-218
tr|A0A521A707|A0A521A707_9FLAO/1-218
tr|A0A4R1ADB4|A0A4R1ADB4_9FLAO/1-218
tr|A0A2K1E4D4|A0A2K1E4D4_9FLAO/1-218
tr|A0A7C7NUB3|A0A7C7NUB3_9FLAO/1-218
tr|A0A2T1N7F5|A0A2T1N7F5_9FLAO/1-218
tr|A0A974YTJ1|A0A974YTJ1_9FLAO/1-218
tr|A0A2K8WUH0|A0A2K8WUH0_9FLAO/1-218
tr|A0A2K8XKU5|A0A2K8XKU5_9FLAO/1-218
tr|A0A4P7WWL0|A0A4P7WWL0_9FLAO/1-218
tr|A0A6H3Q0I8|A0A6H3Q0I8_9FLAO/1-218
tr|A0A7J5AM04|A0A7J5AM04_9FLAO/1-217
tr|A0A3R9UWY3|A0A3R9UWY3_9FLAO/1-218
tr|A0A4Q4B8L6|A0A4Q4B8L6_9FLAO/1-218
tr|A0A4Q9RVR2|A0A4Q9RVR2_9FLAO/1-218
tr|A0A958YLD0|A0A958YLD0_9FLAO/1-172
tr|A0A5C7AKA0|A0A5C7AKA0_9FLAO/1-220
tr|A0A934KUA6|A0A934KUA6_9FLAO/1-218
tr|A0A4Q0XFS2|A0A4Q0XFS2_9FLAO/1-218
tr|A0A7W2M6E2|A0A7W2M6E2_9FLAO/1-218
tr|A0A3S9MWN6|A0A3S9MWN6_9FLAO/1-220
tr|L7W7I2|L7W7I2_NONDD/1-220
tr|A0A7L5FCA2|A0A7L5FCA2_9FLAO/1-220
tr|A0A2K8X7S1|A0A2K8X7S1_9FLAO/1-220
tr|A0A7W1ZDP8|A0A7W1ZDP8_9FLAO/1-220
tr|A0A2M7DGJ9|A0A2M7DGJ9_9FLAO/1-217
```

tr|A0A2N2YEL0|A0A2N2YEL0_9BACT/1-217
tr|A0A6B3R576|A0A6B3R576_9FLAO/1-223
tr|A0A2I0D5V6|A0A2I0D5V6_9FLAO/1-219
tr|A0A7C7IKB4|A0A7C7IKB4_9FLAO/1-220
tr|A0A2D8UGI4|A0A2D8UGI4_9FLAO/1-220
tr|A0A967DZE1|A0A967DZE1_9FLAO/1-217
tr|A0A4U5TRC5|A0A4U5TRC5_9FLAO/1-220
tr|A0A426KL65|A0A426KL65_9FLAO/1-220
tr|A0A975FUQ1|A0A975FUQ1_9FLAO/1-220
tr|A0A2U8QSH2|A0A2U8QSH2_9FLAO/1-218
tr|A0A2T1NDX9|A0A2T1NDX9_9FLAO/1-216
tr|A0A8J7J200|A0A8J7J200_9FLAO/1-226
tr|A0A8J6QC96|A0A8J6QC96_9FLAO/1-220
tr|A0A8J6QSW6|A0A8J6QSW6_9FLAO/1-218
tr|A0A8J6UCW4|A0A8J6UCW4_9FLAO/1-218
tr|A0A5M4B843|A0A5M4B843_9FLAO/1-208
tr|A0A250G556|A0A250G556_9FLAO/1-213
tr|A0A3A1YHI4|A0A3A1YHI4_9FLAO/1-213
tr|A0A250G616|A0A250G616_9FLAO/1-210
tr|A0A0B7HEV4|A0A0B7HEV4_9FLAO/1-210
tr|A0A0B7HI86|A0A0B7HI86_9FLAO/1-210
tr|A0A250E5Q3|A0A250E5Q3_9FLAO/1-208
tr|A0A250EPU5|A0A250EPU5_9FLAO/1-208
tr|A0A4S3M4E3|A0A4S3M4E3_9FLAO/1-218
tr|A0A9E6ZXB8|A0A9E6ZXB8_9FLAO/1-214
tr|A0A6L9EZA5|A0A6L9EZA5_9FLAO/1-218
tr|A0A6P0UQ62|A0A6P0UQ62_9FLAO/1-218
tr|A0A918VWB2|A0A918VWB2_9FLAO/1-220
tr|A0A9E8DIJ5|A0A9E8DIJ5_9FLAO/1-219
tr|A0A958M7U1|A0A958M7U1_9FLAO/1-226
tr|A0A5C6ZUK6|A0A5C6ZUK6_9FLAO/1-221
tr|A0A5B7X264|A0A5B7X264_9FLAO/1-221
tr|A0A5B8YJ73|A0A5B8YJ73_9FLAO/1-219
tr|A0A7Y2J7I6|A0A7Y2J7I6_9FLAO/1-218
tr|A0A7Y3DVZ8|A0A7Y3DVZ8_9FLAO/1-218
tr|A0A947EID3|A0A947EID3_9BACT/1-218
tr|A0A3D2XPK3|A0A3D2XPK3_9FLAO/1-218
tr|A0A7Y3A2J1|A0A7Y3A2J1_9FLAO/1-225
tr|A0A7Y3HXH0|A0A7Y3HXH0_9FLAO/1-221
tr|A0A523GDR2|A0A523GDR2_9BACT/1-218
tr|A0A523GWB9|A0A523GWB9_9BACT/1-218
tr|A0A523H102|A0A523H102_9BACT/1-218
tr|A0A2E0YQE6|A0A2E0YQE6_9FLAO/1-218
tr|A0A7Y2Z4W3|A0A7Y2Z4W3_9FLAO/1-218
tr|A0A6H9LDM3|A0A6H9LDM3_FLASP/1-219
tr|A0A2S0HYI9|A0A2S0HYI9_9FLAO/1-219
tr|A0A7Y1ZBT9|A0A7Y1ZBT9_9FLAO/1-218
tr|A0A2D8FXY1|A0A2D8FXY1_9FLAO/1-218
tr|A0A958U460|A0A958U460_9FLAO/1-219
tr|A0A2D9B7Q3|A0A2D9B7Q3_ALTSX/1-206
tr|A0A2E2IYE6|A0A2E2IYE6_9FLAO/1-219
tr|A0A352BS49|A0A352BS49_9FLAO/1-222
tr|A0A3D2R7Y4|A0A3D2R7Y4_9FLAO/1-221
tr|A0A3B9JQZ3|A0A3B9JQZ3_9FLAO/1-216
tr|A0A3B8T677|A0A3B8T677_9FLAO/1-216
tr|A0A348N9H2|A0A348N9H2_9FLAO/1-216
tr|A0A3D1I5J0|A0A3D1I5J0_9FLAO/1-180
tr|A0A926JVM0|A0A926JVM0_9FLAO/1-221
tr|A0A329N3G7|A0A329N3G7_9FLAO/1-212
tr|A0A3N0E7D8|A0A3N0E7D8_9FLAO/1-215
tr|A0A362XLC6|A0A362XLC6_9BACT/1-220
tr|A0A7Y1VM03|A0A7Y1VM03_9FLAO/1-223
tr|A0A946Z8B0|A0A946Z8B0_9BACT/1-220
tr|A0A942NH41|A0A942NH41_9FLAO/1-219
tr|A0A838ZMR9|A0A838ZMR9_9FLAO/1-217
tr|A0A970I5N6|A0A970I5N6_9FLAO/1-219
tr|A0A942L8I1|A0A942L8I1_9BACT/1-218
tr|A0A2A4RM25|A0A2A4RM25_9FLAO/1-218
tr|A0A2M8A843|A0A2M8A843_9FLAO/1-217
tr|A0A2N2ZI16|A0A2N2ZI16_9BACT/1-217
tr|A0A5C7AP33|A0A5C7AP33_9FLAO/1-219
tr|A0A958TR27|A0A958TR27_9FLAO/1-219
tr|A0A2K9PW15|A0A2K9PW15_9FLAO/1-223
tr|A0A848SAZ3|A0A848SAZ3_9FLAO/1-221
tr|A0A5M7BMW5|A0A5M7BMW5_9FLAO/1-223
tr|A0A2U2X6Z4|A0A2U2X6Z4_9FLAO/1-227
tr|A0A5C7GGA4|A0A5C7GGA4_9FLAO/1-219
tr|A0A5D0HF35|A0A5D0HF35_9FLAO/1-217
tr|A0A2N3HJB4|A0A2N3HJB4_9FLAO/1-222
tr|A0A5C4SM99|A0A5C4SM99_9FLAO/1-218
tr|A0A4Q9FMD1|A0A4Q9FMD1_9FLAO/1-232
tr|A0A4V6MT38|A0A4V6MT38_9FLAO/1-232
tr|A0A923HGC4|A0A923HGC4_9FLAO/1-229
tr|A0A2G6F6D1|A0A2G6F6D1_9FLAO/1-218
tr|A0A2G6LP93|A0A2G6LP93_9FLAO/1-212
tr|A0A958UYZ7|A0A958UYZ7_9FLAO/1-179
tr|A0A432ILQ4|A0A432ILQ4_9BACT/1-217
tr|A0A2A4NME0|A0A2A4NME0_9FLAO/1-217
tr|A0A2A4R0S0|A0A2A4R0S0_9FLAO/1-217
tr|A0A2A4TG87|A0A2A4TG87_9FLAO/1-217
tr|A0A662BPB5|A0A662BPB5_9BACT/1-215
tr|A0A2A4M0G7|A0A2A4M0G7_9FLAO/1-218
tr|A0A4Q1IS01|A0A4Q1IS01_9FLAO/1-218
tr|A0A4Y8AP17|A0A4Y8AP17_9FLAO/1-218
tr|A0A958UAX1|A0A958UAX1_9FLAO/1-201
tr|A0A7Y2TK56|A0A7Y2TK56_9FLAO/1-218
tr|A0A7X8N614|A0A7X8N614_9FLAO/1-218
tr|A0A6P1DGY8|A0A6P1DGY8_9FLAO/1-218
tr|A0A3D1KNQ2|A0A3D1KNQ2_9FLAO/1-218
tr|A0A2N2XC14|A0A2N2XC14_9BACT/1-218
tr|A0A3R6NHH0|A0A3R6NHH0_9CLOT/1-212
tr|A0A373T7X5|A0A373T7X5_9FIRM/1-212
tr|A0A373ZZ53|A0A373ZZ53_9FIRM/1-212
tr|A0A374JHY8|A0A374JHY8_9FIRM/1-212
tr|A0A374E1Q5|A0A374E1Q5_9FIRM/1-212
tr|A0A374BYG6|A0A374BYG6_9FIRM/1-212
tr|A0A417BL20|A0A417BL20_9FIRM/1-212
tr|A0A1H9UYP1|A0A1H9UYP1_BUTFI/1-210
tr|A0A9D1F6T2|A0A9D1F6T2_9FIRM/1-211
tr|A0A371JG13|A0A371JG13_9FIRM/1-213
tr|A0A7C6JXH7|A0A7C6JXH7_9FIRM/1-213
tr|A0A373NGQ8|A0A373NGQ8_9FIRM/1-211
tr|A0A8I0AIR0|A0A8I0AIR0_9FIRM/1-211
tr|A0A3R8LWV0|A0A3R8LWV0_9FIRM/1-211
tr|A0A7X5EK17|A0A7X5EK17_9FIRM/1-211
tr|A0A6L9GY28|A0A6L9GY28_9FIRM/1-211
tr|A0A3R6NNR3|A0A3R6NNR3_9CLOT/1-210
tr|A0A3A6ESC5|A0A3A6ESC5_9FIRM/1-210
tr|A0A3E4UWH7|A0A3E4UWH7_RUMGN/1-210
tr|A0A7X2M750|A0A7X2M750_9FIRM/1-209
tr|A0A373NN34|A0A373NN34_9FIRM/1-213
tr|A0A373SX58|A0A373SX58_9FIRM/1-213
tr|A0A373MN89|A0A373MN89_9FIRM/1-213

```
tr|A0A374JHK0|A0A374JHK0_9FIRM/1-213
tr|A0A396PVY6|A0A396PVY6_9FIRM/1-213
tr|A0A374BW90|A0A374BW90_9FIRM/1-213
tr|A0A417BK07|A0A417BK07_9FIRM/1-213
tr|A0A417FLH4|A0A417FLH4_9FIRM/1-213
tr|A0A494ZRM5|A0A494ZRM5_9FIRM/1-213
tr|A0A4Q8TU42|A0A4Q8TU42_9FIRM/1-213
tr|A0A3D0M0A6|A0A3D0M0A6_9FIRM/1-213
tr|A0A1C5XQA8|A0A1C5XQA8_9CLOT/1-216
tr|A0A7X5KT01|A0A7X5KT01_9FIRM/1-216
tr|A0A3D0EF98|A0A3D0EF98_9FIRM/1-216
tr|A0A9D1H331|A0A9D1H331_9FIRM/1-216
tr|A0A9D1MBN2|A0A9D1MBN2_9FIRM/1-215
tr|A0A9D1SEP5|A0A9D1SEP5_9FIRM/1-216
tr|A0A9D8S1J9|A0A9D8S1J9_9FIRM/1-214
tr|A0A3B8S5C8|A0A3B8S5C8_9FIRM/1-216
tr|A0A943KVM1|A0A943KVM1_9FIRM/1-216
tr|C6LHF1|C6LHF1_9FIRM/1-239
tr|A0A6L5YF64|A0A6L5YF64_9FIRM/1-215
tr|A0A970FLT3|A0A970FLT3_9CLOT/1-214
tr|A0A373W2Z1|A0A373W2Z1_9FIRM/1-214
tr|A0A416T3S2|A0A416T3S2_9FIRM/1-214
tr|A0A416W1X2|A0A416W1X2_9FIRM/1-214
tr|A0A939RAE6|A0A939RAE6_9FIRM/1-210
tr|A0A9D2MNV5|A0A9D2MNV5_9FIRM/1-216
tr|A0A954Z2T9|A0A954Z2T9_9BACT/1-220
tr|A0A7C4CAJ3|A0A7C4CAJ3_9BACT/1-221
tr|A0A2Z4WEG2|A0A2Z4WEG2_9CLOT/1-216
tr|A0A353BT66|A0A353BT66_UNCFI/1-214
tr|A0A7V6H5K2|A0A7V6H5K2_9FIRM/1-211
tr|A0A7X8DB42|A0A7X8DB42_9FIRM/1-211
tr|A0A970PNF4|A0A970PNF4_9FIRM/1-220
tr|A0A6G3ZR39|A0A6G3ZR39_9BACL/1-213
tr|A0A927WJG5|A0A927WJG5_SELRU/1-212
tr|A0A7V3L486|A0A7V3L486_UNCAM/1-226
tr|A0A970G438|A0A970G438_9FIRM/1-216
tr|A0A7C6BEX7|A0A7C6BEX7_9FIRM/1-216
tr|A0A971VVU8|A0A971VVU8_9FIRM/1-216
tr|A0A924VWU5|A0A924VWU5_UNCCL/1-218
tr|A0A960M1W2|A0A960M1W2_9BACT/1-214
tr|A0A943G7L2|A0A943G7L2_9FIRM/1-229
tr|A0A928C456|A0A928C456_9BACT/1-216
tr|A0A9D1ZUQ6|A0A9D1ZUQ6_9FIRM/1-216
tr|A0A496Y082|A0A496Y082_UNCDE/1-218
tr|A0A7C5DV82|A0A7C5DV82_9BACT/1-218
tr|A0A968VWE4|A0A968VWE4_9BACT/1-214
tr|A0A6C2UNS3|A0A6C2UNS3_9BACT/1-213
tr|A0A7X8PTC6|A0A7X8PTC6_9CLOT/1-223
tr|A0A9E5S008|A0A9E5S008_9CLOT/1-223
tr|A0A3C1AZS5|A0A3C1AZS5_9FIRM/1-217
tr|A0A971G7I6|A0A971G7I6_9FIRM/1-219
tr|A0A941WPW3|A0A941WPW3_9FIRM/1-216
tr|A0A969YK23|A0A969YK23_9FIRM/1-213
tr|A0A927SSC9|A0A927SSC9_9FIRM/1-215
tr|A0A971AJB6|A0A971AJB6_9FIRM/1-212
tr|A0A2Z2K5F9|A0A2Z2K5F9_9BACL/1-214
tr|A0A348NL83|A0A348NL83_9MOLU/1-211
tr|A0A254NL48|A0A254NL48_9BACL/1-208
tr|A0A970STA3|A0A970STA3_9FIRM/1-215
tr|A0A972B661|A0A972B661_9CLOT/1-239
tr|A0A7X6XDY2|A0A7X6XDY2_9BACT/1-224
tr|A0A349UN88|A0A349UN88_9BACT/1-218
tr|A0A7C6UQC8|A0A7C6UQC8_UNCFI/1-216
tr|A0A943CTK5|A0A943CTK5_UNCFI/1-211
tr|A0A355A495|A0A355A495_UNCFI/1-225
tr|A0A357T0H6|A0A357T0H6_UNCFI/1-225
tr|A0A8J6I2P7|A0A8J6I2P7_9FIRM/1-219
tr|A0A971N855|A0A971N855_UNCFI/1-219
tr|A0A357CMU7|A0A357CMU7_UNCFI/1-218
tr|A0A357D6J1|A0A357D6J1_UNCFI/1-229
tr|A0A7C6CJR1|A0A7C6CJR1_UNCFI/1-218
tr|A0A972D713|A0A972D713_UNCFI/1-218
tr|A0A7X7UQ91|A0A7X7UQ91_9MOLU/1-210
tr|A0A9J6RAX0|A0A9J6RAX0_9BACI/1-217
tr|A0A2N1PZ26|A0A2N1PZ26_9BACT/1-210
tr|A0A975QXM8|A0A975QXM8_9BACT/1-210
tr|A0A9E2GA88|A0A9E2GA88_UNCFI/1-215
tr|A0A2N1QFC7|A0A2N1QFC7_9BACT/1-210
tr|A0A356GLT1|A0A356GLT1_9MOLU/1-210
tr|A0A385NVB4|A0A385NVB4_9BACI/1-215
tr|A0A2C1Z8R5|A0A2C1Z8R5_9BACI/1-220
tr|A0A2P8MHN3|A0A2P8MHN3_9CLOT/1-215
tr|A0A2C1MAU6|A0A2C1MAU6_BACCE/1-215
tr|A0A443THL5|A0A443THL5_BACMY/1-215
tr|A0A3P3UCT1|A0A3P3UCT1_9BACL/1-213
tr|A0A3P3U4D7|A0A3P3U4D7_9BACL/1-213
tr|A0A3P3U4F3|A0A3P3U4F3_9BACL/1-213
tr|A0A090ZBE8|A0A090ZBE8_PAEMA/1-213
tr|A0A090ZE99|A0A090ZE99_PAEMA/1-213
tr|A0A6A8D924|A0A6A8D924_9BACI/1-213
tr|A0A1E3L9Q8|A0A1E3L9Q8_9BACL/1-212
tr|A0A4U3FBI8|A0A4U3FBI8_9BACL/1-212
tr|A0A9D5M4B4|A0A9D5M4B4_9FIRM/1-208
tr|A0A7C4U3E1|A0A7C4U3E1_9BACT/1-185
tr|A0A9D8TQX4|A0A9D8TQX4_UNCCL/1-167
tr|A0A9D9BY49|A0A9D9BY49_9FIRM/1-164
tr|A0A928RQL4|A0A928RQL4_9FIRM/1-211
tr|A0A413I1Z2|A0A413I1Z2_9FIRM/1-210
tr|A0A416BDS6|A0A416BDS6_9FIRM/1-210
tr|A0A3D2Q0R6|A0A3D2Q0R6_9FIRM/1-211
tr|A0A373L5K1|A0A373L5K1_9FIRM/1-211
tr|A0A373UV12|A0A373UV12_9FIRM/1-211
tr|A0A396LLZ0|A0A396LLZ0_9FIRM/1-211
tr|A0A3R6M5B0|A0A3R6M5B0_9FIRM/1-211
tr|A0A3R6MTX1|A0A3R6MTX1_9FIRM/1-211
tr|A0A3R6PNK9|A0A3R6PNK9_9FIRM/1-211
tr|A0A3R6QEH4|A0A3R6QEH4_9FIRM/1-211
tr|A0A416WN28|A0A416WN28_9FIRM/1-211
tr|A0A374BAP7|A0A374BAP7_9FIRM/1-211
tr|A0A2G3E3J3|A0A2G3E3J3_9FIRM/1-216
tr|A0A9D1JBY3|A0A9D1JBY3_9FIRM/1-211
tr|A0A351V7G2|A0A351V7G2_9FIRM/1-210
tr|A0A927YI29|A0A927YI29_9FIRM/1-210
tr|A0A414HWY3|A0A414HWY3_9FIRM/1-212
tr|A0A173UMW3|A0A173UMW3_9FIRM/1-212
tr|A0A396FQT8|A0A396FQT8_9FIRM/1-212
tr|A0A1Q6SEA6|A0A1Q6SEA6_9FIRM/1-212
tr|A0A349QMM1|A0A349QMM1_9FIRM/1-212
tr|A0A3R6DH83|A0A3R6DH83_9FIRM/1-212
tr|A0A844KJS7|A0A844KJS7_9FIRM/1-211
tr|A0A351R0Z3|A0A351R0Z3_9FIRM/1-211
tr|A0A2N1QJ23|A0A2N1QJ23_9BACT/1-211
tr|A0A7X8WJA4|A0A7X8WJA4_9MOLU/1-218
```

```
tr|A0A859DNM8|A0A859DNM8_9FIRM/1-211
tr|A0A7G8TA93|A0A7G8TA93_9FIRM/1-225
tr|A0A524MS92|A0A524MS92_9BACT/1-218
tr|A0A371J4N1|A0A371J4N1_9FIRM/1-215
tr|A0A934RZK9|A0A934RZK9_9BACT/1-220
tr|A0A5C8M3Q8|A0A5C8M3Q8_9BACL/1-214
tr|A0A7X6VRP9|A0A7X6VRP9_9MOLU/1-213
tr|A0A353J950|A0A353J950_UNCFI/1-213
tr|A0A949K1I9|A0A949K1I9_9FIRM/1-225
tr|A0A3R6XFR1|A0A3R6XFR1_9FIRM/1-210
tr|A0A7X5E9R4|A0A7X5E9R4_9FIRM/1-215
tr|A0A6P1MAM1|A0A6P1MAM1_9BACT/1-212
tr|A0A933ZXV6|A0A933ZXV6_UNCDE/1-240
tr|A0A3F3K4W0|A0A3F3K4W0_GARVA/1-215
tr|A0A3E2CJ24|A0A3E2CJ24_GARVA/1-215
tr|A0A2I1KNB1|A0A2I1KNB1_GARVA/1-215
tr|A0A1H1LTJ3|A0A1H1LTJ3_GARVA/1-215
tr|E3D7N8|E3D7N8_GARV3/1-215
tr|F5LX24|F5LX24_GARVA/1-215
tr|A0A3E2D5B5|A0A3E2D5B5_GARVA/1-169
tr|A0A395Y1E3|A0A395Y1E3_BIFLN/1-213
tr|A0A2K9AGM1|A0A2K9AGM1_BIFBR/1-213
tr|A0A2K9B2B4|A0A2K9B2B4_BIFBR/1-213
tr|A0A9D2H8J8|A0A9D2H8J8_9FIRM/1-212
tr|A0A348T1E3|A0A348T1E3_9FIRM/1-212
tr|A0A943YSA9|A0A943YSA9_9FIRM/1-212
tr|A0A9D0ZE62|A0A9D0ZE62_9FIRM/1-213
tr|D3R1S2|D3R1S2_MAGIU/1-211
tr|A0A3C0P0D3|A0A3C0P0D3_9FIRM/1-215
tr|A0A927XF88|A0A927XF88_9STRE/1-214
tr|A0A2T3G162|A0A2T3G162_9FIRM/1-215
tr|A0A6N8H7Z2|A0A6N8H7Z2_9FIRM/1-212
tr|A0A3D2X3V8|A0A3D2X3V8_9FIRM/1-211
tr|E6LRG8|E6LRG8_9FIRM/1-215
tr|A0A496N621|A0A496N621_9FIRM/1-215
tr|I0R477|I0R477_9FIRM/1-215
tr|A0A496NEG7|A0A496NEG7_9FIRM/1-214
tr|J5GJ17|J5GJ17_9FIRM/1-214
tr|A0A3P3Q432|A0A3P3Q432_9FIRM/1-214
tr|A0A929Y1Z3|A0A929Y1Z3_9FIRM/1-214
tr|A0A2N2BBY9|A0A2N2BBY9_9FIRM/1-212
tr|A0A943QFI8|A0A943QFI8_CLOSP/1-211
tr|A0A9D2DNM5|A0A9D2DNM5_UNCFI/1-210
tr|A0A7D4PVR1|A0A7D4PVR1_9MICO/1-212
tr|A0A7Y2Q3C7|A0A7Y2Q3C7_CELFI/1-214
tr|A0A3D1FR17|A0A3D1FR17_9FIRM/1-214
tr|A0A352P1D8|A0A352P1D8_9FIRM/1-214
tr|A0A353MSH0|A0A353MSH0_CLOSP/1-214
tr|A0A174W2C0|A0A174W2C0_FLAPL/1-230
tr|A0A2V2CGE6|A0A2V2CGE6_UNCCL/1-216
tr|A0A354MP07|A0A354MP07_9FIRM/1-209
tr|A0A354WKQ3|A0A354WKQ3_9FIRM/1-212
tr|A0A3D3ZA04|A0A3D3ZA04_9FIRM/1-212
tr|A0A923MI90|A0A923MI90_9FIRM/1-210
tr|A0A9D1CM36|A0A9D1CM36_9FIRM/1-213
tr|A0A943TZQ4|A0A943TZQ4_UNCFI/1-212
tr|A0A971HY76|A0A971HY76_9CLOT/1-221
tr|A0A9J6R7T3|A0A9J6R7T3_9BACI/1-214
tr|C0CX73|C0CX73_9FIRM/1-218
tr|A0A943EX77|A0A943EX77_9FIRM/1-218
tr|A0A9D1ACQ4|A0A9D1ACQ4_9FIRM/1-216
tr|A0A928HZH4|A0A928HZH4_9FIRM/1-211
tr|A0A3B9WRT3|A0A3B9WRT3_9FIRM/1-218
tr|A0A9D2G643|A0A9D2G643_9FIRM/1-212
tr|A0A3R6KIH0|A0A3R6KIH0_9CLOT/1-213
tr|A0A355VZH3|A0A355VZH3_9FIRM/1-213
tr|A0A2N2DRI6|A0A2N2DRI6_9FIRM/1-224
tr|A0A975CWN5|A0A975CWN5_9FIRM/1-214
tr|A0A6A7K4Y3|A0A6A7K4Y3_9FIRM/1-212
tr|A0A972HUZ2|A0A972HUZ2_9FIRM/1-213
tr|A0A7X8MKT2|A0A7X8MKT2_9FIRM/1-220
tr|A0A3D5UM10|A0A3D5UM10_9FIRM/1-212
tr|A0A970Q2P7|A0A970Q2P7_9FIRM/1-215
tr|A0A7X7S818|A0A7X7S818_9FIRM/1-214
tr|A0A970VXF9|A0A970VXF9_9FIRM/1-211
tr|A0A7X8KV77|A0A7X8KV77_9FIRM/1-211
tr|A0A970D8S1|A0A970D8S1_9FIRM/1-211
tr|A0A970SWS3|A0A970SWS3_9FIRM/1-211
tr|A0A7X9BGM9|A0A7X9BGM9_9FIRM/1-211
tr|A0A7X8YQS0|A0A7X8YQS0_9FIRM/1-211
tr|A0A847FAF6|A0A847FAF6_9CHLR/1-211
tr|A0A7T8B8U3|A0A7T8B8U3_9SPIR/1-215
tr|A0A7X7QNJ1|A0A7X7QNJ1_9CLOT/1-214
tr|A0A7C6U4E6|A0A7C6U4E6_9MOLU/1-214
tr|A0A3P3R120|A0A3P3R120_9FIRM/1-214
tr|A0A8J7H5B0|A0A8J7H5B0_9FIRM/1-213
tr|A0A7C6IZB5|A0A7C6IZB5_9FIRM/1-213
tr|A0A7C6SIS5|A0A7C6SIS5_9FIRM/1-213
tr|A0A521JMI2|A0A521JMI2_9CHLR/1-214
tr|A0A9D5RU12|A0A9D5RU12_9FIRM/1-213
tr|A0A354J647|A0A354J647_9FIRM/1-213
tr|A0A359MT80|A0A359MT80_9FIRM/1-213
tr|F5T7B4|F5T7B4_9FIRM/1-211
tr|A0A413FB02|A0A413FB02_9FIRM/1-210
tr|A0A2T3FGT9|A0A2T3FGT9_9FIRM/1-210
tr|A0A926DFC3|A0A926DFC3_9FIRM/1-223
tr|A0A9D0ZMS1|A0A9D0ZMS1_9FIRM/1-214
tr|A0A351UTA4|A0A351UTA4_9FIRM/1-213
tr|A0A7C6BRX3|A0A7C6BRX3_9FIRM/1-211
tr|A0A3D2GZR0|A0A3D2GZR0_9FIRM/1-212
tr|A0A941XLD4|A0A941XLD4_9FIRM/1-212
tr|A0A9E1B2X0|A0A9E1B2X0_9FIRM/1-212
tr|A0A3D5IGJ3|A0A3D5IGJ3_9FIRM/1-212
tr|A0A975H9Z7|A0A975H9Z7_9FIRM/1-212
tr|A0A7X6WM69|A0A7X6WM69_9FIRM/1-219
tr|A0A972I264|A0A972I264_9FIRM/1-211
tr|A0A4Q2K9Z2|A0A4Q2K9Z2_9FIRM/1-212
tr|A0A859DV40|A0A859DV40_9FIRM/1-232
tr|A0A173VHL6|A0A173VHL6_9FIRM/1-213
tr|C7G6E5|C7G6E5_9FIRM/1-213
tr|A0A414RP98|A0A414RP98_9FIRM/1-213
tr|A0A414T4B2|A0A414T4B2_9FIRM/1-213
tr|A0A7X5E5U4|A0A7X5E5U4_9FIRM/1-215
tr|A0A970MD62|A0A970MD62_9CLOT/1-212
tr|A0A416E7I1|A0A416E7I1_9FIRM/1-212
tr|A0A9D1CPP3|A0A9D1CPP3_9FIRM/1-214
tr|A0A9D5M4M2|A0A9D5M4M2_9FIRM/1-214
tr|A0A9D1DKT0|A0A9D1DKT0_9FIRM/1-214
tr|A0A9D1G3J8|A0A9D1G3J8_9FIRM/1-213
tr|A0A3B8SF75|A0A3B8SF75_9FIRM/1-212
tr|A0A7M2RE44|A0A7M2RE44_9FIRM/1-212
tr|A0A943UB80|A0A943UB80_9FIRM/1-213
tr|A0A9D2Q237|A0A9D2Q237_9FIRM/1-213
```

tr|A0A174U4G1|A0A174U4G1_9CLOT/1-211
tr|A0A7G5MTJ4|A0A7G5MTJ4_9FIRM/1-220
tr|A0A8G1WUE2|A0A8G1WUE2_9FIRM/1-216
tr|A0A9D0Z206|A0A9D0Z206_9FIRM/1-214
tr|A0A7G5N2P2|A0A7G5N2P2_9FIRM/1-212
tr|A0A8G1WU79|A0A8G1WU79_9FIRM/1-212
tr|A0A9D2U6I1|A0A9D2U6I1_9FIRM/1-212
tr|A0A3E3K4M2|A0A3E3K4M2_9FIRM/1-227
tr|A0A417TLC3|A0A417TLC3_9FIRM/1-214
tr|A0A374PFV0|A0A374PFV0_9CLOT/1-213
tr|A0A413XCR6|A0A413XCR6_9CLOT/1-217
tr|A0A4S2HG55|A0A4S2HG55_9FIRM/1-212
tr|A0A3D3RM59|A0A3D3RM59_9FIRM/1-210
tr|A0A9D1J4P7|A0A9D1J4P7_9FIRM/1-211
tr|A0A9D1KT67|A0A9D1KT67_9FIRM/1-211
tr|A0A9E0Z2U4|A0A9E0Z2U4_9FIRM/1-212
tr|A0A173SG09|A0A173SG09_9FIRM/1-223
tr|A0A373IVE1|A0A373IVE1_9FIRM/1-214
tr|A0A173TWX5|A0A173TWX5_9FIRM/1-214
tr|C0B6R8|C0B6R8_9FIRM/1-214
tr|A0A174EB09|A0A174EB09_9FIRM/1-214
tr|A0A1C5LG76|A0A1C5LG76_9FIRM/1-214
tr|A0A3E4GN66|A0A3E4GN66_9FIRM/1-214
tr|A0A416NNM7|A0A416NNM7_9FIRM/1-211
tr|A0A373UDQ7|A0A373UDQ7_9FIRM/1-213
tr|A0A3R6Y6I7|A0A3R6Y6I7_9FIRM/1-213
tr|A0A417TR52|A0A417TR52_9FIRM/1-213
tr|A0A416WM42|A0A416WM42_9FIRM/1-213
tr|A0A374B392|A0A374B392_9FIRM/1-213
tr|A0A3R6QIZ0|A0A3R6QIZ0_9FIRM/1-213
tr|A0A350S3J4|A0A350S3J4_9FIRM/1-213
tr|A0A3D2PPG9|A0A3D2PPG9_9FIRM/1-213
tr|E0S1I0|E0S1I0_BUTPB/1-213
tr|A0A927YLG3|A0A927YLG3_9FIRM/1-209
tr|A0A1C5Y2M4|A0A1C5Y2M4_9CLOT/1-218
tr|A0A417E8K5|A0A417E8K5_9CLOT/1-218
tr|A0A354MPF0|A0A354MPF0_9FIRM/1-211
tr|A0A3D3ZA88|A0A3D3ZA88_9FIRM/1-211
tr|A0A4S2H7Y4|A0A4S2H7Y4_9FIRM/1-212
tr|A0A927ZE29|A0A927ZE29_9FIRM/1-213
tr|A0A3D4X856|A0A3D4X856_9FIRM/1-211
tr|A0A351UX66|A0A351UX66_9FIRM/1-211
tr|A0A3D4XEN3|A0A3D4XEN3_9FIRM/1-213
tr|A0A2E0DI44|A0A2E0DI44_9FLAO/1-216
tr|A0A2E1RJK7|A0A2E1RJK7_9FLAO/1-222
tr|A0A2E8R6X6|A0A2E8R6X6_9FLAO/1-214
tr|A0A3R7T2Q0|A0A3R7T2Q0_9FLAO/1-214
tr|A0A424NPL2|A0A424NPL2_9FLAO/1-214
tr|A0A925BJJ8|A0A925BJJ8_9CHLR/1-223
tr|A0A8T7K8H1|A0A8T7K8H1_UNCCH/1-219
tr|A0A7W1S710|A0A7W1S710_9CHLR/1-216
tr|A0A2N6MNN7|A0A2N6MNN7_9CYAN/1-219
tr|A0A2M8PPR8|A0A2M8PPR8_9CHLR/1-220
tr|A0A8T6PJD6|A0A8T6PJD6_UNCCH/1-213
tr|A0A7V8BXM7|A0A7V8BXM7_9CHLR/1-213
tr|A0A934LWM4|A0A934LWM4_UNCCH/1-214
tr|A0A956Y817|A0A956Y817_9CHLR/1-218
tr|A0A6L9ISI7|A0A6L9ISI7_UNCCH/1-221
tr|A0A7Y5QHI2|A0A7Y5QHI2_9CHLR/1-217
tr|A0A2W4NUE6|A0A2W4NUE6_UNCCH/1-225
tr|A0A3M1XXS6|A0A3M1XXS6_UNCCH/1-221
tr|A0A2M8N7Z7|A0A2M8N7Z7_9CHLR/1-221
tr|A0A2M8NKL0|A0A2M8NKL0_9CHLR/1-221
tr|A0A916GQ65|A0A916GQ65_9CHLR/1-221
tr|A0A916GCX9|A0A916GCX9_9CHLR/1-234
tr|A0A952PXB8|A0A952PXB8_9CHLR/1-218
tr|A0A8T7EZH2|A0A8T7EZH2_UNCCH/1-213
tr|A0A934GZI5|A0A934GZI5_UNCCH/1-211
tr|A0A3M1YKY6|A0A3M1YKY6_UNCCH/1-214
tr|A0A923P7L3|A0A923P7L3_UNCCH/1-214
tr|A0A952C0B1|A0A952C0B1_9CHLR/1-214
tr|A0A931AWV1|A0A931AWV1_9FIRM/1-217
tr|A0A8A7KHX8|A0A8A7KHX8_9FIRM/1-218
tr|A0A8T5X5X0|A0A8T5X5X0_UNCFI/1-218
tr|A0A7C3FSA7|A0A7C3FSA7_9CHLR/1-218
tr|A0A7C1CN15|A0A7C1CN15_9BACT/1-218
tr|A0A947VNG8|A0A947VNG8_9BACT/1-218
tr|A0A3N5RCX6|A0A3N5RCX6_UNCCH/1-240
tr|A0A935CND1|A0A935CND1_9CHLR/1-227
tr|A0A972CRT9|A0A972CRT9_UNCFI/1-211
tr|A0A838DU63|A0A838DU63_9CHLR/1-214
tr|A0A838JCS2|A0A838JCS2_9CHLR/1-214
tr|A0A535BZS5|A0A535BZS5_UNCCH/1-214
tr|A0A535S5X0|A0A535S5X0_UNCCH/1-214
tr|A0A3D1SIJ2|A0A3D1SIJ2_9CHLR/1-214
tr|A0A535ZYI9|A0A535ZYI9_UNCCH/1-214
tr|A0A535IXV1|A0A535IXV1_UNCCH/1-193
tr|A0A535PB08|A0A535PB08_UNCCH/1-214
tr|A0A535FKP6|A0A535FKP6_UNCCH/1-213
tr|A0A535CAG4|A0A535CAG4_UNCCH/1-192
tr|A0A535CDQ8|A0A535CDQ8_UNCCH/1-214
tr|A0A535EZX4|A0A535EZX4_UNCCH/1-204
tr|A0A535VHW3|A0A535VHW3_UNCCH/1-214
tr|A0A352XS15|A0A352XS15_9CHLR/1-215
tr|A0A535S2C8|A0A535S2C8_UNCCH/1-215
tr|A0A535PAW9|A0A535PAW9_UNCCH/1-220
tr|A0A535UIE0|A0A535UIE0_UNCCH/1-215
tr|A0A535SLJ8|A0A535SLJ8_UNCCH/1-215
tr|A0A535ITZ7|A0A535ITZ7_UNCCH/1-215
tr|A0A535C4J6|A0A535C4J6_UNCCH/1-215
tr|A0A535HCF3|A0A535HCF3_UNCCH/1-215
tr|A0A535QI25|A0A535QI25_UNCCH/1-214
tr|A0A4P6JRL1|A0A4P6JRL1_KTERU/1-214
tr|A0A9E3BBP2|A0A9E3BBP2_9CHLR/1-215
tr|A0A353P5F4|A0A353P5F4_9BACT/1-216
tr|A0A916FH50|A0A916FH50_9CHLR/1-233
tr|A0A933J3C6|A0A933J3C6_UNCCH/1-227
tr|A0A920YGJ5|A0A920YGJ5_9CHLR/1-226
tr|A0A3N5UTL4|A0A3N5UTL4_UNCCH/1-218
tr|A0A2N2MBQ2|A0A2N2MBQ2_9CHLR/1-216
tr|A0A2U2SAW8|A0A2U2SAW8_9CHLR/1-215
tr|A0A2N2N0G9|A0A2N2N0G9_9CHLR/1-236
tr|A0A7W0J8E8|A0A7W0J8E8_9CHLR/1-238
tr|A0A2N2NBD8|A0A2N2NBD8_9CHLR/1-238
tr|A0A355GRV9|A0A355GRV9_UNCFI/1-220
tr|A0A7K1IIW1|A0A7K1IIW1_9ACTN/1-217
tr|A0A926V046|A0A926V046_9CYAN/1-230
tr|A0A351L8V7|A0A351L8V7_9CYAN/1-222
tr|A0A8J7DGC4|A0A8J7DGC4_9CYAN/1-228
tr|A0A8J6ZWE7|A0A8J6ZWE7_DESMC/1-205
tr|A0A235HLZ7|A0A235HLZ7_9NOSO/1-233
tr|A0A235IPP6|A0A235IPP6_9NOSO/1-232
tr|A0A7C9PTE9|A0A7C9PTE9_9CYAN/1-235

tr|A0A7C3ZHW7|A0A7C3ZHW7_9CYAN/1-209
tr|A0A978S014|A0A978S014_9CYAN/1-209
tr|A0A9D9PGE5|A0A9D9PGE5_9CYAN/1-231
tr|A0A522XJV4|A0A522XJV4_9CYAN/1-233
tr|A0A6N2DGC3|A0A6N2DGC3_9CYAN/1-234
tr|A0A8K1MQV9|A0A8K1MQV9_9CYAN/1-234
tr|A0A2T1GEC3|A0A2T1GEC3_9CYAN/1-232
tr|A0A7V4P314|A0A7V4P314_9BACT/1-235
tr|A0A7S6MG31|A0A7S6MG31_9BACT/1-219
tr|A0A662EI55|A0A662EI55_9BACT/1-219
tr|A0A7X9J869|A0A7X9J869_9CHLR/1-218
tr|A0A645E3T3|A0A645E3T3_9ZZZZ/1-159
tr|A0A7C4PJM9|A0A7C4PJM9_9CHLR/1-214
tr|A0A3D1JE57|A0A3D1JE57_9CHLR/1-214
tr|E8MZH3|E8MZH3_ANATU/1-214
tr|A0A3B9A4R5|A0A3B9A4R5_9CHLR/1-216
tr|A0A353GZT3|A0A353GZT3_9CHLR/1-217
tr|A0A351X7Y8|A0A351X7Y8_9CHLR/1-217
tr|A0A9D5IWQ1|A0A9D5IWQ1_9CHLR/1-217
tr|A0A2D7MQ89|A0A2D7MQ89_9FLAO/1-220
tr|A0A7Y3HRQ6|A0A7Y3HRQ6_9FLAO/1-219
tr|A0A4Q8QEN0|A0A4Q8QEN0_9FLAO/1-219
tr|A0A5C8V4B9|A0A5C8V4B9_9FLAO/1-219
tr|A0A316KTI5|A0A316KTI5_9FLAO/1-219
tr|A0A3M8GE74|A0A3M8GE74_9FLAO/1-219
tr|A0A371JMB6|A0A371JMB6_9FLAO/1-220
tr|A0A4U1LYU2|A0A4U1LYU2_9FLAO/1-219
tr|A0A3A1N2C4|A0A3A1N2C4_9FLAO/1-219
tr|A0A850NRU5|A0A850NRU5_9FLAO/1-219
tr|A0A2N2P283|A0A2N2P283_9CHLR/1-226
tr|A0A2T3N6T0|A0A2T3N6T0_9GAMM/1-211
tr|A0A2T3NWX3|A0A2T3NWX3_9GAMM/1-211
tr|A0A5N3QY43|A0A5N3QY43_9VIBR/1-210
tr|A0A5P9B873|A0A5P9B873_9VIBR/1-210
tr|A0A3G2QLJ5|A0A3G2QLJ5_9VIBR/1-211
tr|B8K3G5|B8K3G5_9VIBR/1-211
tr|K5UTN2|K5UTN2_9VIBR/1-211
tr|A0A812I023|A0A812I023_9VIBR/1-211
tr|A0A7Y4BJT8|A0A7Y4BJT8_9VIBR/1-210
tr|A0A925IHU4|A0A925IHU4_9CHLR/1-217
tr|A0A6P2B314|A0A6P2B314_9CHLR/1-209
tr|A0A925WLE2|A0A925WLE2_9CHLR/1-180
tr|A0A3M1N3E7|A0A3M1N3E7_UNCCH/1-220
tr|A0A2M8PGV1|A0A2M8PGV1_9CHLR/1-220
tr|A0A2M8P479|A0A2M8P479_9CHLR/1-220
tr|A0A660THP9|A0A660THP9_UNCSP/1-219
tr|A0A949JPU4|A0A949JPU4_9SPIR/1-214
tr|A0A359LKH5|A0A359LKH5_9BACT/1-218
tr|A0A2N3AHQ6|A0A2N3AHQ6_9BACT/1-186
tr|A0A3C0LAK6|A0A3C0LAK6_9BACT/1-218
tr|A0A9C9GDE4|A0A9C9GDE4_UNCFI/1-231
tr|A0A496ZY66|A0A496ZY66_UNCCH/1-180
tr|A0A521I237|A0A521I237_UNCCH/1-218
tr|A0A2N2NJF0|A0A2N2NJF0_9CHLR/1-213
tr|A0A3M2DAU2|A0A3M2DAU2_9BACT/1-219
tr|A0A2J6WYM9|A0A2J6WYM9_9CHLR/1-219
tr|A0A3M1WAE2|A0A3M1WAE2_UNCCH/1-220
tr|A0A7C2PIT8|A0A7C2PIT8_9CHLR/1-220
tr|A0A9D9JA92|A0A9D9JA92_9CHLR/1-216
tr|A0A9D9JR77|A0A9D9JR77_9CHLR/1-216
tr|A0A7C1F068|A0A7C1F068_UNCCH/1-224
tr|A0A7C4QFC1|A0A7C4QFC1_UNCCH/1-215
tr|A0A932DL07|A0A932DL07_UNCCH/1-215
tr|A0A933APS2|A0A933APS2_UNCCH/1-215
tr|A0A424Y5D1|A0A424Y5D1_9FIRM/1-216
tr|A0A7X8HTH3|A0A7X8HTH3_9FIRM/1-217
tr|A0A970GS22|A0A970GS22_9FIRM/1-214
tr|A0A661VRL6|A0A661VRL6_UNCCH/1-216
tr|A0A972GCT6|A0A972GCT6_9CHLR/1-218
tr|A0A9D0KIF1|A0A9D0KIF1_UNCCH/1-233
tr|A0A2E7H324|A0A2E7H324_9FLAO/1-217
tr|A0A7L9BHF9|A0A7L9BHF9_9BACT/1-217
tr|A0A6N2E9T3|A0A6N2E9T3_9SPIO/1-218
tr|A0A8B5WV52|A0A8B5WV52_9SPIO/1-230
tr|A0A7Y2L4T7|A0A7Y2L4T7_9THEO/1-216
tr|A0A831NNK0|A0A831NNK0_THESB/1-216
tr|A0A7C6RE11|A0A7C6RE11_UNCCL/1-215
tr|A0A3B8MVH3|A0A3B8MVH3_THESB/1-215
tr|A0A948AHU7|A0A948AHU7_9GAMM/1-217
tr|A0A948HYM8|A0A948HYM8_9GAMM/1-222
tr|A0A9D5K4W1|A0A9D5K4W1_9BACT/1-217
tr|A0A934VUD5|A0A934VUD5_9BACT/1-219
tr|A0A6C1QWD1|A0A6C1QWD1_9SPIO/1-219
tr|A0A6C1RS37|A0A6C1RS37_9SPIO/1-219
tr|A0A6I7P3V6|A0A6I7P3V6_9SPIO/1-219
tr|A0A7X7G378|A0A7X7G378_9BACT/1-223
tr|A0A970FIT1|A0A970FIT1_9CLOT/1-217
tr|A0A847H2H6|A0A847H2H6_9BACT/1-234
tr|A0A6L5YAN2|A0A6L5YAN2_9BACT/1-236
tr|A0A6A7KBD5|A0A6A7KBD5_9FIRM/1-218
tr|A0A9E1EQV6|A0A9E1EQV6_UNCFI/1-223
tr|A0A7C6ADE7|A0A7C6ADE7_9BACT/1-217
tr|A0A353EX45|A0A353EX45_9FIRM/1-215
tr|A0A970PWA0|A0A970PWA0_9FIRM/1-208
tr|A0A971LT75|A0A971LT75_9FIRM/1-218
tr|A0A7V6HT91|A0A7V6HT91_9MOLU/1-227
tr|A0A925HK03|A0A925HK03_9BACT/1-222
tr|A0A3M2B327|A0A3M2B327_UNCPL/1-212
tr|A0A952VIS2|A0A952VIS2_9BACT/1-226
tr|A0A959PYS2|A0A959PYS2_9BACT/1-223
tr|A0A7X8Q7H1|A0A7X8Q7H1_9SPIR/1-218
tr|A0A413LRS7|A0A413LRS7_9CLOT/1-221
tr|A0A2V2DUZ4|A0A2V2DUZ4_9FIRM/1-227
tr|A0A949NHB5|A0A949NHB5_9FIRM/1-213
tr|A0A7T4YGS1|A0A7T4YGS1_UNCPL/1-220
tr|A0A918NDA1|A0A918NDA1_9GAMM/1-212
tr|A0A3M1KFU1|A0A3M1KFU1_9PROT/1-234
tr|A0A4D7DGF0|A0A4D7DGF0_9SPHN/1-213
tr|A0A432VNS9|A0A432VNS9_9SPHN/1-222
tr|A0A2W5AER5|A0A2W5AER5_9SPHN/1-215
tr|A0A847D962|A0A847D962_9LACT/1-229
tr|A0A2A4QJ03|A0A2A4QJ03_9GAMM/1-217
tr|A0A2N1HPC9|A0A2N1HPC9_9GAMM/1-221
tr|A0A516N0Z1|A0A516N0Z1_9GAMM/1-220
tr|A0A2E0ENG8|A0A2E0ENG8_9RICK/1-228
tr|A0A917YWN6|A0A917YWN6_9ALTE/1-218
tr|A0A8J6INJ0|A0A8J6INJ0_9ALTE/1-216
tr|A0A939DKL6|A0A939DKL6_9ALTE/1-220
tr|A0A920CYN2|A0A920CYN2_9BACL/1-226
tr|A0A948FLC0|A0A948FLC0_9BACT/1-221
tr|A0A9D8EC04|A0A9D8EC04_9ACTN/1-221
tr|A0A3C1I6P2|A0A3C1I6P2_9BACI/1-222
tr|A0A919XDW3|A0A919XDW3_9BACI/1-222

tr|A0A919WQM4|A0A919WQM4_9BACI/1-223
tr|A0A5D8QCD8|A0A5D8QCD8_9THEO/1-217
tr|A0A4Z0WAB4|A0A4Z0WAB4_9GAMM/1-220
tr|A0A923HK67|A0A923HK67_9BURK/1-218
tr|A0A353BV30|A0A353BV30_UNCFI/1-219
tr|A0A2W5UWQ0|A0A2W5UWQ0_9BACT/1-238
tr|A0A248JYS7|A0A248JYS7_9PROT/1-218
tr|A0A7H0LQZ3|A0A7H0LQZ3_9SPHN/1-213
tr|A0A4S1X121|A0A4S1X121_9SPHN/1-219
tr|A0A853FRE5|A0A853FRE5_9SPHN/1-212
tr|A0A926G5N0|A0A926G5N0_9SPHN/1-218
tr|A0A069RGS3|A0A069RGS3_PEPLI/1-220
tr|A0A971E4S9|A0A971E4S9_UNCFI/1-216
tr|A0A955C0H9|A0A955C0H9_9BACT/1-223
tr|A0A951XYI0|A0A951XYI0_9BACT/1-230
tr|A0A3M1R1A9|A0A3M1R1A9_UNCPL/1-223
tr|A0A7S6M5K1|A0A7S6M5K1_9BACT/1-220
tr|A0A3M1BNI4|A0A3M1BNI4_9CHLR/1-185
tr|A0A420ECR6|A0A420ECR6_9ALTE/1-213
tr|A0A411YIQ6|A0A411YIQ6_9ACTN/1-232
tr|A0A2D7MCW9|A0A2D7MCW9_9FLAO/1-220
tr|A0A2E3SDA9|A0A2E3SDA9_9FLAO/1-219
tr|A0A2Z4LV37|A0A2Z4LV37_9FLAO/1-219
tr|A0A964WYG5|A0A964WYG5_9FLAO/1-219
tr|A0A3A1NS19|A0A3A1NS19_9FLAO/1-219
tr|A0A432J1I0|A0A432J1I0_9BACT/1-219
tr|A0A5N5IRU1|A0A5N5IRU1_9FLAO/1-219
tr|A0A6I5KYQ5|A0A6I5KYQ5_9FLAO/1-219
tr|A0A3A1NIW3|A0A3A1NIW3_9FLAO/1-219
tr|A0A418NA24|A0A418NA24_9FLAO/1-219
tr|A0A958M313|A0A958M313_9FLAO/1-219
tr|A0A444VLA9|A0A444VLA9_9FLAO/1-219
tr|A0A4S8RJY9|A0A4S8RJY9_9FLAO/1-219
tr|A0A9D9ETF6|A0A9D9ETF6_9FLAO/1-219
tr|A0A2E1AP15|A0A2E1AP15_9FLAO/1-219
tr|A0A6G7IZE7|A0A6G7IZE7_9FLAO/1-219
tr|A0A2D9C0T3|A0A2D9C0T3_9FLAO/1-219
tr|A0A358G125|A0A358G125_9FLAO/1-219
tr|A0A3R7TJF6|A0A3R7TJF6_9FLAO/1-219
tr|A0A7K3NAR0|A0A7K3NAR0_9FLAO/1-219
tr|A0A355FUL3|A0A355FUL3_9FLAO/1-187
tr|A0A381ICC5|A0A381ICC5_CLODI/1-42
tr|K8EJG4|K8EJG4_CARML/1-66
tr|A0A9D7U495|A0A9D7U495_9BACT/1-215
tr|A0A3C1TIG0|A0A3C1TIG0_9BACT/1-219
tr|A0A344TFT3|A0A344TFT3_9BACT/1-220
tr|A0A369I702|A0A369I702_9BACT/1-220
tr|A0A917DM63|A0A917DM63_9BACT/1-216
tr|A0A2N4XB88|A0A2N4XB88_9BACT/1-217
tr|A0A838YXC7|A0A838YXC7_9BACT/1-216
tr|A0A372F4F1|A0A372F4F1_9BACT/1-216
tr|A0A4Q5LZ47|A0A4Q5LZ47_9BACT/1-216
tr|A0A939GB77|A0A939GB77_9BACT/1-223
tr|A0A939GHW7|A0A939GHW7_9BACT/1-223
tr|A0A3P1C0Y3|A0A3P1C0Y3_9BACT/1-221
tr|A0A368JJK3|A0A368JJK3_9BACT/1-221
tr|A0A5N1JSE8|A0A5N1JSE8_9BACT/1-221
tr|A0A3P1CXB5|A0A3P1CXB5_9BACT/1-220
tr|A0A7K0EL79|A0A7K0EL79_9BACT/1-221
tr|A0A924JBZ9|A0A924JBZ9_9BACT/1-220
tr|A0A418MEZ4|A0A418MEZ4_9BACT/1-218
tr|A0A7L5DV62|A0A7L5DV62_9BACT/1-218
tr|A0A6M5Y898|A0A6M5Y898_9BACT/1-220
tr|A0A4Q2UM04|A0A4Q2UM04_9BACT/1-220
tr|A0A6G9ATR2|A0A6G9ATR2_9BACT/1-219
tr|A0A6P1WAU7|A0A6P1WAU7_9BACT/1-219
tr|A0A6L9L6U8|A0A6L9L6U8_9BACT/1-221
tr|A0A6M0IFD0|A0A6M0IFD0_9BACT/1-219
tr|A0A927B4C9|A0A927B4C9_9BACT/1-218
tr|A0A2K8YSG3|A0A2K8YSG3_9BACT/1-218
tr|A0A926Y314|A0A926Y314_9BACT/1-219
tr|A0A7G5H006|A0A7G5H006_9BACT/1-218
tr|A0A515A1P3|A0A515A1P3_9BACT/1-218
tr|A0A7K1SGH3|A0A7K1SGH3_9BACT/1-218
tr|A0A3D1PZU7|A0A3D1PZU7_9FIRM/1-219
tr|A0A6B8REM0|A0A6B8REM0_9BACL/1-211
tr|A0A940GDI9|A0A940GDI9_9BACL/1-212
tr|A0A494XWU2|A0A494XWU2_9BACL/1-214
tr|A0A974PBU2|A0A974PBU2_9BACL/1-209
tr|A0A7X2L2Q6|A0A7X2L2Q6_9BACL/1-216
tr|A0A2Z2KPE4|A0A2Z2KPE4_9BACL/1-209
tr|A0A5B0W427|A0A5B0W427_9BACL/1-211
tr|A0A850EUP3|A0A850EUP3_9BACL/1-210
tr|A0A7X3LJI1|A0A7X3LJI1_9BACL/1-211
tr|A0A7X3GLY6|A0A7X3GLY6_9BACL/1-211
tr|A0A9J6ZB73|A0A9J6ZB73_9BACL/1-211
tr|A0A3Q9I8T3|A0A3Q9I8T3_9BACL/1-211
tr|A0A3S1DET8|A0A3S1DET8_9BACL/1-212
tr|A0A090Z802|A0A090Z802_PAEMA/1-211
tr|A0A5D0CSC9|A0A5D0CSC9_9BACL/1-213
tr|A0A7X2ZL62|A0A7X2ZL62_9BACL/1-214
tr|C6J752|C6J752_9BACL/1-210
tr|A0A4Q0VXZ1|A0A4Q0VXZ1_9BACI/1-215
tr|A0A7S7L7B2|A0A7S7L7B2_9BACI/1-215
tr|A0A969ZK39|A0A969ZK39_9CLOT/1-188
tr|A0A410DR16|A0A410DR16_9CLOT/1-214
tr|A0A848M6T4|A0A848M6T4_PAELE/1-223
tr|A0A3S1CBM2|A0A3S1CBM2_9BACL/1-222
tr|A0A3Q8S8S9|A0A3Q8S8S9_9BACL/1-220
tr|A0A7X3CPR6|A0A7X3CPR6_9BACL/1-220
tr|A0A3S9UT90|A0A3S9UT90_9BACL/1-219
tr|A0A090ZMP4|A0A090ZMP4_PAEMA/1-216
tr|A0A3P3U1R0|A0A3P3U1R0_9BACL/1-216
tr|A0A7X2ZN70|A0A7X2ZN70_9BACL/1-219
tr|C6IYD6|C6IYD6_9BACL/1-219
tr|A0A7X3INF6|A0A7X3INF6_9BACL/1-223
tr|A0A259T5E1|A0A259T5E1_9BACL/1-223
tr|A0A5J5FZC3|A0A5J5FZC3_9BACL/1-220
tr|A0A3N9P8V6|A0A3N9P8V6_9BACL/1-220
tr|A0A6M1PHJ7|A0A6M1PHJ7_9BACL/1-220
tr|A0A850F158|A0A850F158_9BACL/1-220
tr|A0A0E4H9W8|A0A0E4H9W8_9BACL/1-220
tr|A0A974P7G4|A0A974P7G4_9BACL/1-220
tr|A0A2Z2K3Q3|A0A2Z2K3Q3_9BACL/1-220
tr|A0A5B0WI41|A0A5B0WI41_9BACL/1-218
tr|A0A357MC44|A0A357MC44_PAESP/1-218
tr|A0A7X2H3W4|A0A7X2H3W4_9BACL/1-220
tr|A0A5D0CP25|A0A5D0CP25_9BACL/1-223
tr|A0A2M9MRZ9|A0A2M9MRZ9_9BACL/1-216
tr|A0A328W9J5|A0A328W9J5_PAELA/1-216
tr|A0A385TW43|A0A385TW43_PAELA/1-216
tr|F3MCV1|F3MCV1_9BACL/1-216
tr|A0A2A5LJC9|A0A2A5LJC9_PAELA/1-216

```
tr|A0A3S1D643|A0A3S1D643_9BACL/1-216
tr|A0A2S0U5S7|A0A2S0U5S7_9BACL/1-217
tr|A0A359LA07|A0A359LA07_9BACT/1-215
tr|A0A919RXV3|A0A919RXV3_9CLOT/1-222
tr|A0A9E8W0C1|A0A9E8W0C1_9CLOT/1-215
tr|A0A7D4Q8G8|A0A7D4Q8G8_9SPHI/1-216
tr|A0A425XXQ3|A0A425XXQ3_9BACT/1-219
tr|A0A4Q1JNZ6|A0A4Q1JNZ6_9BACT/1-218
tr|A0A434AYL0|A0A434AYL0_9BACT/1-215
tr|A0A6L5DZ10|A0A6L5DZ10_9BACT/1-215
tr|A0A2V3ZYA3|A0A2V3ZYA3_9BACT/1-215
tr|A0A425YEL9|A0A425YEL9_9BACT/1-215
tr|A0A2A4NBX2|A0A2A4NBX2_9BACT/1-215
tr|A0A7C6MQE9|A0A7C6MQE9_9BACT/1-215
tr|A0A7V6H2P0|A0A7V6H2P0_9BACT/1-215
tr|A0A9J6ZNI4|A0A9J6ZNI4_9BACT/1-215
tr|A0A972A7S1|A0A972A7S1_9BACT/1-215
tr|A0A353TWQ7|A0A353TWQ7_9BACT/1-215
tr|A0A359G0B5|A0A359G0B5_9BACT/1-215
tr|A0A971IUZ9|A0A971IUZ9_9BACT/1-219
tr|A0A5R9QYP3|A0A5R9QYP3_9BACT/1-218
tr|A0A2U2B8B3|A0A2U2B8B3_9BACT/1-214
tr|A0A4Q8RZG6|A0A4Q8RZG6_9BACT/1-215
tr|A0A553GN04|A0A553GN04_9BACT/1-217
tr|A0A941IYN4|A0A941IYN4_9BACT/1-217
tr|A0A7V3IU06|A0A7V3IU06_9FLAO/1-196
tr|A0A847LZS4|A0A847LZS4_9SPIR/1-182
tr|A0A1C6BQQ4|A0A1C6BQQ4_9CLOT/1-216
tr|A0A949K1X9|A0A949K1X9_9FIRM/1-216
tr|A0A970J5C3|A0A970J5C3_UNCFI/1-219
tr|A0A7C6CG50|A0A7C6CG50_UNCFI/1-225
tr|A0A353MPE8|A0A353MPE8_UNCFI/1-234
tr|A0A7C6RUK4|A0A7C6RUK4_UNCFI/1-220
tr|A0A3C0H6Z4|A0A3C0H6Z4_UNCFI/1-222
tr|A0A3C1QKI5|A0A3C1QKI5_UNCFI/1-222
tr|A0A524QK34|A0A524QK34_9BACT/1-215
tr|A0A6L8TXF6|A0A6L8TXF6_9BACT/1-215
tr|A0A3D3HFU8|A0A3D3HFU8_9BACT/1-173
tr|A0A7X7ETP7|A0A7X7ETP7_9BACT/1-215
tr|A0A7X7CZJ7|A0A7X7CZJ7_9BACT/1-216
tr|A0A959U7H7|A0A959U7H7_9BACT/1-216
tr|A0A661Z9S3|A0A661Z9S3_9BACT/1-215
tr|A0A7G2MC46|A0A7G2MC46_9BACT/1-223
tr|A0A662BR29|A0A662BR29_9BACT/1-219
tr|A0A969T6Q3|A0A969T6Q3_9CHLR/1-219
tr|A0A3C1PKY9|A0A3C1PKY9_9BACT/1-197
tr|A0A5C1QDA7|A0A5C1QDA7_9SPIO/1-212
tr|A0A7Y8NV10|A0A7Y8NV10_9BACT/1-216
tr|A0A353Z4I7|A0A353Z4I7_9BACT/1-218
tr|A0A353Z2B4|A0A353Z2B4_9BACT/1-217
tr|A0A8J6YTQ3|A0A8J6YTQ3_9BACL/1-217
tr|A0A355RXH5|A0A355RXH5_9CLOT/1-216
tr|A0A7C4RYV1|A0A7C4RYV1_FERPE/1-215
tr|A0A7V4KD59|A0A7V4KD59_FERPE/1-215
tr|A0A832IJY9|A0A832IJY9_FERPE/1-215
tr|A0A6P1Y983|A0A6P1Y983_9FIRM/1-215
tr|A0A7X8MT45|A0A7X8MT45_UNCCL/1-214
tr|A0A3D5KH93|A0A3D5KH93_CLOSP/1-214
tr|A0A1M4N8G2|A0A1M4N8G2_9CLOT/1-214
tr|A0A970DSF3|A0A970DSF3_9FIRM/1-214
tr|A0A925ZAH6|A0A925ZAH6_9CLOT/1-215
tr|A0A0R3K2R0|A0A0R3K2R0_CALMK/1-214
tr|A0A5B7TDW8|A0A5B7TDW8_9CLOT/1-215
tr|A0A6P1P040|A0A6P1P040_9BACT/1-219
tr|A0A7X9FAE8|A0A7X9FAE8_9BACT/1-225
tr|A0A7K1SX44|A0A7K1SX44_9SPHI/1-215
tr|A0A4Q3FTQ7|A0A4Q3FTQ7_9SPHI/1-215
tr|A0A4V1T2H1|A0A4V1T2H1_9SPHI/1-215
tr|A0A7D3XXA3|A0A7D3XXA3_9BACT/1-219
tr|A0A355IZ67|A0A355IZ67_9BACT/1-218
tr|A0A3D5YV59|A0A3D5YV59_9BACT/1-214
tr|A0A923IJ05|A0A923IJ05_9BACT/1-227
tr|A0A2S4GSH8|A0A2S4GSH8_9FIRM/1-213
tr|A0A8G1TM15|A0A8G1TM15_9FIRM/1-213
tr|A0A970UN95|A0A970UN95_UNCFI/1-224
tr|A0A7C4XDH4|A0A7C4XDH4_9BACT/1-224
tr|A0A7C6CWV1|A0A7C6CWV1_UNCFI/1-220
tr|A0A970EQB1|A0A970EQB1_UNCFI/1-221
tr|A0A353LYK8|A0A353LYK8_UNCFI/1-222
tr|A0A356BYL6|A0A356BYL6_UNCFI/1-218
tr|A0A965A8I6|A0A965A8I6_9BACT/1-223
tr|A0A352MH11|A0A352MH11_9BACT/1-219
tr|A0A7X8L6G4|A0A7X8L6G4_9CLOT/1-222
tr|A0A7C7E630|A0A7C7E630_9CLOT/1-217
tr|A0A971EMG8|A0A971EMG8_9CLOT/1-217
tr|A0A9D1C1U5|A0A9D1C1U5_9CLOT/1-217
tr|A0A7Z2VRF0|A0A7Z2VRF0_9BACL/1-220
tr|A0A2U8E525|A0A2U8E525_9BACT/1-224
tr|A0A3D0ZD38|A0A3D0ZD38_9SPIR/1-190
tr|A0A7C6Z7L3|A0A7C6Z7L3_UNCFI/1-235
tr|A0A2U8E7X5|A0A2U8E7X5_9BACT/1-234
tr|A0A7C3LRC9|A0A7C3LRC9_UNCCH/1-213
tr|A0A3N5UIL3|A0A3N5UIL3_UNCCH/1-211
tr|A0A925EGP2|A0A925EGP2_9CHLR/1-219
tr|A0A9J6R7S0|A0A9J6R7S0_9BACI/1-218
tr|A0A323TIJ2|A0A323TIJ2_9BACI/1-215
tr|A0A7V6LIN9|A0A7V6LIN9_9FIRM/1-218
tr|A0A847WSK5|A0A847WSK5_9BACI/1-225
tr|A0A3E2WQT0|A0A3E2WQT0_9CLOT/1-220
tr|A0A941XFI3|A0A941XFI3_9BACT/1-221
tr|A0A267ML21|A0A267ML21_9CLOT/1-216
tr|A0A355WDG9|A0A355WDG9_9CLOT/1-215
tr|A0A7C9H944|A0A7C9H944_UNCFI/1-217
tr|A0A7C9L8F0|A0A7C9L8F0_UNCFI/1-216
tr|A0A847NEH2|A0A847NEH2_9FIRM/1-217
tr|A0A9J6P2Q3|A0A9J6P2Q3_9CLOT/1-214
tr|A0A3C2D3Z5|A0A3C2D3Z5_9BACT/1-228
tr|A0A3B9ZMA9|A0A3B9ZMA9_9BACT/1-220
tr|A0A925SEP4|A0A925SEP4_9BACT/1-217
tr|A0A2N2X5J8|A0A2N2X5J8_9BACT/1-215
tr|A0A5N7IK65|A0A5N7IK65_9CLOT/1-218
tr|A0A7Y3WSU1|A0A7Y3WSU1_9CLOT/1-218
tr|A0A2M8TEL8|A0A2M8TEL8_9CLOT/1-216
tr|A0A7Y3NX47|A0A7Y3NX47_9FIRM/1-217
tr|A0A410PQR4|A0A410PQR4_9CLOT/1-216
tr|A0A662AFB3|A0A662AFB3_9BACT/1-218
tr|A0A929FZI3|A0A929FZI3_9BACT/1-219
tr|A0A849UQ33|A0A849UQ33_9BACT/1-218
tr|A0A3D4TWH3|A0A3D4TWH3_9BACT/1-218
tr|A0A8I2A732|A0A8I2A732_9BACT/1-217
tr|A0A6I0E233|A0A6I0E233_9BACT/1-219
tr|A0A7C1CR58|A0A7C1CR58_9BACT/1-220
tr|A0A7C4JX68|A0A7C4JX68_9BACT/1-220
```

tr|A0A9E0A3R9|A0A9E0A3R9_9BACT/1-220
tr|A0A958ZWN6|A0A958ZWN6_9BACT/1-217
tr|A0A7K3XHR9|A0A7K3XHR9_9BACT/1-217
tr|A0A174GBF5|A0A174GBF5_9CLOT/1-212
tr|A0A3E4U5Z0|A0A3E4U5Z0_9CLOT/1-212
tr|A0A662C096|A0A662C096_9BACT/1-218
tr|A0A6P1TLU5|A0A6P1TLU5_9FIRM/1-217
tr|A0A6I3Q6S5|A0A6I3Q6S5_9FIRM/1-218
tr|A0A975EYE9|A0A975EYE9_9SPIR/1-223
tr|A0A6I5IBC8|A0A6I5IBC8_9BACT/1-216
tr|A0A937K482|A0A937K482_9CLOT/1-216
tr|A0A936QCI0|A0A936QCI0_9RHOO/1-216
tr|A0A7X7WMJ4|A0A7X7WMJ4_9BACT/1-222
tr|A0A7C6EVN3|A0A7C6EVN3_9BACT/1-222
tr|A0A7V4LQP8|A0A7V4LQP8_9BACT/1-222
tr|A0A1G6L589|A0A1G6L589_9BACT/1-216
tr|A0A7C4CCG7|A0A7C4CCG7_9BACT/1-215
tr|A0A7X7TT34|A0A7X7TT34_9THEM/1-216
tr|A0A3D3DYU7|A0A3D3DYU7_9BACT/1-216
tr|A0A3D2GHT6|A0A3D2GHT6_9BACT/1-216
tr|A0A250FVZ5|A0A250FVZ5_9FLAO/1-210
tr|A0A958WMJ4|A0A958WMJ4_9BACT/1-217
tr|A0A4R0N653|A0A4R0N653_9SPHI/1-221
tr|A0A2R7L2I6|A0A2R7L2I6_9SPHI/1-224
tr|A0A7K0FRB6|A0A7K0FRB6_9SPHI/1-223
tr|A0A0B8XTZ1|A0A0B8XTZ1_9SPHI/1-223
tr|A0A5C0VIB6|A0A5C0VIB6_9SPHI/1-223
tr|A0A928UWN0|A0A928UWN0_9SPHI/1-214
tr|A0A519U789|A0A519U789_9SPHI/1-220
tr|A0A520C7A9|A0A520C7A9_9SPHI/1-215
tr|A0A4U1C266|A0A4U1C266_9SPHI/1-216
tr|A0A4U1CLQ8|A0A4U1CLQ8_9SPHI/1-216
tr|A0A2T7BN19|A0A2T7BN19_9BACT/1-216
tr|A0A924MCR6|A0A924MCR6_9SPHI/1-217
tr|A0A519U1F7|A0A519U1F7_9SPHI/1-222
tr|A0A924KE93|A0A924KE93_9SPHI/1-223
tr|A0A6B9Z3S8|A0A6B9Z3S8_9SPHI/1-226
tr|A0A7S9Q196|A0A7S9Q196_9SPHI/1-229
tr|A0A369Q1N7|A0A369Q1N7_9SPHI/1-228
tr|A0A7K0FUC8|A0A7K0FUC8_9SPHI/1-228
tr|A0A3N0BNP4|A0A3N0BNP4_9SPHI/1-232
tr|A0A3N7F0P2|A0A3N7F0P2_9SPHI/1-232
tr|A0A4V2JHA6|A0A4V2JHA6_9SPHI/1-232
tr|A0A366KPM1|A0A366KPM1_9SPHI/1-227
tr|A0A4R0PH76|A0A4R0PH76_9SPHI/1-228
tr|A0A7G9QGD9|A0A7G9QGD9_9SPHI/1-228
tr|A0A317EJQ4|A0A317EJQ4_9SPHI/1-227
tr|A0A4Q3BC72|A0A4Q3BC72_9SPHI/1-225
tr|A0A519VZ35|A0A519VZ35_9SPHI/1-226
tr|A0A519Y945|A0A519Y945_9SPHI/1-226
tr|A0A317F284|A0A317F284_9SPHI/1-220
tr|A0A4Q3TDK3|A0A4Q3TDK3_9BACT/1-226
tr|A0A519RVU4|A0A519RVU4_9SPHI/1-226
tr|A0A3G8X109|A0A3G8X109_9SPHI/1-226
tr|A0A4R0P5N9|A0A4R0P5N9_9SPHI/1-226
tr|A0A354BXE9|A0A354BXE9_9BACT/1-239
tr|A0A357VRR7|A0A357VRR7_9BACT/1-214
tr|A0A848J3A2|A0A848J3A2_9BACT/1-216
tr|A0A9D6JUI6|A0A9D6JUI6_9BURK/1-219
tr|A0A932W944|A0A932W944_9BURK/1-216
tr|A0A923KUD7|A0A923KUD7_9BURK/1-217
tr|A0A941E5R8|A0A941E5R8_9BURK/1-216
tr|A0A941I485|A0A941I485_9BURK/1-216
tr|A0A959DXH4|A0A959DXH4_9BACT/1-219
tr|A0A963ZY35|A0A963ZY35_9BACT/1-219
tr|A0A5D6V5T2|A0A5D6V5T2_9BACT/1-222
tr|A0A3M9MQD1|A0A3M9MQD1_9BACT/1-220
tr|A0A3M9MTV9|A0A3M9MTV9_9BACT/1-220
tr|A0A5B6TCK0|A0A5B6TCK0_9BACT/1-228
tr|A0A5M8Q761|A0A5M8Q761_9BACT/1-220
tr|A0A2S7IQQ5|A0A2S7IQQ5_9BACT/1-220
tr|A0A261PXL8|A0A261PXL8_9BACT/1-220
tr|A0A2N7BG64|A0A2N7BG64_9BACT/1-220
tr|A0A267T991|A0A267T991_9BACT/1-216
tr|A0A923UNS2|A0A923UNS2_9BACT/1-214
tr|A0A521VJP1|A0A521VJP1_9BACT/1-217
tr|A0A2E0VW36|A0A2E0VW36_9BACT/1-221
tr|A0A354S873|A0A354S873_9FLAO/1-151
tr|A0A257IYR6|A0A257IYR6_9BACT/1-216
tr|A0A924KDJ1|A0A924KDJ1_9SPHI/1-218
tr|A0A2S5A0N6|A0A2S5A0N6_9SPHI/1-216
tr|A0A2W5F1P0|A0A2W5F1P0_9SPHI/1-216
tr|A0A3E1NW56|A0A3E1NW56_9BACT/1-217
tr|A0A5B2VNX8|A0A5B2VNX8_9BACT/1-213
tr|A0A7K1U2G3|A0A7K1U2G3_9BACT/1-216
tr|A0A5C6LQA1|A0A5C6LQA1_9BACT/1-219
tr|A0A6B9ZQU2|A0A6B9ZQU2_9BACT/1-219
tr|A0A3E2NJB2|A0A3E2NJB2_9SPHI/1-214
tr|A0A444MMU5|A0A444MMU5_9SPHI/1-217
tr|A0A4Y8S7H9|A0A4Y8S7H9_9SPHI/1-217
tr|A0A2A2S7Z0|A0A2A2S7Z0_9SPHI/1-216
tr|A0A7V1SQJ3|A0A7V1SQJ3_9BACT/1-216
tr|A0A2J6H3G8|A0A2J6H3G8_9SPHI/1-216
tr|A0A372NYU0|A0A372NYU0_9SPHI/1-216
tr|A0A563U8M7|A0A563U8M7_9SPHI/1-216
tr|A0A934PS59|A0A934PS59_9SPHI/1-218
tr|A0A4Q5LPE1|A0A4Q5LPE1_9SPHI/1-216
tr|A0A563U3D7|A0A563U3D7_9SPHI/1-217
tr|A0A4Y8AJ15|A0A4Y8AJ15_9SPHI/1-216
tr|A0A926NS96|A0A926NS96_9SPHI/1-217
tr|A0A4R0MY26|A0A4R0MY26_9SPHI/1-216
tr|A0A519VA02|A0A519VA02_9SPHI/1-216
tr|A0A923ITX3|A0A923ITX3_9SPHI/1-216
tr|A0A520ADG5|A0A520ADG5_9SPHI/1-216
tr|A0A520CXY4|A0A520CXY4_9SPHI/1-216
tr|A0A519P0V5|A0A519P0V5_FLASP/1-216
tr|A0A519VEM7|A0A519VEM7_9SPHI/1-216
tr|A0A249SYJ5|A0A249SYJ5_9BACT/1-216
tr|A0A4Q6EI92|A0A4Q6EI92_9SPHI/1-216
tr|A0A5B8UZQ3|A0A5B8UZQ3_9SPHI/1-216
tr|A0A929KS25|A0A929KS25_9SPHI/1-216
tr|A0A5B8W339|A0A5B8W339_9SPHI/1-215
tr|A0A494VZ94|A0A494VZ94_9SPHI/1-216
tr|A0A5C1I0S5|A0A5C1I0S5_9SPHI/1-216
tr|A0A364WPZ9|A0A364WPZ9_9SPHI/1-216
tr|A0A8A5K0W8|A0A8A5K0W8_9SPHI/1-216
tr|A0A6I4I2G7|A0A6I4I2G7_9SPHI/1-216
tr|A0A966DV99|A0A966DV99_9SPHI/1-216
tr|A0A5P2G3K4|A0A5P2G3K4_9BACT/1-219
tr|A0A2U2PFT9|A0A2U2PFT9_9SPHI/1-216
tr|A0A4Q0M9E7|A0A4Q0M9E7_9SPHI/1-216
tr|A0A5M9HD13|A0A5M9HD13_9SPHI/1-216
tr|A0A6B3PHD4|A0A6B3PHD4_9BACT/1-216

```
tr|A0A2M6GEM9|A0A2M6GEM9_9BACT/1-217
tr|A0A7G7GEI2|A0A7G7GEI2_9BACT/1-221
tr|A0A7L7LBB4|A0A7L7LBB4_9BACT/1-219
tr|A0A2T2YFM9|A0A2T2YFM9_9BACT/1-220
tr|A0A354W6G1|A0A354W6G1_9FLAO/1-215
tr|A0A2T4WBR4|A0A2T4WBR4_9BACT/1-215
tr|A0A348TVY7|A0A348TVY7_9FLAO/1-215
tr|A0A937IRS5|A0A937IRS5_9FLAO/1-215
tr|A0A3C2ARH3|A0A3C2ARH3_9FLAO/1-217
tr|A0A3C0U6R5|A0A3C0U6R5_9FLAO/1-214
tr|A0A2E2Y952|A0A2E2Y952_9FLAO/1-214
tr|A0A7W2GGQ0|A0A7W2GGQ0_9FLAO/1-214
tr|A0A3E0MUK2|A0A3E0MUK2_9BACT/1-214
tr|A0A3E0R5M4|A0A3E0R5M4_9BACT/1-214
tr|A0A2B4FHP8|A0A2B4FHP8_9BACI/1-223
tr|A0A7S8HGU5|A0A7S8HGU5_9BACI/1-215
tr|A0A959QGY5|A0A959QGY5_9BACT/1-223
tr|A0A2E0KQA5|A0A2E0KQA5_9FLAO/1-218
tr|A0A9D7W5G9|A0A9D7W5G9_9BACT/1-218
tr|A0A2D9X3N0|A0A2D9X3N0_9EURY/1-219
tr|A0A9E1YMG0|A0A9E1YMG0_9FLAO/1-218
tr|A0A651I0F2|A0A651I0F2_9BACT/1-222
tr|A0A7Y2TFZ3|A0A7Y2TFZ3_9BACT/1-218
tr|A0A9D7ZDY5|A0A9D7ZDY5_9BACT/1-217
tr|A0A351GJU9|A0A351GJU9_9BACT/1-217
tr|A0A2A5E5D8|A0A2A5E5D8_9BACT/1-216
tr|A0A355U9P6|A0A355U9P6_9BACT/1-216
tr|A0A3D1H5L9|A0A3D1H5L9_9BACT/1-216
tr|A0A352P1W8|A0A352P1W8_9FIRM/1-216
tr|A0A3N5K1B1|A0A3N5K1B1_9BACT/1-215
tr|A0A353P141|A0A353P141_9BACT/1-214
tr|A0A3N5M9B9|A0A3N5M9B9_UNCCH/1-235
tr|A0A971MSH9|A0A971MSH9_UNCFI/1-214
tr|A0A969FJ56|A0A969FJ56_9CYAN/1-236
tr|A0A0P8A2H1|A0A0P8A2H1_9CYAN/1-234
tr|A0A928VY80|A0A928VY80_9CYAN/1-231
tr|A0A9C8TXE2|A0A9C8TXE2_9BACT/1-232
tr|A0A971IYZ5|A0A971IYZ5_UNCFI/1-233
tr|A0A316S4B4|A0A316S4B4_9FIRM/1-227
tr|A0A973JI40|A0A973JI40_9CELL/1-226
tr|A0A9D1CRQ9|A0A9D1CRQ9_9FIRM/1-220
tr|A0A416R6D0|A0A416R6D0_9FIRM/1-219
tr|D1Y0W5|D1Y0W5_9BACT/1-134
tr|A0A925WPK5|A0A925WPK5_9CHLR/1-221
tr|A0A7Y5QCJ4|A0A7Y5QCJ4_9CHLR/1-219
tr|A0A3M1YAS6|A0A3M1YAS6_UNCCH/1-216
tr|A0A2M8NBM8|A0A2M8NBM8_9CHLR/1-216
tr|A0A6P2B6D1|A0A6P2B6D1_9CHLR/1-215
tr|A0A7S8EAY2|A0A7S8EAY2_9CHLR/1-225
tr|A0A2E1AKS4|A0A2E1AKS4_9CHLR/1-218
tr|A0A958C9G4|A0A958C9G4_9CHLR/1-220
tr|A0A958HGE2|A0A958HGE2_9CHLR/1-220
tr|A0A7Y3U9K7|A0A7Y3U9K7_UNCPL/1-229
tr|A0A7C2KX86|A0A7C2KX86_9BACT/1-226
tr|A0A936IG38|A0A936IG38_9BACT/1-234
tr|A0A357T8S6|A0A357T8S6_9CLOT/1-213
tr|A0A939KIE5|A0A939KIE5_9CLOT/1-214
tr|A0A7Z0PGJ9|A0A7Z0PGJ9_9FUSO/1-213
tr|D0GPR0|D0GPR0_9FUSO/1-212
tr|A0A510JCK8|A0A510JCK8_9FUSO/1-212
tr|A0A7Y2BDI0|A0A7Y2BDI0_9BACT/1-216
tr|A0A958X400|A0A958X400_9BACT/1-229
tr|A0A2E4SY61|A0A2E4SY61_9FLAO/1-224
tr|A0A2E6BUU3|A0A2E6BUU3_9FLAO/1-215
tr|A0A2E4S911|A0A2E4S911_9FLAO/1-217
tr|A0A2E7CTZ1|A0A2E7CTZ1_9FLAO/1-215
tr|A0A9D0YUC6|A0A9D0YUC6_9FIRM/1-220
tr|A0A6N2CUN7|A0A6N2CUN7_9BACT/1-225
tr|A0A7M3MY71|A0A7M3MY71_9BACT/1-223
tr|A0A3D2X4W5|A0A3D2X4W5_9FIRM/1-220
tr|A0A1S8TMM9|A0A1S8TMM9_9CLOT/1-215
tr|A0A7D6ZR44|A0A7D6ZR44_9CLOT/1-215
tr|A0A496QCW2|A0A496QCW2_9BACT/1-217
tr|A0A940D975|A0A940D975_9BACT/1-217
tr|A0A7X6XZ71|A0A7X6XZ71_CLOSP/1-214
tr|A0A358S4M4|A0A358S4M4_9CLOT/1-213
tr|A0A941CPS2|A0A941CPS2_9CLOT/1-213
tr|A0A7M1TA59|A0A7M1TA59_9BACT/1-228
tr|A0A936GEU0|A0A936GEU0_9BACT/1-219
tr|A0A2A5XD73|A0A2A5XD73_9FLAO/1-219
tr|A0A2D8KHR6|A0A2D8KHR6_9FLAO/1-218
tr|A0A424PI60|A0A424PI60_9FLAO/1-215
tr|A0A2E1P5R3|A0A2E1P5R3_9FLAO/1-215
tr|A0A2T2VSF6|A0A2T2VSF6_9BACT/1-201
tr|A0A520UM58|A0A520UM58_9FLAO/1-202
tr|A0A520UE19|A0A520UE19_9FLAO/1-216
tr|A0A937GR99|A0A937GR99_9FLAO/1-213
tr|A0A2E8R6J7|A0A2E8R6J7_9FLAO/1-214
tr|A0A2D7A4M3|A0A2D7A4M3_9FLAO/1-216
tr|A0A2D5LT60|A0A2D5LT60_9FLAO/1-216
tr|A0A2E8FXZ4|A0A2E8FXZ4_9FLAO/1-216
tr|A0A2E5Z3G8|A0A2E5Z3G8_9FLAO/1-221
tr|A0A2E8VCQ4|A0A2E8VCQ4_9FLAO/1-218
tr|A0A2D5HN19|A0A2D5HN19_9FLAO/1-223
tr|A0A2D8JA99|A0A2D8JA99_9FLAO/1-223
tr|A0A3R7V3V1|A0A3R7V3V1_9FLAO/1-217
tr|A0A2U2RX44|A0A2U2RX44_9BACT/1-216
tr|A0A661XDF1|A0A661XDF1_9BACT/1-222
tr|A0A936VZ49|A0A936VZ49_9BACT/1-230
tr|A0A935G7S3|A0A935G7S3_9BACT/1-225
tr|A0A958XSL7|A0A958XSL7_9BACT/1-229
tr|A0A9D1LNM1|A0A9D1LNM1_9FIRM/1-211
tr|A0A496R7T1|A0A496R7T1_UNCSP/1-223
tr|A0A946JYB4|A0A946JYB4_9GAMM/1-216
tr|A0A845LMU6|A0A845LMU6_9FIRM/1-218
tr|A0A2E7VR85|A0A2E7VR85_9FLAO/1-215
tr|A0A2D9ANU6|A0A2D9ANU6_9FLAO/1-215
tr|A0A2E5XDX2|A0A2E5XDX2_9FLAO/1-215
tr|A0A2E1CED5|A0A2E1CED5_9FLAO/1-214
tr|A0A2E8TM48|A0A2E8TM48_9FLAO/1-215
tr|A0A938HDM9|A0A938HDM9_9SPHN/1-214
tr|A0A9D8BDV6|A0A9D8BDV6_9BACT/1-215
tr|A0A3D5ZF02|A0A3D5ZF02_9FLAO/1-221
tr|A0A2E1FNJ6|A0A2E1FNJ6_9FLAO/1-222
tr|A0A2E4TCF1|A0A2E4TCF1_9FLAO/1-221
tr|A0A975A957|A0A975A957_9BACT/1-214
tr|A0A4R0Z662|A0A4R0Z662_9BACL/1-216
tr|A0A917JNG2|A0A917JNG2_9GAMM/1-224
tr|I1E091|I1E091_9GAMM/1-215
tr|A0A2N1YA76|A0A2N1YA76_9GAMM/1-215
tr|A0A350NTD8|A0A350NTD8_9FLAO/1-219
tr|A0A353VHL3|A0A353VHL3_9FLAO/1-215
tr|A0A449AS24|A0A449AS24_9MOLU/1-54
```

```
tr|A0A7Y2F3L8|A0A7Y2F3L8_9FLAO/1-216
tr|A0A2H0IKC0|A0A2H0IKC0_9BACT/1-216
tr|A0A1V6GZJ9|A0A1V6GZJ9_9BACT/1-220
tr|A0A7X7TFI7|A0A7X7TFI7_9THEM/1-220
tr|A0A972E3J3|A0A972E3J3_9THEM/1-207
tr|A0A7V6UY75|A0A7V6UY75_9CLOT/1-219
tr|A0A920YR95|A0A920YR95_9CLOT/1-219
tr|A0A7V4I8D9|A0A7V4I8D9_9BACT/1-225
tr|A0A401UE11|A0A401UE11_9BACT/1-224
tr|A0A958VUZ0|A0A958VUZ0_9BACT/1-214
tr|A0A952N2P4|A0A952N2P4_9BACT/1-213
tr|A0A9E3KYW4|A0A9E3KYW4_9BACT/1-213
tr|A0A952MRX6|A0A952MRX6_9BACT/1-214
tr|A0A9E3KGW0|A0A9E3KGW0_9BACT/1-213
tr|A0A952M922|A0A952M922_9BACT/1-214
tr|A0A3D6EV76|A0A3D6EV76_9BACT/1-206
tr|A0A952MEI8|A0A952MEI8_9BACT/1-214
tr|A0A3B8W6Q6|A0A3B8W6Q6_9BACT/1-214
tr|A0A977L9K7|A0A977L9K7_9BACT/1-214
tr|A0A2W4NU48|A0A2W4NU48_9BACT/1-220
tr|A0A953K1W0|A0A953K1W0_9BACT/1-214
tr|A0A968WL72|A0A968WL72_9BACT/1-208
tr|A0A364Y397|A0A364Y397_9BACT/1-230
tr|A0A385SXQ8|A0A385SXQ8_9BACT/1-214
tr|A0A7Y4XL45|A0A7Y4XL45_9BACT/1-214
tr|A0A6N4EJI3|A0A6N4EJI3_9RHOO/1-214
tr|A0A936L3Z9|A0A936L3Z9_9BACT/1-221
tr|A0A8J3CGT3|A0A8J3CGT3_9BURK/1-217
tr|A0A3B9UDD1|A0A3B9UDD1_9FLAO/1-205
tr|A0A2Z4G7M3|A0A2Z4G7M3_9BACT/1-216
tr|A0A8U0M4G2|A0A8U0M4G2_9BACT/1-224
tr|A0A926H603|A0A926H603_9BACT/1-225
tr|A0A521GEU9|A0A521GEU9_9BACT/1-229
tr|A0A7J5U5G6|A0A7J5U5G6_9BACT/1-226
tr|A0A3G3GQ14|A0A3G3GQ14_9BACT/1-219
tr|A0A3C0UK34|A0A3C0UK34_9BACT/1-219
tr|A0A976HGK2|A0A976HGK2_9BACT/1-223
tr|A0A976HSX2|A0A976HSX2_9BACT/1-223
tr|A0A9E6RWU5|A0A9E6RWU5_9BACT/1-223
tr|A0A6N6KPX0|A0A6N6KPX0_9BACT/1-225
tr|A0A959J695|A0A959J695_9BACT/1-225
tr|A0A9E2AJ21|A0A9E2AJ21_9BACT/1-228
tr|A0A4D7JWF4|A0A4D7JWF4_9BACT/1-219
tr|A0A4V6BJ72|A0A4V6BJ72_9BACT/1-218
tr|A0A5R9KSD6|A0A5R9KSD6_9BACT/1-214
tr|A0A4R4KK61|A0A4R4KK61_9BACT/1-220
tr|A0A7C9BFK9|A0A7C9BFK9_9BACT/1-218
tr|A0A2Z2KCI3|A0A2Z2KCI3_9BACL/1-235
tr|A0A923LLQ1|A0A923LLQ1_9FIRM/1-218
tr|A0A413LV85|A0A413LV85_9CLOT/1-216
tr|D3AQX6|D3AQX6_9CLOT/1-217
tr|A0A3R6VHX4|A0A3R6VHX4_9CLOT/1-212
tr|A0A4S2HJU0|A0A4S2HJU0_9FIRM/1-214
tr|A0A7U9SB37|A0A7U9SB37_9FIRM/1-214
tr|A0A2E3SQ06|A0A2E3SQ06_9BACT/1-219
tr|A0A3B9N1G5|A0A3B9N1G5_9BACT/1-219
tr|A0A2N5ZZ95|A0A2N5ZZ95_9BACT/1-218
tr|A0A3C1JW13|A0A3C1JW13_9BACT/1-220
tr|A0A3N7FCC9|A0A3N7FCC9_9SPHI/1-227
tr|A0A924UN92|A0A924UN92_9SPHI/1-216
tr|A0A847M0W4|A0A847M0W4_9SPIR/1-215
tr|A0A4P7DTR8|A0A4P7DTR8_9SPHI/1-216
tr|A0A7G9BSW5|A0A7G9BSW5_9SPHI/1-213
tr|A0A420W1E6|A0A420W1E6_9SPHI/1-215
tr|A0A349R4B3|A0A349R4B3_9SPHI/1-215
tr|A0A3C0Z8Q9|A0A3C0Z8Q9_9SPHI/1-215
tr|A0A6G7BZS2|A0A6G7BZS2_9SPHI/1-215
tr|A0A363NPE6|A0A363NPE6_9SPHI/1-215
tr|A0A7G5E3Q0|A0A7G5E3Q0_SPHMU/1-215
tr|A0A349VWP8|A0A349VWP8_9SPHI/1-215
tr|A0A374A9J8|A0A374A9J8_9BACT/1-220
tr|A0A5B3G4F1|A0A5B3G4F1_9BACT/1-220
tr|A0A943ATQ8|A0A943ATQ8_9BACT/1-220
tr|A0A9D1TYN6|A0A9D1TYN6_9BACT/1-223
tr|A0A5C7FMH0|A0A5C7FMH0_9BACI/1-219
tr|A7VQZ8|A7VQZ8_9FIRM/1-229
tr|A0A970EWF7|A0A970EWF7_9FIRM/1-216
tr|A0A3M8DRS7|A0A3M8DRS7_9BACL/1-217
tr|A0A3D3FK06|A0A3D3FK06_9SPIR/1-219
tr|A0A7X5EGM7|A0A7X5EGM7_9FIRM/1-185
tr|D7UWL3|D7UWL3_LISGR/1-216
tr|A0A3D0WVU6|A0A3D0WVU6_9FIRM/1-228
tr|A0A353RC55|A0A353RC55_9FIRM/1-228
tr|A0A357Z3J2|A0A357Z3J2_9FIRM/1-228
tr|A0A9D1DET0|A0A9D1DET0_9FIRM/1-216
tr|A0A9D1FRP0|A0A9D1FRP0_9FIRM/1-216
tr|A0A942Z2S8|A0A942Z2S8_9BACI/1-215
tr|A0A942TF39|A0A942TF39_9BACI/1-215
tr|A0A942TS19|A0A942TS19_9BACI/1-215
tr|A0A424Y5T2|A0A424Y5T2_9CLOT/1-215
tr|A0A250F6E6|A0A250F6E6_CAPSP/1-210
tr|A0A250FIR9|A0A250FIR9_CAPSP/1-210
tr|A0A2A3N7F2|A0A2A3N7F2_CAPSP/1-210
tr|A0A1Z4BLH8|A0A1Z4BLH8_9FLAO/1-210
tr|A0A2S0LC47|A0A2S0LC47_9FLAO/1-202
tr|J0WN13|J0WN13_9FLAO/1-207
tr|I9DYT9|I9DYT9_9FLAO/1-207
tr|A0A7H8YR96|A0A7H8YR96_9FLAO/1-207
tr|A0A2S0LII3|A0A2S0LII3_9FLAO/1-207
tr|A0A7Y8VU48|A0A7Y8VU48_9FLAO/1-207
tr|E4MT34|E4MT34_CAPOC/1-207
tr|A0A9D9TMQ0|A0A9D9TMQ0_9BACT/1-213
tr|A0A5B0FKB9|A0A5B0FKB9_9BACT/1-218
tr|A0A3D8Y503|A0A3D8Y503_9BACT/1-222
tr|A0A4R5DQJ2|A0A4R5DQJ2_9BACT/1-222
tr|A0A9E8NBJ2|A0A9E8NBJ2_9BACT/1-219
tr|A0A5Q0Q6K4|A0A5Q0Q6K4_9SPHI/1-216
tr|A0A6N8L2H8|A0A6N8L2H8_9SPHI/1-217
tr|A0A9E2SE81|A0A9E2SE81_9BACT/1-228
tr|A0A2T8HFM6|A0A2T8HFM6_9SPHI/1-219
tr|A0A7K1Y981|A0A7K1Y981_9SPHI/1-224
tr|A0A7G8V2K3|A0A7G8V2K3_9SPHI/1-226
tr|A0A929PLQ8|A0A929PLQ8_9SPHI/1-221
tr|A0A553F0Z8|A0A553F0Z8_9BACT/1-220
tr|A0A7H0VIS6|A0A7H0VIS6_9FLAO/1-219
tr|A0A2D8SHY2|A0A2D8SHY2_9FLAO/1-219
tr|A0A353FEP6|A0A353FEP6_9FLAO/1-219
tr|A0A2G4HBJ0|A0A2G4HBJ0_9FLAO/1-211
tr|A0A3B8QVX5|A0A3B8QVX5_9FLAO/1-216
tr|A0A937KHR9|A0A937KHR9_9FLAO/1-219
tr|A0A966P5J6|A0A966P5J6_9BACT/1-219
tr|A0A975I9Y1|A0A975I9Y1_9FLAO/1-214
tr|A0A9E4LBX6|A0A9E4LBX6_9FLAO/1-214
```

```
tr|A0A7Y6YG29|A0A7Y6YG29_9BACT/1-216
tr|A0A6L3ZCX4|A0A6L3ZCX4_9FLAO/1-216
tr|A0A6N6RHI6|A0A6N6RHI6_9FLAO/1-216
tr|A0A850L8H3|A0A850L8H3_9BACT/1-216
tr|A0A6M1T502|A0A6M1T502_9BACT/1-218
tr|A0A8T5XWD6|A0A8T5XWD6_UNCFI/1-216
tr|A0A969HMJ8|A0A969HMJ8_9BACT/1-221
tr|A0A9E5QSU3|A0A9E5QSU3_9BACT/1-219
tr|A0A7C3IR57|A0A7C3IR57_9SPIR/1-220
tr|A0A7Y4Z5E3|A0A7Y4Z5E3_9BACT/1-218
tr|A0A7Y7TTK7|A0A7Y7TTK7_9BACT/1-218
tr|A0A7X9HJQ5|A0A7X9HJQ5_9BACT/1-223
tr|A0A352F6V8|A0A352F6V8_9BACT/1-213
tr|A0A970UT58|A0A970UT58_9BACT/1-213
tr|A0A5B9Y9C2|A0A5B9Y9C2_9FIRM/1-216
tr|A0A0B0HSI4|A0A0B0HSI4_9BACL/1-139
tr|A0A370ANK3|A0A370ANK3_9SPIO/1-228
tr|A0A5C1QHN2|A0A5C1QHN2_9SPIO/1-223
tr|A0A2M8H0S5|A0A2M8H0S5_9VIBR/1-208
tr|A0A3D8ZUE7|A0A3D8ZUE7_9ENTE/1-222
tr|A0A553S872|A0A553S872_ENTAV/1-222
tr|A0A4P8KF62|A0A4P8KF62_ENTAV/1-222
tr|A0A6I7ZIX6|A0A6I7ZIX6_9ENTE/1-222
tr|A0A437UR67|A0A437UR67_ENTAV/1-222
tr|A0A2N8PXT0|A0A2N8PXT0_ENTAV/1-222
tr|A0A833HLM1|A0A833HLM1_9CLOT/1-217
tr|A0A3C0BFB4|A0A3C0BFB4_9SPIR/1-215
tr|A0A842HHY8|A0A842HHY8_9BACT/1-224
tr|A0A9D2TSM7|A0A9D2TSM7_9FIRM/1-216
tr|A0A975F5M0|A0A975F5M0_9SPIR/1-223
tr|A0A923L1S7|A0A923L1S7_9FIRM/1-225
tr|A0A2S0KNZ4|A0A2S0KNZ4_9FIRM/1-224
tr|A0A416RSG0|A0A416RSG0_9CLOT/1-212
tr|A0A090ZA67|A0A090ZA67_PAEMA/1-216
tr|A0A9E1F8R8|A0A9E1F8R8_CLOSP/1-215
tr|A0A2G6LL13|A0A2G6LL13_9FLAO/1-216
tr|A0A959LBH2|A0A959LBH2_9BACT/1-225
tr|A0A959B0R8|A0A959B0R8_9BACT/1-219
tr|A0A959C130|A0A959C130_9BACT/1-234
tr|A0A959AQ81|A0A959AQ81_9BACT/1-223
tr|A0A959HBT6|A0A959HBT6_9BACT/1-220
tr|A0A959HGT9|A0A959HGT9_9BACT/1-212
tr|A0A959BT20|A0A959BT20_9BACT/1-220
tr|A0A959ECK6|A0A959ECK6_9BACT/1-220
tr|A0A7Y5J2U8|A0A7Y5J2U8_9BACT/1-225
tr|A0A7Y5R8I8|A0A7Y5R8I8_9BACT/1-221
tr|A0A959JPP7|A0A959JPP7_9BACT/1-216
tr|A0A2S2DVI8|A0A2S2DVI8_9BACT/1-215
tr|A0A437PWN1|A0A437PWN1_9BACT/1-216
tr|A0A4Q9BAX1|A0A4Q9BAX1_9BACT/1-215
tr|A0A9D9WC15|A0A9D9WC15_9BACT/1-215
tr|A0A1V5W036|A0A1V5W036_9BACT/1-218
tr|A0A7X8J636|A0A7X8J636_9BACT/1-221
tr|A0A2S9J2W6|A0A2S9J2W6_9SPHI/1-219
tr|A0A5D4H4D7|A0A5D4H4D7_9SPHI/1-219
tr|A0A2S9JSW4|A0A2S9JSW4_9SPHI/1-219
tr|A0A6M1NG71|A0A6M1NG71_9SPHI/1-219
tr|A0A368DH08|A0A368DH08_9FLAO/1-219
tr|A0A2D5K6Q6|A0A2D5K6Q6_9FLAO/1-213
tr|A0A2E4Q9L7|A0A2E4Q9L7_9FLAO/1-217
tr|A0A959M4I1|A0A959M4I1_9BACT/1-223
tr|A0A7C5JMT9|A0A7C5JMT9_9BACT/1-216
tr|A0A7V1HX50|A0A7V1HX50_9BACT/1-213
tr|A0A3D5HMK5|A0A3D5HMK5_9BACT/1-220
tr|A0A2E4IA74|A0A2E4IA74_9BACT/1-220
tr|A0A3C0HBF4|A0A3C0HBF4_9BACT/1-220
tr|A0A3M1YXP3|A0A3M1YXP3_9BACT/1-221
tr|A0A959FV44|A0A959FV44_9BACT/1-219
tr|A0A660R2D3|A0A660R2D3_9BACT/1-173
tr|A0A6B2M3A3|A0A6B2M3A3_9BACT/1-224
tr|A0A212JPK8|A0A212JPK8_9FIRM/1-217
tr|A0A7X9IJT9|A0A7X9IJT9_9DELT/1-214
tr|A0A329KVL0|A0A329KVL0_9BACL/1-221
tr|A0A7X8FUD2|A0A7X8FUD2_9SPIR/1-217
tr|A0A970D3C5|A0A970D3C5_9CLOT/1-215
tr|A0A9D2KKB2|A0A9D2KKB2_9FIRM/1-217
tr|A0A7V4YSD7|A0A7V4YSD7_9BACT/1-233
tr|A0A951WCA8|A0A951WCA8_9BACT/1-214
tr|A0A316KFH1|A0A316KFH1_9FLAO/1-210
tr|A0A9D9XCK4|A0A9D9XCK4_9FLAO/1-209
tr|A0A358Y2B0|A0A358Y2B0_9FLAO/1-211
tr|A0A965J266|A0A965J266_9FLAO/1-211
tr|A0A9D9ML39|A0A9D9ML39_9FLAO/1-211
tr|A0A946R7A2|A0A946R7A2_9FLAO/1-222
tr|A0A9E5GPF4|A0A9E5GPF4_9BACT/1-222
tr|A0A4S4C659|A0A4S4C659_9BACL/1-220
tr|A0A2B8A351|A0A2B8A351_9BACT/1-213
tr|A0A7Y1U727|A0A7Y1U727_9FLAO/1-217
tr|A0A7Y7BDT0|A0A7Y7BDT0_9BACT/1-216
tr|A0A9E1CH41|A0A9E1CH41_UNCFI/1-216
tr|A0A9E3KH38|A0A9E3KH38_9BACT/1-220
tr|A0A2D5BAW2|A0A2D5BAW2_9FLAO/1-223
tr|A0A2A2GE36|A0A2A2GE36_9BACT/1-220
tr|A0A967SPN4|A0A967SPN4_9BACT/1-220
tr|A0A2D9HQ65|A0A2D9HQ65_9FLAO/1-218
tr|A0A3D1BF62|A0A3D1BF62_9FLAO/1-218
tr|A0A2E8WYL5|A0A2E8WYL5_9FLAO/1-218
tr|A0A3C0DYN1|A0A3C0DYN1_9FLAO/1-218
tr|A0A351DUD5|A0A351DUD5_9FLAO/1-220
tr|A0A2D5AGA5|A0A2D5AGA5_9FLAO/1-221
tr|A0A357K704|A0A357K704_9FLAO/1-221
tr|A0A3D3A770|A0A3D3A770_9FLAO/1-221
tr|A0A2A5FKY6|A0A2A5FKY6_9BACT/1-221
tr|A0A946WX24|A0A946WX24_9FLAO/1-217
tr|A0A9E1Q8I9|A0A9E1Q8I9_9FLAO/1-217
tr|A0A424QLF9|A0A424QLF9_9FLAO/1-219
tr|A0A2D5E0H1|A0A2D5E0H1_9FLAO/1-219
tr|A0A356MMK9|A0A356MMK9_9FLAO/1-217
tr|A0A3R7USQ0|A0A3R7USQ0_9FLAO/1-217
tr|A0A2E4ZPP6|A0A2E4ZPP6_9FLAO/1-217
tr|A0A352WQU1|A0A352WQU1_9FLAO/1-217
tr|A0A2D7DE54|A0A2D7DE54_9FLAO/1-217
tr|A0A2E7TXL2|A0A2E7TXL2_9FLAO/1-217
tr|A0A2E9YF63|A0A2E9YF63_9FLAO/1-219
tr|A0A946EYD3|A0A946EYD3_9FLAO/1-219
tr|A0A350I6W6|A0A350I6W6_9BACT/1-220
tr|A0A2T4WJW5|A0A2T4WJW5_9BACT/1-219
tr|A0A7Y2BXK4|A0A7Y2BXK4_9BACT/1-217
tr|A0A7Y3A6F1|A0A7Y3A6F1_9BACT/1-223
tr|A0A7K3WLK0|A0A7K3WLK0_9FLAO/1-216
tr|A0A5C6RJT1|A0A5C6RJT1_9BACT/1-222
tr|A0A959KTL3|A0A959KTL3_9BACT/1-226
tr|A0A9D7F6Q3|A0A9D7F6Q3_9BACT/1-224
```

```
tr|A0A959LCL0|A0A959LCL0_9BACT/1-221
tr|A0A934YE48|A0A934YE48_9BACT/1-225
tr|A0A9E0I9M5|A0A9E0I9M5_9BACT/1-225
tr|A0A959DRB0|A0A959DRB0_9BACT/1-225
tr|A0A7D5RGP2|A0A7D5RGP2_9BACT/1-225
tr|A0A959D2E4|A0A959D2E4_9BACT/1-225
tr|A0A6B2G7Y8|A0A6B2G7Y8_9PROT/1-217
tr|A0A2M7VNN4|A0A2M7VNN4_9FLAO/1-217
tr|A0A7Y1YGF5|A0A7Y1YGF5_9BACT/1-224
tr|A0A2S5A747|A0A2S5A747_9SPHI/1-220
tr|A0A7C5GAX4|A0A7C5GAX4_9BACT/1-219
tr|A0A9D0XKA8|A0A9D0XKA8_9BACT/1-219
tr|A0A848SYF7|A0A848SYF7_9BACT/1-224
tr|A0A7G3FPT0|A0A7G3FPT0_9BACT/1-216
tr|A0A9D9AXY9|A0A9D9AXY9_9BACT/1-216
tr|A0A848SWL5|A0A848SWL5_9BACT/1-220
tr|A0A9E7B366|A0A9E7B366_9BACT/1-227
tr|A0A7Y3H325|A0A7Y3H325_9BACT/1-217
tr|A0A3D4AJU1|A0A3D4AJU1_9BACT/1-227
tr|A0A2Z4UEY3|A0A2Z4UEY3_9FIRM/1-220
tr|A0A844J729|A0A844J729_9CLOT/1-224
tr|A0A2T3JPU3|A0A2T3JPU3_9GAMM/1-216
tr|A0A242FTX5|A0A242FTX5_ENTFC/1-225
tr|A0A8E2RPH8|A0A8E2RPH8_ENTFC/1-225
tr|A0A133CUD9|A0A133CUD9_ENTFC/1-225
tr|A0A828QB99|A0A828QB99_ENTFC/1-225
tr|A0A3D5CSG5|A0A3D5CSG5_ENTSX/1-225
tr|A0A6B3Q6F8|A0A6B3Q6F8_ENTFC/1-225
tr|A0A934XGA4|A0A934XGA4_9BACT/1-232
tr|A0A9D9U6K1|A0A9D9U6K1_9BACT/1-233
tr|A0A935M2A2|A0A935M2A2_9BACT/1-229
tr|A0A935ZW79|A0A935ZW79_9BACT/1-229
tr|A0A9D7SC64|A0A9D7SC64_9BACT/1-229
tr|A0A2P6MKR0|A0A2P6MKR0_9BACI/1-227
tr|A0A969PXP3|A0A969PXP3_9BACI/1-222
tr|A0A959G1Q5|A0A959G1Q5_9BACT/1-224
tr|A0A6N7SIK5|A0A6N7SIK5_9CLOT/1-224
tr|A0A4Q6XYL9|A0A4Q6XYL9_9SPHI/1-218
tr|A0A7T4UK34|A0A7T4UK34_9SPHI/1-215
tr|A0A7X5YWV5|A0A7X5YWV5_9SPHI/1-215
tr|A0A7K0BAU2|A0A7K0BAU2_9SPHI/1-215
tr|A0A3D5ADR6|A0A3D5ADR6_9SPHI/1-215
tr|A0A976GS29|A0A976GS29_9BACT/1-216
tr|A0A944T5I3|A0A944T5I3_9FLAO/1-217
tr|A0A4Q1C2K5|A0A4Q1C2K5_9BACT/1-216
tr|A0A3R7UT75|A0A3R7UT75_9FLAO/1-216
tr|A0A3D4CD23|A0A3D4CD23_9FLAO/1-216
tr|A0A965I0P5|A0A965I0P5_9PROT/1-219
tr|A0A6N9MLW7|A0A6N9MLW7_ECOLX/1-219
tr|A0A7W3ERE7|A0A7W3ERE7_ESCFE/1-219
tr|A0A2K3TT93|A0A2K3TT93_ECOLX/1-219
tr|A0A3J9BD20|A0A3J9BD20_ECOLX/1-219
tr|A0A4P8C1K3|A0A4P8C1K3_ECOLX/1-219
tr|A0A5B9ACD8|A0A5B9ACD8_ECOLX/1-219
tr|A0A066R5B8|A0A066R5B8_ECOLX/1-219
tr|A0A6H2GEX4|A0A6H2GEX4_9ESCH/1-219
tr|A0A836ZEJ4|A0A836ZEJ4_ECOLX/1-219
tr|A0A8B3KJN0|A0A8B3KJN0_SHISO/1-219
tr|A7ZLE6|A7ZLE6_ECO24/1-219
tr|D8E798|D8E798_ECOLX/1-219
tr|E9TF70|E9TF70_ECOLX/1-219
tr|F4NGG0|F4NGG0_ECOLX/1-219
tr|I2UFD4|I2UFD4_ECOLX/1-219
tr|A0A6D0ILD8|A0A6D0ILD8_ECOLX/1-188
tr|A0A7Z1J0T0|A0A7Z1J0T0_SHIFL/1-219
tr|D7XPQ4|D7XPQ4_ECOLX/1-219
tr|E1J7Q7|E1J7Q7_ECOLX/1-219
tr|E6BN25|E6BN25_ECOLX/1-219
tr|A0A3Y3MK13|A0A3Y3MK13_ECOLX/1-219
tr|A0A829CNM0|A0A829CNM0_ECOLX/1-219
tr|J7QE14|J7QE14_ECOLX/1-219
tr|A0A7H9LSV0|A0A7H9LSV0_ECOLX/1-219
tr|A0A8S7NEG7|A0A8S7NEG7_ECOLX/1-219
tr|A0A827Q5E0|A0A827Q5E0_ECOLX/1-219
tr|A0A0J2BWA0|A0A0J2BWA0_ECOLX/1-219
tr|A0A4Y8GPX4|A0A4Y8GPX4_ECOLX/1-219
tr|A0A641BIY2|A0A641BIY2_ECOLX/1-219
tr|A0A7I0KQV0|A0A7I0KQV0_ECOLX/1-219
tr|A0A828UA74|A0A828UA74_ECOLX/1-219
tr|A0A0J2BLH3|A0A0J2BLH3_ECOLX/1-219
tr|A0A826L6Y6|A0A826L6Y6_ECOLX/1-219
tr|A0A4U9THK9|A0A4U9THK9_ECOLX/1-219
tr|A0A3L4JQG4|A0A3L4JQG4_ECOLX/1-219
tr|A0A775TCT2|A0A775TCT2_ECOLX/1-219
tr|A0A4P7TQP9|A0A4P7TQP9_SHIFM/1-219
tr|F5NTT1|F5NTT1_SHIFL/1-219
tr|A0A376W5Z1|A0A376W5Z1_ECOLX/1-219
tr|A0A376LR95|A0A376LR95_ECOLX/1-219
tr|D8A1X6|D8A1X6_ECOMS/1-219
tr|A0A241QR90|A0A241QR90_ECOLX/1-219
tr|D7XYH0|D7XYH0_ECOM1/1-219
tr|A0A232PR10|A0A232PR10_ECOLX/1-219
tr|B1LGB1|B1LGB1_ECOSM/1-219
tr|A0A0B1I124|A0A0B1I124_ECOLX/1-219
tr|E1IVD0|E1IVD0_ECOLX/1-219
tr|A0A6D0IFH1|A0A6D0IFH1_ECOLX/1-207
tr|A0A8H9DBH5|A0A8H9DBH5_SHIFL/1-219
tr|A0A0H3PQ51|A0A0H3PQ51_ECO5C/1-219
tr|C3TBL2|C3TBL2_ECOLX/1-219
tr|A0A891SFN0|A0A891SFN0_ECOLX/1-219
tr|A0A0D8VW63|A0A0D8VW63_ECOLX/1-219
tr|A0A8H9DZQ0|A0A8H9DZQ0_SHIBO/1-219
tr|A0A2X3JUI0|A0A2X3JUI0_ECOLX/1-66
tr|A0A2H4TLT3|A0A2H4TLT3_ECOLX/1-219
tr|E2QLS6|E2QLS6_ECOLX/1-219
tr|A0A8E0KTH1|A0A8E0KTH1_ECOLX/1-219
tr|A0A839B003|A0A839B003_ECOLX/1-219
tr|A0A7W4KMM5|A0A7W4KMM5_9ESCH/1-219
tr|A0A793XWM6|A0A793XWM6_ECOLX/1-219
tr|A0A6L8Z4X1|A0A6L8Z4X1_ECOLX/1-219
tr|A0A6C8T533|A0A6C8T533_ECOLX/1-219
tr|A0A4Q9CF35|A0A4Q9CF35_ECOLX/1-219
tr|A0A3U5W9X1|A0A3U5W9X1_SHIBO/1-219
tr|A0A398Q2S1|A0A398Q2S1_SHIBO/1-219
tr|A0A377MXQ4|A0A377MXQ4_ECOLX/1-219
tr|A0A377D3R6|A0A377D3R6_ECOLX/1-169
tr|A0A376K2P9|A0A376K2P9_ECOLX/1-219
tr|A0A2X1PMM3|A0A2X1PMM3_ECOLX/1-219
tr|A0A2G9A6G1|A0A2G9A6G1_ECOLX/1-219
tr|A0A2A6QCH5|A0A2A6QCH5_ECOLX/1-219
tr|A0A2A3VWU5|A0A2A3VWU5_ECOLX/1-219
tr|A0A246NY93|A0A246NY93_ECOLX/1-219
tr|A0A1M0D2S2|A0A1M0D2S2_ECOLX/1-219
```

tr|A0A192EN96|A0A192EN96_ECOLX/1-219
tr|A0A0H0KPT7|A0A0H0KPT7_ECOLX/1-219
tr|A0A0B1FYS1|A0A0B1FYS1_ECOLX/1-219
tr|A0A0A1A3I6|A0A0A1A3I6_ECOLX/1-219
tr|A0A085NXU5|A0A085NXU5_ECOLX/1-219
sp|P77366|PGMB_ECOLI/1-219
tr|A0A0E1T3D3|A0A0E1T3D3_ECOLX/1-219
tr|A0A7I6H572|A0A7I6H572_ECOHS/1-219
tr|W8TAK6|W8TAK6_ECOLX/1-219
tr|A0A377DSY8|A0A377DSY8_ECOLX/1-40
tr|A0A4P8S906|A0A4P8S906_9GAMM/1-234
tr|A0A2N5E6X2|A0A2N5E6X2_9GAMM/1-231
tr|A0A2N5E4F2|A0A2N5E4F2_9GAMM/1-233
tr|A0A2N5EPP8|A0A2N5EPP8_9GAMM/1-231
tr|A0A376RJ72|A0A376RJ72_ECOLX/1-88
tr|A0A2P8VLA1|A0A2P8VLA1_9ENTR/1-227
tr|A0A2T7ATI8|A0A2T7ATI8_9ENTR/1-227
tr|A0A2T7B494|A0A2T7B494_9ENTR/1-225
tr|A0A423XTY2|A0A423XTY2_9ENTR/1-225
tr|A0A7V7URV8|A0A7V7URV8_CROSK/1-225
tr|A0A855WS62|A0A855WS62_CROSK/1-225
tr|A0A3R9GB97|A0A3R9GB97_9ENTR/1-227
tr|A0A5J6WAV5|A0A5J6WAV5_9ENTR/1-227
tr|A0A3C0H257|A0A3C0H257_9ENTR/1-227
tr|A0A5D4YF74|A0A5D4YF74_9ENTR/1-235
tr|A0A6I2I8T9|A0A6I2I8T9_9ENTR/1-226
tr|A0A8H9YRV2|A0A8H9YRV2_9PSED/1-226
tr|A0A3S5XQJ6|A0A3S5XQJ6_LELAM/1-226
tr|A0A085AIE2|A0A085AIE2_9ENTR/1-226
tr|A0A329HJY7|A0A329HJY7_9ENTR/1-226
tr|A0A4R0GCP1|A0A4R0GCP1_9ENTR/1-224
tr|A0A5Q2K339|A0A5Q2K339_9ENTR/1-224
tr|D2ZCV5|D2ZCV5_9ENTR/1-224
tr|A0A7D6YL95|A0A7D6YL95_9ENTR/1-221
tr|A0A443UD69|A0A443UD69_ENTCL/1-221
tr|A0A8B2SF55|A0A8B2SF55_9ENTR/1-221
tr|A0A443UQ96|A0A443UQ96_ENTCL/1-221
tr|A0A855M287|A0A855M287_9ENTR/1-221
tr|A0A2L2JIW3|A0A2L2JIW3_9ENTR/1-221
tr|A0A330F588|A0A330F588_ENTCL/1-221
tr|A0A8B3JW64|A0A8B3JW64_9ENTR/1-221
tr|A0A443XC50|A0A443XC50_ENTCL/1-221
tr|A0A5C1C0T5|A0A5C1C0T5_9ENTR/1-221
tr|A0A330GEA5|A0A330GEA5_ENTCL/1-221
tr|A0A2T9UYP9|A0A2T9UYP9_9ENTR/1-221
tr|A0A8I1FY10|A0A8I1FY10_ENTAS/1-221
tr|A0A2S4SQP4|A0A2S4SQP4_9ENTR/1-222
tr|A0A2U3F5K8|A0A2U3F5K8_9ENTR/1-222
tr|A0A411GH31|A0A411GH31_ENTCL/1-221
tr|A0A263VKD5|A0A263VKD5_ENTAS/1-221
tr|A0A808IFR6|A0A808IFR6_9ENTR/1-221
tr|A0A7W3CCP7|A0A7W3CCP7_ENTAS/1-221
tr|A0A8B2UNU1|A0A8B2UNU1_9ENTR/1-221
tr|A0A330DC03|A0A330DC03_9ENTR/1-224
tr|A0A2W0GBR0|A0A2W0GBR0_9ENTR/1-224
tr|A0A8B2Y1N9|A0A8B2Y1N9_9ENTR/1-224
tr|A0A7H8UGH1|A0A7H8UGH1_ENTCL/1-221
tr|A0A4Q4A4R6|A0A4Q4A4R6_ENTCL/1-221
tr|A0A4R0FUA8|A0A4R0FUA8_9ENTR/1-221
tr|A0A3A3ZT51|A0A3A3ZT51_9ENTR/1-221
tr|A0A7G3F1R5|A0A7G3F1R5_ENTCL/1-221
tr|A0A801DRF0|A0A801DRF0_9ENTR/1-222
tr|A0A7T9PDM7|A0A7T9PDM7_9ENTR/1-221
tr|A0A431SNY7|A0A431SNY7_9ENTR/1-221
tr|A0A7H9JTK9|A0A7H9JTK9_ENTCL/1-221
tr|A0A9E7BFV8|A0A9E7BFV8_9ENTR/1-221
tr|A0A7T1QU30|A0A7T1QU30_9ENTR/1-221
tr|F5RXW1|F5RXW1_9ENTR/1-221
tr|A0A8B3V1I0|A0A8B3V1I0_9ENTR/1-221
tr|A0A5Q5DZV2|A0A5Q5DZV2_9ENTR/1-221
tr|A0A6G4MKR0|A0A6G4MKR0_9ENTR/1-221
tr|A0A4Q2QUN4|A0A4Q2QUN4_9ENTR/1-221
tr|A0A6L3Y319|A0A6L3Y319_9ENTR/1-221
tr|A0A2J0QTU2|A0A2J0QTU2_9ENTR/1-221
tr|A0A5B7XTA5|A0A5B7XTA5_9ENTR/1-224
tr|A0A9J6Q2A3|A0A9J6Q2A3_9ENTR/1-225
tr|A0A839BFR9|A0A839BFR9_9ENTR/1-225
tr|A0A4V6JJJ6|A0A4V6JJJ6_9ENTR/1-225
tr|A0A6G7TFG1|A0A6G7TFG1_9ENTR/1-225
tr|A0A3R9W614|A0A3R9W614_9ENTR/1-225
tr|A0A7W2V3P6|A0A7W2V3P6_9ENTR/1-225
tr|A0A3R9A122|A0A3R9A122_ENTCL/1-225
tr|A0A851G6K3|A0A851G6K3_9ENTR/1-225
tr|A0A7V7M9B6|A0A7V7M9B6_ENTCL/1-225
tr|A0A7Z1J8A4|A0A7Z1J8A4_ENTCL/1-225
tr|A0A8H9S660|A0A8H9S660_ENTCL/1-225
tr|A0A2Y9U023|A0A2Y9U023_9GAMM/1-227
tr|A0A4Q9ECR5|A0A4Q9ECR5_9GAMM/1-225
tr|A0A2A2MEF9|A0A2A2MEF9_9GAMM/1-225
tr|A0A3G2I9R5|A0A3G2I9R5_9ENTR/1-225
tr|A0A085GDN2|A0A085GDN2_9ENTR/1-225
tr|A0A3A5K0P0|A0A3A5K0P0_9ENTR/1-225
tr|A0A702BM40|A0A702BM40_SALBN/1-219
tr|A0A750P494|A0A750P494_SALER/1-219
tr|A0A5U3DMS3|A0A5U3DMS3_SALER/1-219
tr|A0A248K990|A0A248K990_SALBN/1-219
tr|A0A8F8AXV8|A0A8F8AXV8_SALBN/1-219
tr|A0A5Y3AZ91|A0A5Y3AZ91_SALER/1-219
tr|A0A702DCF4|A0A702DCF4_SALDZ/1-219
tr|A0A735RFA2|A0A735RFA2_SALDZ/1-219
tr|A0A8F5RKZ9|A0A8F5RKZ9_SALDZ/1-219
tr|A0A7U6BK13|A0A7U6BK13_SALER/1-219
tr|A0A5Y9Q3N3|A0A5Y9Q3N3_SALER/1-219
tr|A0A5U3EHZ8|A0A5U3EHZ8_SALET/1-219
tr|A0A2I5HGD9|A0A2I5HGD9_SALDZ/1-219
tr|A0A3K0PE37|A0A3K0PE37_SALER/1-219
tr|A0A658B6N0|A0A658B6N0_SALET/1-219
tr|A0A8E9YA46|A0A8E9YA46_SALDZ/1-219
tr|A0A3F3J3L1|A0A3F3J3L1_SALER/1-219
tr|A0A3R0QY55|A0A3R0QY55_SALEN/1-219
tr|A0A3S4K1C8|A0A3S4K1C8_SALER/1-219
tr|A0A3U4W4K1|A0A3U4W4K1_SALDZ/1-219
tr|A0A701UP12|A0A701UP12_SALER/1-219
tr|A0A7U5YGR9|A0A7U5YGR9_SALDZ/1-219
tr|A0A7Z1TCD0|A0A7Z1TCD0_SALET/1-219
tr|A0A5I0BS86|A0A5I0BS86_SALET/1-219
tr|A0A5W2LQ92|A0A5W2LQ92_SALET/1-219
tr|A0A742R6E4|A0A742R6E4_SALER/1-219
tr|A0A5I5DTK4|A0A5I5DTK4_SALET/1-219
tr|A0A5T3NHZ6|A0A5T3NHZ6_SALER/1-219
tr|A0A2C9NZX0|A0A2C9NZX0_SALET/1-219
tr|A0A737H5P2|A0A737H5P2_SALER/1-219
tr|A0A735M785|A0A735M785_SALER/1-219

```
tr|A0A402XKM0|A0A402XKM0_SALER/1-219
tr|A0A737Y0M2|A0A737Y0M2_SALER/1-219
tr|A0A6C7D4Y0|A0A6C7D4Y0_SALER/1-219
tr|A0A701VVQ0|A0A701VVQ0_SALER/1-219
tr|A0A737IN59|A0A737IN59_SALER/1-219
tr|A0A5Y1WD24|A0A5Y1WD24_SALER/1-219
tr|A0A3V8I0E2|A0A3V8I0E2_SALER/1-219
tr|A0A731N980|A0A731N980_SALER/1-219
tr|A0A379Q6B0|A0A379Q6B0_SALER/1-219
tr|A0A379QZK5|A0A379QZK5_SALER/1-219
tr|A0A3U3WLW6|A0A3U3WLW6_SALER/1-219
tr|A0A603B752|A0A603B752_SALER/1-219
tr|A0A701V3C6|A0A701V3C6_SALER/1-219
tr|A0A729AUB3|A0A729AUB3_SALER/1-219
tr|A0A732GQC5|A0A732GQC5_SALER/1-219
tr|A0A734HZK1|A0A734HZK1_SALER/1-219
tr|A0A735PYC1|A0A735PYC1_SALER/1-219
tr|A0A379QL84|A0A379QL84_SALER/1-219
tr|A0A5Y3USH2|A0A5Y3USH2_SALER/1-219
tr|A0A6C7CC70|A0A6C7CC70_SALER/1-219
tr|A0A5Y2QE11|A0A5Y2QE11_SALER/1-219
tr|A0A737EWT0|A0A737EWT0_SALER/1-219
tr|A0A753A1F6|A0A753A1F6_SALER/1-219
tr|A0A6L3RX51|A0A6L3RX51_CROSK/1-225
tr|A0A2P5GV12|A0A2P5GV12_9ENTR/1-220
tr|A0A4P7IZQ4|A0A4P7IZQ4_9ENTR/1-222
tr|A0A0P8LZL1|A0A0P8LZL1_CITFR/1-222
tr|A0A0D7M0W9|A0A0D7M0W9_CITFR/1-222
tr|A0A8B5QF49|A0A8B5QF49_9ENTR/1-222
tr|A0A7D6TDF7|A0A7D6TDF7_CITFR/1-222
tr|A0A7W3F0Y6|A0A7W3F0Y6_CITFR/1-222
tr|A0A4U6GZI2|A0A4U6GZI2_9ENTR/1-222
tr|A0A5P2MEC7|A0A5P2MEC7_9ENTR/1-222
tr|A0A6I5AJV4|A0A6I5AJV4_9ENTR/1-222
tr|A0A7X1BM73|A0A7X1BM73_9ENTR/1-222
tr|A0A2Z3X7P5|A0A2Z3X7P5_9ENTR/1-222
tr|D4B909|D4B909_9ENTR/1-222
tr|A0A4P6WKL8|A0A4P6WKL8_9ENTR/1-222
tr|A0A2S4QAX9|A0A2S4QAX9_CITFR/1-222
tr|A0A7L6IFY2|A0A7L6IFY2_9ENTR/1-222
tr|A0A7L6U242|A0A7L6U242_9ENTR/1-222
tr|A0A7L6JME2|A0A7L6JME2_9ENTR/1-222
tr|A0A7L6DSP3|A0A7L6DSP3_9ENTR/1-222
tr|A0A7W2VJ41|A0A7W2VJ41_9ENTR/1-222
tr|A0A2I8S6S4|A0A2I8S6S4_9ENTR/1-222
tr|A0A1R0FUW5|A0A1R0FUW5_CITBR/1-222
tr|A0A7L6BSF2|A0A7L6BSF2_9ENTR/1-222
tr|A0A7L6PIS1|A0A7L6PIS1_9ENTR/1-222
tr|A0A7W3DXZ1|A0A7W3DXZ1_9ENTR/1-222
tr|A0A2S4S1C1|A0A2S4S1C1_CITAM/1-222
tr|A0A379SKS1|A0A379SKS1_SALER/1-101
tr|A0A7T8FKZ1|A0A7T8FKZ1_SALET/1-222
tr|A0A737F4V6|A0A737F4V6_SALER/1-222
tr|A0A740VJH4|A0A740VJH4_SALET/1-222
tr|A0A379S437|A0A379S437_SALER/1-222
tr|A0A3J8Q185|A0A3J8Q185_SALER/1-222
tr|A0A5W1EP41|A0A5W1EP41_SALER/1-222
tr|A0A739C2U3|A0A739C2U3_SALER/1-222
tr|A0A5T8WXD8|A0A5T8WXD8_SALER/1-222
tr|A0A632U6F0|A0A632U6F0_SALER/1-222
tr|A0A741N0G2|A0A741N0G2_SALER/1-222
tr|A0A8H9TVP6|A0A8H9TVP6_9ENTR/1-220
tr|A0A9C7QL52|A0A9C7QL52_CITAM/1-220
tr|A0A6B1TZG6|A0A6B1TZG6_CITAM/1-220
tr|A0A7L6YQQ3|A0A7L6YQQ3_9ENTR/1-220
tr|A0A7L6Z1H4|A0A7L6Z1H4_9ENTR/1-220
tr|A0A7L6UGQ3|A0A7L6UGQ3_9ENTR/1-220
tr|A0A3S7DC38|A0A3S7DC38_9ENTR/1-220
tr|A0A482PKA5|A0A482PKA5_CITRO/1-220
tr|A0A7L6XJV6|A0A7L6XJV6_9ENTR/1-220
tr|A0A7L6YBL4|A0A7L6YBL4_9ENTR/1-220
tr|A0A387ANI5|A0A387ANI5_9LACO/1-221
tr|A0A8I2JKA0|A0A8I2JKA0_9LACO/1-221
tr|A0A1L8CHJ7|A0A1L8CHJ7_9LACO/1-221
tr|A0A410K7G3|A0A410K7G3_LATCU/1-223
tr|A0A4Q4IK05|A0A4Q4IK05_9BACL/1-226
tr|A0A410DCS0|A0A410DCS0_9BACL/1-226
tr|A0A6N9I1X5|A0A6N9I1X5_9LACO/1-224
tr|A0A349MX35|A0A349MX35_9LACO/1-223
tr|A0A5P0ZES5|A0A5P0ZES5_9LACO/1-227
tr|A0A241RPA0|A0A241RPA0_LACPE/1-223
tr|A0A2S9VTX9|A0A2S9VTX9_LACPE/1-223
tr|A0A098R767|A0A098R767_9LACO/1-223
tr|A0A8E4BMR6|A0A8E4BMR6_9LACO/1-223
tr|A0A0G9FDY8|A0A0G9FDY8_LACPN/1-223
tr|A0A2I0YTB3|A0A2I0YTB3_LACPN/1-223
tr|A0A4Z0JHG5|A0A4Z0JHG5_9LACO/1-223
tr|A0A2N7AVQ4|A0A2N7AVQ4_9LACO/1-228
tr|A0A2P4R9E8|A0A2P4R9E8_9LACO/1-227
tr|A0A386PU09|A0A386PU09_9LACO/1-227
tr|A0A5B7SYC4|A0A5B7SYC4_9LACO/1-223
tr|A0A7L7KXZ3|A0A7L7KXZ3_9LACO/1-223
tr|A0A5P0ZWT3|A0A5P0ZWT3_9LACO/1-228
tr|A0A9D2CPT8|A0A9D2CPT8_9LACO/1-228
tr|A0A202FER0|A0A202FER0_9LACO/1-223
tr|A0A210P9X4|A0A210P9X4_9LACO/1-223
tr|A0A5P8QM02|A0A5P8QM02_9LACO/1-223
tr|A0A3R8I4C5|A0A3R8I4C5_9LACO/1-222
tr|A0A5B8TJP0|A0A5B8TJP0_9LACO/1-222
tr|A0A2S9VWN4|A0A2S9VWN4_LACPE/1-224
tr|A0A370ABG3|A0A370ABG3_9LACO/1-228
tr|A0A4Q9Y8T9|A0A4Q9Y8T9_9LACO/1-228
tr|A0A241RP56|A0A241RP56_LACPE/1-225
tr|A0A2S9VSM4|A0A2S9VSM4_LACPE/1-225
tr|A0A8E4G054|A0A8E4G054_9LACO/1-223
tr|A0A2S1T423|A0A2S1T423_LACPN/1-223
tr|A0A0G9GSZ0|A0A0G9GSZ0_LACPN/1-223
tr|A0A2S3U0N0|A0A2S3U0N0_LACPN/1-223
tr|D7VED1|D7VED1_LACPN/1-223
tr|A0A1E3KPU1|A0A1E3KPU1_LACPN/1-223
tr|A0A426D7C6|A0A426D7C6_9LACO/1-223
tr|A0A2R3JPJ8|A0A2R3JPJ8_9LACO/1-223
tr|A0A2R3JWU6|A0A2R3JWU6_9LACO/1-223
tr|A0A2V1N6L2|A0A2V1N6L2_9LACO/1-223
tr|A0A7X3C3N0|A0A7X3C3N0_9LACO/1-223
tr|A0A1Z5IAE1|A0A1Z5IAE1_9LACO/1-223
tr|A0A1Z5J4N5|A0A1Z5J4N5_9LACO/1-223
tr|A0A1Z5IIJ5|A0A1Z5IIJ5_9LACO/1-223
tr|A0A1Z5IKX3|A0A1Z5IKX3_9LACO/1-223
tr|A0A1Z5IX45|A0A1Z5IX45_9LACO/1-223
tr|A0A425XP49|A0A425XP49_9LACO/1-223
tr|A0A6P1E7Y7|A0A6P1E7Y7_LENHI/1-223
tr|C0XMI3|C0XMI3_LENH9/1-223
```

```
tr|A0A844EBW8|A0A844EBW8_9LACO/1-224
tr|A0A5R9CXE9|A0A5R9CXE9_9LACO/1-223
tr|A0A8E1X419|A0A8E1X419_LENBU/1-223
tr|A0A269Y5R9|A0A269Y5R9_9LACO/1-223
tr|A0A6N4HV21|A0A6N4HV21_9LACO/1-224
tr|A0A8D4IVU0|A0A8D4IVU0_LATCU/1-223
tr|A0A7Z6QKQ0|A0A7Z6QKQ0_LEVBR/1-217
tr|A0A2P4R4Z3|A0A2P4R4Z3_9LACO/1-223
tr|A0A9D1QQ21|A0A9D1QQ21_9LACO/1-223
tr|A0A4Z0J4X3|A0A4Z0J4X3_9LACO/1-223
tr|A0A7Z6MLW9|A0A7Z6MLW9_LEVBR/1-223
tr|A0A0D0FHH5|A0A0D0FHH5_LEVBR/1-223
tr|A0A4Q0VGV8|A0A4Q0VGV8_9LACO/1-223
tr|A0A921EEJ0|A0A921EEJ0_9LACO/1-223
tr|A0A1Y6JVE1|A0A1Y6JVE1_9LACO/1-223
tr|A0A9D1U5H2|A0A9D1U5H2_9LACO/1-223
tr|A0A1Z5H3B4|A0A1Z5H3B4_9LACO/1-226
tr|A0A921B3N9|A0A921B3N9_9LACO/1-224
tr|A0A6C2C3D2|A0A6C2C3D2_9LACO/1-223
tr|C5RBI0|C5RBI0_WEIPA/1-220
tr|A0A4Q7IW25|A0A4Q7IW25_WEIPA/1-225
tr|A0A5M9ER82|A0A5M9ER82_WEIPA/1-225
tr|A0A512PSK0|A0A512PSK0_9LACO/1-225
tr|A0A921SQT2|A0A921SQT2_9LACO/1-225
tr|A0A4Y4FZ08|A0A4Y4FZ08_WEIHE/1-222
tr|A0A5B8TP16|A0A5B8TP16_WEIHE/1-223
tr|A0A5M9DPW2|A0A5M9DPW2_WEIPA/1-223
tr|A0A3P2RDN0|A0A3P2RDN0_WEIVI/1-221
tr|A0A7L8CI15|A0A7L8CI15_WEIVI/1-218
tr|A0A5B8SYX1|A0A5B8SYX1_LEUPS/1-224
tr|A0A7L8V0L6|A0A7L8V0L6_9LACO/1-224
tr|A0A3T0TRC7|A0A3T0TRC7_9LACO/1-223
tr|A0A5C4TJQ3|A0A5C4TJQ3_FRUSA/1-223
tr|A0A482PZ42|A0A482PZ42_WEICO/1-223
tr|A0A4Z0RLJ2|A0A4Z0RLJ2_WEICO/1-223
tr|A0A3R5Z475|A0A3R5Z475_WEICO/1-223
tr|A0A4Z0RH56|A0A4Z0RH56_WEICO/1-223
tr|A0A4Z0S096|A0A4Z0S096_WEICO/1-223
tr|A0A846ZFE5|A0A846ZFE5_9LACO/1-224
tr|A0A5B8TBP9|A0A5B8TBP9_LEULA/1-224
tr|A0A5B8T487|A0A5B8T487_LEULA/1-224
tr|A0A6L7A9S0|A0A6L7A9S0_LEULA/1-224
tr|A0A288QX53|A0A288QX53_9LACO/1-224
tr|A0A7G9T472|A0A7G9T472_9LACO/1-224
tr|A0A5P1X0S3|A0A5P1X0S3_9LACO/1-223
tr|A0A7Z0HTH7|A0A7Z0HTH7_9LACO/1-228
tr|A0A843YXB1|A0A843YXB1_LEUME/1-228
tr|A0A5C4TJ58|A0A5C4TJ58_FRUSA/1-223
tr|A0A425VZB5|A0A425VZB5_9ACTN/1-223
tr|A0A6L5GUC1|A0A6L5GUC1_9FIRM/1-223
tr|A0A6N7BRZ5|A0A6N7BRZ5_PEDPE/1-223
tr|A0A7Z0KUT3|A0A7Z0KUT3_9LACO/1-223
tr|A0A843Z3Q5|A0A843Z3Q5_LEUME/1-223
tr|A0A846ZIE6|A0A846ZIE6_9LACO/1-223
tr|A0A5B8T0W8|A0A5B8T0W8_LEUPS/1-223
tr|A0A4Q7IVA9|A0A4Q7IVA9_WEIPA/1-223
tr|C5RBX8|C5RBX8_WEIPA/1-223
tr|A0A4Y4G9L1|A0A4Y4G9L1_WEIHE/1-223
tr|A0A5P8JTN7|A0A5P8JTN7_9LACO/1-228
tr|A0A5Q2NWH1|A0A5Q2NWH1_9LACO/1-229
tr|A0A7T4MYN5|A0A7T4MYN5_PEDPE/1-219
tr|A0A6L5A812|A0A6L5A812_PEDPE/1-219
tr|A0A833T150|A0A833T150_9LACO/1-219
tr|A0A6N7BTY5|A0A6N7BTY5_PEDPE/1-219
tr|A0A4P6YUE6|A0A4P6YUE6_9LACO/1-226
tr|A0A843R7T3|A0A843R7T3_LIMFE/1-222
tr|A0A2V2CWR0|A0A2V2CWR0_LIMFE/1-222
tr|A0A0F4HDH8|A0A0F4HDH8_LIMFE/1-222
tr|C0WVW4|C0WVW4_LIMFE/1-222
tr|D0DSD6|D0DSD6_LIMFE/1-222
tr|A0A855ZNX1|A0A855ZNX1_LIMFE/1-222
tr|A0A3R8G9W4|A0A3R8G9W4_9LACO/1-220
tr|A0A5B8TES4|A0A5B8TES4_9LACO/1-220
tr|A0A5Q2P1C6|A0A5Q2P1C6_9LACO/1-220
tr|A0A556UEA8|A0A556UEA8_9LACO/1-220
tr|A0A4V3RE06|A0A4V3RE06_9LACO/1-221
tr|A0A4Z0GGM8|A0A4Z0GGM8_LACJH/1-220
tr|A0A7L5UJH9|A0A7L5UJH9_9LACO/1-221
tr|A0A2Z6TEH7|A0A2Z6TEH7_9LACO/1-220
tr|A0A315ZN28|A0A315ZN28_LIMMU/1-220
tr|A0A3C1I6Q5|A0A3C1I6Q5_9LACO/1-220
tr|A0A7X9RD99|A0A7X9RD99_9LACO/1-220
tr|A0A9D1U425|A0A9D1U425_9LACO/1-220
tr|C7XVI6|C7XVI6_9LACO/1-219
tr|A0A9D1VGJ3|A0A9D1VGJ3_9LACO/1-220
tr|A0A7W3YLM0|A0A7W3YLM0_9LACO/1-220
tr|A0A7W3TS31|A0A7W3TS31_9LACO/1-220
tr|A0A3M6SAY8|A0A3M6SAY8_LIMRT/1-220
tr|A0A079YT56|A0A079YT56_LIMRT/1-220
tr|A0A828RGY4|A0A828RGY4_LIMRT/1-220
tr|A0A7W3YM02|A0A7W3YM02_9LACO/1-220
tr|A0A2S1EN98|A0A2S1EN98_LIMRT/1-220
tr|A0A855XJG3|A0A855XJG3_LIMRT/1-220
tr|A0A317GGU0|A0A317GGU0_LIMRT/1-220
tr|A0A7Y0SWA9|A0A7Y0SWA9_LIMRT/1-220
tr|A0A7X2KIS1|A0A7X2KIS1_LIMRT/1-220
tr|A0A7W3V2S0|A0A7W3V2S0_9LACO/1-220
tr|A0A6N1EMT9|A0A6N1EMT9_LIMRT/1-220
tr|A0A517D2T8|A0A517D2T8_LIMRT/1-220
tr|A0A1S9AN24|A0A1S9AN24_LIMRT/1-220
tr|A0A081NQZ9|A0A081NQZ9_LIMRT/1-220
tr|A0A8D9S5S6|A0A8D9S5S6_LIMRT/1-220
tr|F8DR87|F8DR87_LIMRS/1-220
tr|Q1KMT6|Q1KMT6_LIMRT/1-220
tr|C8P4W6|C8P4W6_9LACO/1-220
tr|E3C8N0|E3C8N0_9LACO/1-220
tr|A0A5P9QG95|A0A5P9QG95_9LACO/1-221
tr|A0A2J6NNU9|A0A2J6NNU9_9LACO/1-221
tr|A0A9E2KW88|A0A9E2KW88_9LACO/1-220
tr|A0A7W3TYD8|A0A7W3TYD8_9LACO/1-220
tr|C2EU59|C2EU59_9LACO/1-220
tr|A0A9E2KTJ2|A0A9E2KTJ2_9LACO/1-220
tr|C8PC99|C8PC99_9LACO/1-209
tr|A0A6G7B7J2|A0A6G7B7J2_9LACO/1-220
tr|E1NMG2|E1NMG2_9LACO/1-220
tr|E1NT81|E1NT81_9LACO/1-220
tr|A0A916QJS7|A0A916QJS7_9LACO/1-226
tr|A0A3M0PA64|A0A3M0PA64_9LACO/1-225
tr|A0A318M2R7|A0A318M2R7_9LACO/1-225
tr|A0A3M0NXI4|A0A3M0NXI4_9LACO/1-225
tr|A0A2Z3HRQ1|A0A2Z3HRQ1_9LACO/1-225
tr|A0A5P5ZK79|A0A5P5ZK79_9LACO/1-220
tr|A0A6B8YEK7|A0A6B8YEK7_9LACO/1-220
```

tr|A0A558KGE1|A0A558KGE1_9LACO/1-221
tr|A0A558K4C5|A0A558K4C5_9LACO/1-220
tr|A0A133PHP4|A0A133PHP4_LACGS/1-220
tr|D1YL18|D1YL18_LACGS/1-220
tr|A0A2Z6PPL7|A0A2Z6PPL7_9LACO/1-220
tr|A0A6P1EPG1|A0A6P1EPG1_9LACO/1-220
tr|A0A1V3Y1D7|A0A1V3Y1D7_LACGS/1-220
tr|A0A6B2G003|A0A6B2G003_9LACO/1-220
tr|A0A8A4V077|A0A8A4V077_LACGS/1-220
tr|A0A1B3PV14|A0A1B3PV14_LACJH/1-220
tr|A0A1Y4IBJ3|A0A1Y4IBJ3_LACJH/1-220
tr|D0R1V9|D0R1V9_LACJF/1-220
tr|A0A6B9HXS9|A0A6B9HXS9_LACJH/1-220
tr|A0A346MT01|A0A346MT01_LACJH/1-220
tr|C2E3E3|C2E3E3_LACJH/1-220
tr|A0A244CIZ1|A0A244CIZ1_LACJH/1-220
tr|A0A4Z0GDJ5|A0A4Z0GDJ5_LACJH/1-220
tr|D4YU61|D4YU61_9LACO/1-223
tr|A0A2V4EWB1|A0A2V4EWB1_LACHE/1-225
tr|A0A8H9F999|A0A8H9F999_LACHE/1-223
tr|Q5FI05|Q5FI05_LACAC/1-221
tr|A0A809KFM8|A0A809KFM8_LACAI/1-221
tr|A0A437STJ5|A0A437STJ5_9LACO/1-220
tr|A0A1G5WIK1|A0A1G5WIK1_9LACO/1-220
tr|C2EP86|C2EP86_9LACO/1-220
tr|A0A8I1LF53|A0A8I1LF53_9LACO/1-220
tr|A0A854ZDR9|A0A854ZDR9_9LACO/1-220
tr|A0A7X7H2P1|A0A7X7H2P1_9LACO/1-221
tr|A0A3R6BXU4|A0A3R6BXU4_LACAM/1-221
tr|F0THE4|F0THE4_LACAM/1-221
tr|A0A4Q0LSC0|A0A4Q0LSC0_9LACO/1-220
tr|E3R6B5|E3R6B5_9LACO/1-224
tr|A0A109DFH6|A0A109DFH6_9LACO/1-220
tr|C7XJR4|C7XJR4_9LACO/1-224
tr|D0DFQ6|D0DFQ6_9LACO/1-224
tr|D5GZY2|D5GZY2_LACCS/1-220
tr|A0A2A7RIW3|A0A2A7RIW3_9LACO/1-220
tr|A0A5M9Z0U8|A0A5M9Z0U8_9LACO/1-220
tr|A0A558JKE8|A0A558JKE8_LACJE/1-223
tr|D6S3G5|D6S3G5_LACJE/1-223
tr|A0A5N1IFF7|A0A5N1IFF7_LACJE/1-223
tr|A0A6A8MDX6|A0A6A8MDX6_9LACO/1-222
tr|A0A061CMF3|A0A061CMF3_LACDL/1-221
tr|F0HWK4|F0HWK4_LACDL/1-221
tr|A0A4Y3JVH4|A0A4Y3JVH4_9LACO/1-221
tr|A0A844FQ68|A0A844FQ68_9LACO/1-222
tr|A0A1V0Q4M1|A0A1V0Q4M1_LACPA/1-230
tr|A0A5Q8BRY7|A0A5Q8BRY7_LACPA/1-230
tr|C2FGX1|C2FGX1_LACPA/1-232
tr|A0A7M1BYV2|A0A7M1BYV2_LACPA/1-230
tr|A0A7Y7UIP9|A0A7Y7UIP9_LACRH/1-223
tr|A0A873ZKJ8|A0A873ZKJ8_LACRH/1-223
tr|A0A508YSD9|A0A508YSD9_LACRH/1-223
tr|C2JVW5|C2JVW5_LACRM/1-225
tr|A0A7S7JIG3|A0A7S7JIG3_LACRG/1-223
tr|A0A5R8LG36|A0A5R8LG36_LACCA/1-220
tr|A0A5R8M163|A0A5R8M163_LACZE/1-220
tr|A0A5R8LPW9|A0A5R8LPW9_LACZE/1-220
tr|A0A6L3ULK6|A0A6L3ULK6_LACCA/1-220
tr|A0A5P8JMB6|A0A5P8JMB6_9LACO/1-222
tr|A0A6M9CGY8|A0A6M9CGY8_LACPA/1-226
tr|A0A3M1Q672|A0A3M1Q672_9PROT/1-230
tr|A0A2X2BTH1|A0A2X2BTH1_PROMI/1-62
tr|A0A5C7FAT6|A0A5C7FAT6_9BACI/1-212
tr|A0A323TRL1|A0A323TRL1_9BACI/1-220
tr|A0A4Q0VWR4|A0A4Q0VWR4_9BACI/1-226
tr|A0A167T095|A0A167T095_9BACI/1-233
tr|A0A178T7K6|A0A178T7K6_9BACI/1-221
tr|A0A2G5RQC4|A0A2G5RQC4_9BACI/1-221
tr|A0A4S3L2U2|A0A4S3L2U2_9BACI/1-221
tr|A0A859FEI2|A0A859FEI2_9BACI/1-227
tr|A0A5D4RKA5|A0A5D4RKA5_9BACI/1-223
tr|A0A8J6MVY3|A0A8J6MVY3_9BACL/1-221
tr|A0A940X157|A0A940X157_9BACI/1-225
tr|A0A0M0KFR4|A0A0M0KFR4_ALKHA/1-226
tr|A0A4Y7WGD2|A0A4Y7WGD2_ALKHA/1-226
tr|A0A6I5A4T4|A0A6I5A4T4_9BACI/1-222
tr|A0A6I1F8S6|A0A6I1F8S6_9BACI/1-220
tr|A0A7Y0K5K6|A0A7Y0K5K6_9BACI/1-221
tr|A0A941JKN1|A0A941JKN1_NIACI/1-221
tr|A0A3G4ZB91|A0A3G4ZB91_NIACI/1-221
tr|A0A2N0Z0J2|A0A2N0Z0J2_9BACI/1-222
tr|A0A3S2U7Z2|A0A3S2U7Z2_9BACI/1-222
tr|A0A553SGE3|A0A553SGE3_NIACI/1-222
tr|A0A3D8X231|A0A3D8X231_PRIMG/1-235
tr|D5DWM9|D5DWM9_PRIM1/1-235
tr|A0A2B0QSA1|A0A2B0QSA1_PRIMG/1-235
tr|D5DMB8|D5DMB8_PRIM3/1-235
tr|A0A0B6AEC4|A0A0B6AEC4_PRIM2/1-235
tr|A0A6M6DJL5|A0A6M6DJL5_PRIMG/1-235
tr|A0A5B8PU61|A0A5B8PU61_PRIMG/1-235
tr|A0A2C5HGB8|A0A2C5HGB8_9BACI/1-235
tr|A0A6I0BXC7|A0A6I0BXC7_9BACI/1-224
tr|A0A1X7G830|A0A1X7G830_9BACI/1-224
tr|A0A8I1MJW2|A0A8I1MJW2_9BACI/1-221
tr|A0A264YXW5|A0A264YXW5_9BACI/1-221
tr|A0A4Z0D7A2|A0A4Z0D7A2_9FIRM/1-217
tr|A0A972J059|A0A972J059_9FIRM/1-223
tr|A0A6I7NH66|A0A6I7NH66_9MOLU/1-217
tr|A0A2N1QCB9|A0A2N1QCB9_9BACT/1-219
tr|A0A5B9Y3Q9|A0A5B9Y3Q9_9MOLU/1-225
tr|A0A1Y0L2F8|A0A1Y0L2F8_9MOLU/1-222
tr|W6A7G5|W6A7G5_9MOLU/1-218
tr|A0A0L1IJV0|A0A0L1IJV0_9MOLU/1-209
tr|A0A654IRF5|A0A654IRF5_9MOLU/1-209
tr|L5LA05|L5LA05_9MOLU/1-209
tr|A0A0C2ZL35|A0A0C2ZL35_MYCCA/1-209
tr|A0A084ERI9|A0A084ERI9_MYCCA/1-209
tr|E4PTJ3|E4PTJ3_MYCLG/1-209
tr|A0A2T4I9G8|A0A2T4I9G8_9MOLU/1-209
tr|A0A014LZV9|A0A014LZV9_MYCMC/1-209
tr|F4MP69|F4MP69_MYCML/1-209
tr|A0A014L6I7|A0A014L6I7_9MOLU/1-220
tr|A0A2D1UDW3|A0A2D1UDW3_9MOLU/1-223
tr|A0A269TII1|A0A269TII1_9MOLU/1-233
tr|A0A4R0XLU1|A0A4R0XLU1_9MOLU/1-219
tr|A0A4R0XQP0|A0A4R0XQP0_9MOLU/1-219
tr|A0A5Q3TK82|A0A5Q3TK82_MYCSY/1-225
tr|Q4A6U4|Q4A6U4_MYCS5/1-225
tr|Q6KHP8|Q6KHP8_MESM1/1-223
tr|A0A5B8K1Q0|A0A5B8K1Q0_9MOLU/1-226
tr|A0A5B8JY41|A0A5B8JY41_9MOLU/1-226
tr|A0A5B8K9J5|A0A5B8K9J5_9MOLU/1-226

```
tr|D5E6C2|D5E6C2_MYCCM/1-222
tr|D4XWP6|D4XWP6_9MOLU/1-222
tr|A0A6H0V695|A0A6H0V695_9MOLU/1-226
tr|A0A5P3I6Q7|A0A5P3I6Q7_MYCGL/1-225
tr|A0A6H2DUJ6|A0A6H2DUJ6_9MOLU/1-221
tr|E0TKS5|E0TKS5_MESHH/1-222
tr|A0A2D1JL35|A0A2D1JL35_9MOLU/1-226
tr|A0A809S131|A0A809S131_9MOLU/1-223
tr|A0A6N1DV19|A0A6N1DV19_9MOLU/1-222
tr|A0A857DBJ0|A0A857DBJ0_9MOLU/1-225
tr|A0A1D3IQZ5|A0A1D3IQZ5_NEIGO/1-216
tr|A0A378VUU5|A0A378VUU5_NEIGO/1-103
tr|A0A7L7KSP2|A0A7L7KSP2_9BACT/1-223
tr|A0A4Q2A2K1|A0A4Q2A2K1_9FIRM/1-225
tr|A0A942QY31|A0A942QY31_9CLOT/1-223
tr|A0A4S3LN34|A0A4S3LN34_9ENTR/1-227
tr|A0A947QGC9|A0A947QGC9_UNCSP/1-218
tr|A0A2S1JRC0|A0A2S1JRC0_9GAMM/1-238
tr|A0A3N9SG94|A0A3N9SG94_9SPHN/1-213
tr|A0A970K4C8|A0A970K4C8_9FIRM/1-223
tr|A0A8B6I567|A0A8B6I567_CLODI/1-205
tr|A0A1C3SVI1|A0A1C3SVI1_9LACT/1-215
tr|A0A401IRB6|A0A401IRB6_9LACO/1-228
tr|A0A5R9BUH3|A0A5R9BUH3_9LACO/1-231
tr|A0A1D7TRJ7|A0A1D7TRJ7_9LACO/1-223
tr|C2EJL1|C2EJL1_9LACO/1-226
tr|A0A6N9ITK7|A0A6N9ITK7_9LACO/1-223
tr|A0A2U2M0T0|A0A2U2M0T0_9LACO/1-223
tr|A0A089QCY6|A0A089QCY6_9LACO/1-223
tr|A0A1V9QRF9|A0A1V9QRF9_9LACO/1-223
tr|A0A6A8GZM7|A0A6A8GZM7_9LACO/1-223
tr|A0A930RHZ9|A0A930RHZ9_STRSP/1-222
tr|A0A417Z2L2|A0A417Z2L2_9LACO/1-223
tr|A0A347SSQ5|A0A347SSQ5_9LACO/1-223
tr|A0A417ZHN1|A0A417ZHN1_9LACO/1-223
tr|A0A3M0NZH8|A0A3M0NZH8_9LACO/1-221
tr|A0A3M0PJK5|A0A3M0PJK5_9LACO/1-222
tr|A0A318M8M9|A0A318M8M9_9LACO/1-222
tr|A0A3M0NK52|A0A3M0NK52_9LACO/1-219
tr|A0A3M0M161|A0A3M0M161_9LACO/1-219
tr|A0A396SR60|A0A396SR60_9LACO/1-223
tr|A0A417ZM58|A0A417ZM58_9LACO/1-223
tr|A0A3M0MXD7|A0A3M0MXD7_9LACO/1-223
tr|A0A3M0NTF9|A0A3M0NTF9_9LACO/1-223
tr|A0A5B8THR4|A0A5B8THR4_9LACO/1-218
tr|A0A2N8PUH8|A0A2N8PUH8_ENTAV/1-218
tr|A0A4P8KD45|A0A4P8KD45_ENTAV/1-218
tr|A0A6I7ZCH0|A0A6I7ZCH0_9ENTE/1-218
tr|A0A6F9XSC6|A0A6F9XSC6_9LACO/1-223
tr|A0A6F9YNW9|A0A6F9YNW9_9LACO/1-223
tr|A0A6N3D805|A0A6N3D805_STRPA/1-220
tr|F8DHB4|F8DHB4_STREP/1-221
tr|I2NML7|I2NML7_STRPA/1-220
tr|A0A7T3T2D2|A0A7T3T2D2_9LACT/1-220
tr|I2NDM3|I2NDM3_STRPA/1-216
tr|A0A350B8D6|A0A350B8D6_9FIRM/1-219
tr|A0A3D1YS23|A0A3D1YS23_9FIRM/1-219
tr|A0A6M0L9Y1|A0A6M0L9Y1_9LACT/1-220
tr|A0A7X6S063|A0A7X6S063_9STRE/1-219
tr|A0A5R9CPX6|A0A5R9CPX6_9LACO/1-224
tr|A0A3R8JIB9|A0A3R8JIB9_9LACO/1-224
tr|A0A6P1E846|A0A6P1E846_LENHI/1-224
tr|C0XGZ2|C0XGZ2_LENH9/1-224
tr|A0A6N4HLV8|A0A6N4HLV8_9LACO/1-224
tr|A0A224V819|A0A224V819_9LACO/1-224
tr|A0A8E1X3U8|A0A8E1X3U8_LENBU/1-224
tr|A0A844EFI8|A0A844EFI8_9LACO/1-224
tr|A0A1Z5H503|A0A1Z5H503_9LACO/1-231
tr|A0A7X2XWP9|A0A7X2XWP9_9LACO/1-230
tr|A0A929MS34|A0A929MS34_ABIDE/1-221
tr|A0A929MWW5|A0A929MWW5_9LACT/1-221
tr|A0A948TKD2|A0A948TKD2_9LACO/1-223
tr|A0A291KFE6|A0A291KFE6_BROTH/1-214
tr|A0A7G2N1R9|A0A7G2N1R9_LACSP/1-215
tr|A0A6I5XGQ0|A0A6I5XGQ0_LEUCI/1-222
tr|A0A7L8UYI0|A0A7L8UYI0_9LACO/1-222
tr|A0A5C5EDU1|A0A5C5EDU1_9LACT/1-227
tr|A0A069AX49|A0A069AX49_CLODI/1-223
tr|A0A125V9P9|A0A125V9P9_CLODI/1-223
tr|A0A0H3NAV5|A0A0H3NAV5_CLODC/1-223
tr|A0A0B2XMZ6|A0A0B2XMZ6_LATCU/1-222
tr|A0A8E3BL74|A0A8E3BL74_LATCU/1-222
tr|A0A921F7R1|A0A921F7R1_9LACO/1-223
tr|A0A9D1RJU2|A0A9D1RJU2_9LACO/1-222
tr|G8PBC3|G8PBC3_PEDCP/1-224
tr|A0A9D1UWL1|A0A9D1UWL1_9LACO/1-221
tr|A0A3E4MGP9|A0A3E4MGP9_9LACO/1-222
tr|A0A6A8HTF8|A0A6A8HTF8_9LACO/1-222
tr|E7FQ32|E7FQ32_9LACO/1-222
tr|A0A3B8RJU9|A0A3B8RJU9_9LACO/1-220
tr|A0A4V3RNS1|A0A4V3RNS1_9LACO/1-221
tr|A0A2Z4VW08|A0A2Z4VW08_9LACO/1-221
tr|A0A358LWQ5|A0A358LWQ5_9LACO/1-221
tr|A0A4Q2AZU8|A0A4Q2AZU8_9LACO/1-221
tr|A0A6M0I6M2|A0A6M0I6M2_9LACO/1-221
tr|A0A7X3LJG6|A0A7X3LJG6_9BACL/1-220
tr|A0A7T1H1L7|A0A7T1H1L7_9GAMM/1-215
tr|B2PZ81|B2PZ81_PROST/1-215
tr|A0A433ZZG4|A0A433ZZG4_MORMO/1-219
tr|A0A7D6HJP1|A0A7D6HJP1_9GAMM/1-216
tr|A0A6P1RZW6|A0A6P1RZW6_9GAMM/1-223
tr|A0A8B3F843|A0A8B3F843_PECPM/1-220
tr|E6LHM4|E6LHM4_ENTI1/1-223
tr|A0A374JDV7|A0A374JDV7_9ENTE/1-219
tr|A0A414XF14|A0A414XF14_9ENTE/1-227
tr|A0A4S3B8C7|A0A4S3B8C7_9ENTE/1-222
tr|A0A6G8AQS1|A0A6G8AQS1_9ENTE/1-220
tr|A0A8I2AYW0|A0A8I2AYW0_9ENTE/1-220
tr|A0A7X6D749|A0A7X6D749_9ENTE/1-220
tr|A0A8I1VTS3|A0A8I1VTS3_9ENTE/1-220
tr|A0A3S0A543|A0A3S0A543_9ENTE/1-225
tr|A0A3Q9BKW0|A0A3Q9BKW0_9LACT/1-222
tr|A0A6G7KAJ2|A0A6G7KAJ2_9LACT/1-222
tr|A0A9D2I3S0|A0A9D2I3S0_9LACT/1-222
tr|A0A7T2Y991|A0A7T2Y991_9LACT/1-223
tr|A0A352NDJ4|A0A352NDJ4_LACSP/1-221
tr|A0A7T2Y8M0|A0A7T2Y8M0_9LACT/1-221
tr|A0A7T4NKP8|A0A7T4NKP8_9LACT/1-221
tr|A0A098CQ48|A0A098CQ48_9LACT/1-221
tr|A0A6L2ZTU2|A0A6L2ZTU2_9LACT/1-221
tr|A0A4P5PFA7|A0A4P5PFA7_9ENTE/1-222
tr|A0A4D7CNV0|A0A4D7CNV0_9ENTE/1-223
tr|A0A4Z0D3K6|A0A4Z0D3K6_9ENTE/1-223
```

```
tr|A0A4Y9JGE2|A0A4Y9JGE2_9STRE/1-219
tr|A0A0E3WFC2|A0A0E3WFC2_9STRE/1-219
tr|A0A943W3N3|A0A943W3N3_9LACT/1-222
tr|A0A3C1XAL0|A0A3C1XAL0_LACSP/1-219
tr|A0A7T4TI13|A0A7T4TI13_9LACT/1-225
tr|A0A098CXE7|A0A098CXE7_9LACT/1-225
tr|A0A9E4CAY4|A0A9E4CAY4_LACSP/1-225
tr|A0A2R5HKL2|A0A2R5HKL2_9LACT/1-223
tr|A0A5C8K8E1|A0A5C8K8E1_9LACT/1-223
tr|A0A7X1ZBN7|A0A7X1ZBN7_9LACT/1-221
tr|A0A4Q7DT33|A0A4Q7DT33_9LACT/1-221
tr|A0A514Z7U5|A0A514Z7U5_9LACT/1-218
tr|A0A387BLK2|A0A387BLK2_9LACT/1-219
tr|A0A387BTP3|A0A387BTP3_9LACT/1-221
tr|A0A552XK19|A0A552XK19_9LACT/1-225
tr|A2RIG5|A2RIG5_LACLM/1-221
tr|T2F4K2|T2F4K2_LACLC/1-221
tr|A0A3N6LI22|A0A3N6LI22_9LACT/1-221
tr|A0A5D4G7I0|A0A5D4G7I0_LACLL/1-221
sp|P71447|PGMB_LACLA/1-221
tr|A0A1V0NDU3|A0A1V0NDU3_LACLL/1-221
tr|A0A5P3IFK1|A0A5P3IFK1_LACLL/1-221
tr|A0A6I2HTB2|A0A6I2HTB2_LACLC/1-221
tr|A0A2X0R470|A0A2X0R470_9LACT/1-221
tr|A0A6M0MA83|A0A6M0MA83_9LACT/1-208
tr|A0A0V8AP46|A0A0V8AP46_LACLL/1-221
tr|S6EW90|S6EW90_LACLL/1-221
tr|A0A552Z207|A0A552Z207_9LACT/1-221
tr|A0A0H1RRQ5|A0A0H1RRQ5_LACLL/1-221
tr|A0A7X1VGM1|A0A7X1VGM1_9LACT/1-221
tr|A0A2A5SEY2|A0A2A5SEY2_LACLH/1-221
tr|A0A2A9IDT2|A0A2A9IDT2_9LACT/1-221
tr|D2BNW1|D2BNW1_LACLK/1-221
tr|H5SX32|H5SX32_LACLL/1-221
tr|A0A4Y9GJX1|A0A4Y9GJX1_9NEIS/1-224
tr|A0A7H1MF12|A0A7H1MF12_9NEIS/1-224
tr|F5S8F3|F5S8F3_KINKI/1-222
tr|A0A220S0L7|A0A220S0L7_9NEIS/1-221
tr|A0A381E2E8|A0A381E2E8_9GAMM/1-220
tr|C8NDK2|C8NDK2_CARH6/1-220
tr|A0A125WF99|A0A125WF99_NEIME/1-221
```

```
tr|A0A828RMQ4|A0A828RMQ4_NEIME/1-221
tr|A0A112R2H5|A0A112R2H5_NEIME/1-221
tr|A0A0Y5LYF6|A0A0Y5LYF6_NEIME/1-221
tr|D0WAS7|D0WAS7_NEILA/1-221
tr|C0EN14|C0EN14_NEIFL/1-221
tr|A0A378VTJ4|A0A378VTJ4_NEIME/1-221
tr|I4E3Q3|I4E3Q3_NEIME/1-221
tr|A0A0G4BWH2|A0A0G4BWH2_NEIME/1-221
tr|E3D1A4|E3D1A4_NEIM7/1-221
tr|A9M2K1|A9M2K1_NEIM0/1-221
tr|A0A425AMM7|A0A425AMM7_NEIME/1-221
tr|E0N6N2|E0N6N2_NEIME/1-221
tr|A0A9K2KMN5|A0A9K2KMN5_NEIM8/1-221
tr|A0A9D5WRU9|A0A9D5WRU9_NEISI/1-221
tr|C6M175|C6M175_NEISI/1-221
tr|I2NRK1|I2NRK1_NEISI/1-221
tr|D2ZUE0|D2ZUE0_NEIMU/1-221
tr|A0A4D7WHH0|A0A4D7WHH0_NEISU/1-221
tr|A0A2I1XCY5|A0A2I1XCY5_NEISI/1-221
tr|A0A930GV05|A0A930GV05_NEISI/1-221
tr|A0A5Q3S229|A0A5Q3S229_9NEIS/1-221
tr|A0A8I1D9D5|A0A8I1D9D5_NEIME/1-221
tr|E6MX25|E6MX25_NEIMH/1-221
tr|Q9K108|Q9K108_NEIMB/1-221
tr|A0A854WI16|A0A854WI16_9STRE/1-222
tr|A0A1C3SX67|A0A1C3SX67_9LACT/1-223
tr|A0A0D6DX88|A0A0D6DX88_9LACT/1-223
tr|A0A6A0B358|A0A6A0B358_9LACT/1-220
tr|A0A6A0B818|A0A6A0B818_9LACT/1-227
tr|A0A847J262|A0A847J262_9LACT/1-220
tr|A0A965UCB6|A0A965UCB6_9FIRM/1-220
tr|A0A2A5SH88|A0A2A5SH88_9LACT/1-220
tr|A0A8E0REM3|A0A8E0REM3_LACSP/1-220
tr|A0A7V8N0B5|A0A7V8N0B5_9LACT/1-220
tr|A0A3A1Y7W0|A0A3A1Y7W0_9GAMM/1-224
tr|A0A3A1YK75|A0A3A1YK75_9GAMM/1-222
tr|A0A3A1Y398|A0A3A1Y398_9GAMM/1-221
tr|A0A3A1YHQ7|A0A3A1YHQ7_9GAMM/1-222
tr|A0A828QQ74|A0A828QQ74_ENTFL/1-222
tr|A0A125W9Z1|A0A125W9Z1_ENTFL/1-222
tr|A0A1B4XMC2|A0A1B4XMC2_ENTFL/1-222
```

```
tr|A0A4V5V1S1|A0A4V5V1S1_ENTFL/1-142
tr|A0A640MY51|A0A640MY51_BACAN/1-222
tr|E2YY31|E2YY31_ENTFL/1-222
tr|Q836Y8|Q836Y8_ENTFA/1-222
tr|A0A970JPA5|A0A970JPA5_ENTSX/1-221
tr|A0A179EU54|A0A179EU54_ENTTH/1-221
tr|A0A249SGF9|A0A249SGF9_ENTTH/1-221
tr|A0A2S7RN44|A0A2S7RN44_ENTMU/1-221
tr|A0A1V2UE92|A0A1V2UE92_ENTMU/1-221
tr|V5XKN7|V5XKN7_ENTMU/1-221
tr|A0A2T5DF36|A0A2T5DF36_ENTMU/1-221
tr|A0A132P760|A0A132P760_ENTFC/1-221
tr|Q3Y2T0|Q3Y2T0_ENTFD/1-221
tr|A0A7V8C7I6|A0A7V8C7I6_ENTFC/1-221
tr|A0A828QCF7|A0A828QCF7_ENTFC/1-221
tr|A0A2G0EE36|A0A2G0EE36_ENTFC/1-221
tr|A0A3N3L0R1|A0A3N3L0R1_ENTFC/1-221
tr|A0A7Y6QES3|A0A7Y6QES3_ENTFC/1-221
tr|A0A6N2ZU38|A0A6N2ZU38_ENTFC/1-221
tr|A0A8E2CCG1|A0A8E2CCG1_9ENTE/1-221
tr|A0A8H9BFW2|A0A8H9BFW2_ENTFC/1-221
tr|A0A133MW26|A0A133MW26_ENTFC/1-221
tr|A0A8X8GSW4|A0A8X8GSW4_9ENTE/1-221
tr|A0A242BCB4|A0A242BCB4_ENTFC/1-221
tr|A0A4Y3JT47|A0A4Y3JT47_9ENTE/1-221
tr|A0A7W2AL47|A0A7W2AL47_9ENTE/1-221
tr|A0A4U4EIH1|A0A4U4EIH1_9ENTE/1-221
tr|A0A5N0YY33|A0A5N0YY33_9ENTE/1-221
tr|A0A377L238|A0A377L238_9ENTE/1-221
tr|A0A3F3NK74|A0A3F3NK74_ENTFC/1-221
tr|A0A7U6DSS4|A0A7U6DSS4_9ENTE/1-221
tr|A0A921KDM8|A0A921KDM8_SPOPS/1-221
tr|A0A2D0BY13|A0A2D0BY13_ENTFC/1-221
tr|A0A713XYK4|A0A713XYK4_SALTM/1-221
tr|A0A8X8GV29|A0A8X8GV29_9ENTE/1-221
tr|A0A2V2L4B4|A0A2V2L4B4_ENTFC/1-221
tr|A0A2S7MD00|A0A2S7MD00_ENTFC/1-221
tr|A0A377KL22|A0A377KL22_9ENTE/1-221
tr|A0A511J2Q3|A0A511J2Q3_9ENTE/1-221
tr|A0A3D5CSW9|A0A3D5CSW9_ENTSX/1-221
tr|A0A1V8VG55|A0A1V8VG55_ENTHR/1-221
```

tr|A0A838Z7I4|A0A838Z7I4_ENTHR/1-221
tr|A0A9D9XJ16|A0A9D9XJ16_ENTSX/1-220
tr|A0A9E3ZVG2|A0A9E3ZVG2_9ENTE/1-220
tr|A0A345FPL8|A0A345FPL8_9ENTE/1-219
tr|A0A553S8V3|A0A553S8V3_ENTAV/1-219
tr|A0A2N8PX59|A0A2N8PX59_ENTAV/1-219
tr|A0A6I7Z7G7|A0A6I7Z7G7_9ENTE/1-219
tr|A0A437UQQ8|A0A437UQQ8_ENTAV/1-219
tr|A0A4P8KFL4|A0A4P8KFL4_ENTAV/1-219
tr|A0A3D9A1U4|A0A3D9A1U4_9ENTE/1-222
tr|A0A7U6QX42|A0A7U6QX42_ENTAV/1-219
tr|A0A842GKZ3|A0A842GKZ3_9LIST/1-220
tr|A0A842BAJ4|A0A842BAJ4_9LIST/1-220
tr|A0A842GEK2|A0A842GEK2_9LIST/1-220
tr|A0A841YJ57|A0A841YJ57_9LIST/1-220
tr|A0A842AE59|A0A842AE59_9LIST/1-220
tr|A0A841XUJ1|A0A841XUJ1_9LIST/1-220
tr|A0A7X0YX05|A0A7X0YX05_9LIST/1-220
tr|A0A7X1D823|A0A7X1D823_9LIST/1-220
tr|A0A842EMB4|A0A842EMB4_9LIST/1-220
tr|A0A7X0XPF3|A0A7X0XPF3_9LIST/1-220
tr|A0A842FIX4|A0A842FIX4_9LIST/1-220
tr|A0A842EN69|A0A842EN69_9LIST/1-220
tr|A0A842BGJ2|A0A842BGJ2_9LIST/1-220
tr|A0A7X0XXR5|A0A7X0XXR5_9LIST/1-220
tr|A0A841X481|A0A841X481_9LIST/1-220
tr|A0A842C5K4|A0A842C5K4_9LIST/1-220
tr|A0A842EDR8|A0A842EDR8_9LIST/1-220
tr|A0A841XX51|A0A841XX51_9LIST/1-220
tr|A0A841XH92|A0A841XH92_9LIST/1-220
tr|A0A7X1CE47|A0A7X1CE47_9LIST/1-220
tr|A0A7X0WCW8|A0A7X0WCW8_9LIST/1-220
tr|A0A842E831|A0A842E831_9LIST/1-220
tr|A0A7X0XFR8|A0A7X0XFR8_9LIST/1-220
tr|A0A842FVF9|A0A842FVF9_9LIST/1-220
tr|A0A842AYD3|A0A842AYD3_9LIST/1-220
tr|A0A842A5T2|A0A842A5T2_9LIST/1-220
tr|A0A841W8G9|A0A841W8G9_9LIST/1-220
tr|A0A7X1A7R2|A0A7X1A7R2_9LIST/1-220
tr|A0A842D6D8|A0A842D6D8_9LIST/1-220
tr|A0A7X1CZK3|A0A7X1CZK3_9LIST/1-220
tr|A0A841ZWD3|A0A841ZWD3_9LIST/1-220
tr|A0A7X0ZMZ8|A0A7X0ZMZ8_9LIST/1-220
tr|A0A841VWE5|A0A841VWE5_9LIST/1-220
tr|A0A842GH32|A0A842GH32_9LIST/1-219
tr|A0A842E823|A0A842E823_9LIST/1-219
tr|A0A842DGD2|A0A842DGD2_9LIST/1-219
tr|A0A841ZZ13|A0A841ZZ13_9LIST/1-219
tr|A0A842DYY8|A0A842DYY8_9LIST/1-220
tr|A0A841VW89|A0A841VW89_9LIST/1-220
tr|A0A841XCZ6|A0A841XCZ6_9LIST/1-220
tr|A0A7X1DQG5|A0A7X1DQG5_9LIST/1-220
tr|A0A7X0XQI9|A0A7X0XQI9_9LIST/1-220
tr|A0A842BFT4|A0A842BFT4_9LIST/1-220
tr|A0A842GJ35|A0A842GJ35_9LIST/1-219
tr|A0A842FSZ0|A0A842FSZ0_9LIST/1-219
tr|A0A842B1S5|A0A842B1S5_9LIST/1-219
tr|A0A841YLI1|A0A841YLI1_9LIST/1-219
tr|A0A841Y1J6|A0A841Y1J6_9LIST/1-219
tr|A0A841XMM5|A0A841XMM5_9LIST/1-219
tr|A0A7X1C0U5|A0A7X1C0U5_9LIST/1-219
tr|A0A7X0YNY6|A0A7X0YNY6_9LIST/1-219
tr|A0A7X1D5Z5|A0A7X1D5Z5_9LIST/1-219
tr|A0A7X0Y2N5|A0A7X0Y2N5_9LIST/1-219
tr|A0A841Z2W8|A0A841Z2W8_9LIST/1-220
tr|A0A7W1T534|A0A7W1T534_9LIST/1-220
tr|A0A3N9UDR8|A0A3N9UDR8_9LIST/1-219
tr|A0A841YZ04|A0A841YZ04_9LIST/1-219
tr|A0A5C5UPD7|A0A5C5UPD7_9BACL/1-224
tr|A0A2N5G9J5|A0A2N5G9J5_9BACI/1-223
tr|A0A3E2JCE2|A0A3E2JCE2_9BACI/1-225
tr|A0A2A8FPV0|A0A2A8FPV0_9BACI/1-225
tr|A0A2M9NPU6|A0A2M9NPU6_9BACI/1-234
tr|A0A417YY53|A0A417YY53_9BACI/1-224
tr|A0A4Q2I2X6|A0A4Q2I2X6_9BACI/1-228
tr|A0A268JB33|A0A268JB33_9BACI/1-222
tr|A0A6H1NYS3|A0A6H1NYS3_PRIMG/1-227
tr|A0A2C2UUA4|A0A2C2UUA4_9BACI/1-227
tr|A0A841YCT4|A0A841YCT4_9LIST/1-221
tr|A0A7X0X4U3|A0A7X0X4U3_9LIST/1-220
tr|A0A7X0X391|A0A7X0X391_LISSE/1-220
tr|A0A7X0T3V9|A0A7X0T3V9_LISWE/1-220
tr|A0A7X1A1Z1|A0A7X1A1Z1_9LIST/1-220
tr|A0A1E7E7M6|A0A1E7E7M6_LISMN/1-220
tr|A0A842BP50|A0A842BP50_LISIO/1-220
tr|A0A823HUW5|A0A823HUW5_LISMN/1-220
tr|A0A7X0ZDL5|A0A7X0ZDL5_9LIST/1-220
tr|A0A842CQD7|A0A842CQD7_9LIST/1-220
tr|A0A7X0WKE6|A0A7X0WKE6_LISIO/1-220
tr|A0A3T2AHN0|A0A3T2AHN0_LISMN/1-220
tr|A0A459ZHE1|A0A459ZHE1_LISMN/1-220
tr|A0A7X1DEM7|A0A7X1DEM7_9LIST/1-220
tr|A0A823IQR2|A0A823IQR2_LISMN/1-220
tr|A0A608I6J4|A0A608I6J4_LISMN/1-220
tr|A0A4B9HT01|A0A4B9HT01_LISMN/1-220
tr|A0A5K9RME4|A0A5K9RME4_LISMN/1-220
tr|A0A5Y9DN63|A0A5Y9DN63_LISMN/1-220
tr|A0A2Z5C3H3|A0A2Z5C3H3_LISMN/1-220
tr|A0A660JIE3|A0A660JIE3_LISIO/1-220
tr|A0A464T1T7|A0A464T1T7_LISMN/1-220
tr|A0A940P7D9|A0A940P7D9_9ENTE/1-218
tr|A0A6I1FA20|A0A6I1FA20_9BACI/1-219
tr|A0A7Y0PP82|A0A7Y0PP82_9BACI/1-219
tr|A0A941GKE7|A0A941GKE7_NIACI/1-219
tr|A0A268FFD3|A0A268FFD3_NIACI/1-219
tr|A0A3Q9BLG5|A0A3Q9BLG5_9LACT/1-220
tr|A0A6G7KB78|A0A6G7KB78_9LACT/1-220
tr|A0A6G7WK98|A0A6G7WK98_9LACT/1-223
tr|A0A847BBK9|A0A847BBK9_9LACT/1-223
tr|A0A940P492|A0A940P492_9ENTE/1-220
tr|A0A956NVC9|A0A956NVC9_9LACT/1-219
tr|A0A0R2J741|A0A0R2J741_CARML/1-219
tr|A0A2N5JHU6|A0A2N5JHU6_CARML/1-219
tr|A0A8B5W6F9|A0A8B5W6F9_9LACT/1-219
tr|A0A4V1ENS5|A0A4V1ENS5_ENTAV/1-215
tr|A0A6I7ZGT6|A0A6I7ZGT6_9ENTE/1-215
tr|A0A8B5VUS2|A0A8B5VUS2_ENTAV/1-215
tr|A0A7H0FNP0|A0A7H0FNP0_ENTFL/1-216
tr|A0A4U3MF36|A0A4U3MF36_ENTFL/1-216
tr|A0A855UHG4|A0A855UHG4_ENTFL/1-216
tr|A0A828QM83|A0A828QM83_ENTFL/1-216
tr|E2Z0N3|E2Z0N3_ENTFL/1-216
tr|A0A8B3RWL5|A0A8B3RWL5_ENTFL/1-216

tr|A0A1G1SAA9|A0A1G1SAA9_ENTFL/1-216
tr|A0A6I4XUI0|A0A6I4XUI0_ENTFL/1-216
tr|A0A9D9RQ01|A0A9D9RQ01_9FUSO/1-214
tr|A0A3G5FBH3|A0A3G5FBH3_9ENTE/1-216
tr|A0A3G5F7I1|A0A3G5F7I1_9ENTE/1-216
tr|A0A2H6CCD8|A0A2H6CCD8_TETHA/1-216
tr|A0A3G5FG75|A0A3G5FG75_TETHA/1-216
tr|A0A829UCW9|A0A829UCW9_TETHA/1-216
tr|A0A6I5YDB2|A0A6I5YDB2_TETHA/1-216
tr|A0A2H6DD88|A0A2H6DD88_TETHA/1-216
tr|A0A851HWJ3|A0A851HWJ3_TETHA/1-216
tr|A0A2H6CYH3|A0A2H6CYH3_TETHA/1-216
tr|A0A2W3Z0J1|A0A2W3Z0J1_9ENTE/1-216
tr|D7UWF4|D7UWF4_LISGR/1-216
tr|A0A940SYA9|A0A940SYA9_9ENTE/1-216
tr|A0A1Y3UQ19|A0A1Y3UQ19_9ENTE/1-216
tr|A0A9D2QNY7|A0A9D2QNY7_9ENTE/1-216
tr|A0A7X9NLV5|A0A7X9NLV5_9ENTE/1-216
tr|A0A413AKU4|A0A413AKU4_9ENTE/1-216
tr|A0A0H2Q4T5|A0A0H2Q4T5_9ENTE/1-216
tr|A0A9D2TUN0|A0A9D2TUN0_9ENTE/1-216
tr|A0A9D2PGW7|A0A9D2PGW7_9FIRM/1-216
tr|A0A9D2F7P3|A0A9D2F7P3_9ENTE/1-216
tr|A0A970FAW4|A0A970FAW4_ENTSX/1-216
tr|A0A3D1J8R6|A0A3D1J8R6_ENTSX/1-216
tr|A0A1V8Z0I3|A0A1V8Z0I3_ENTGA/1-216
tr|A0A7U6DUP4|A0A7U6DUP4_9ENTE/1-216
tr|A0A376H0I4|A0A376H0I4_ENTGA/1-216
tr|A0A415EPS8|A0A415EPS8_ENTCA/1-216
tr|A0A7T4GD50|A0A7T4GD50_ENTCA/1-216
tr|A0A8G1WPT8|A0A8G1WPT8_ENTCA/1-216
tr|A0A494SJC1|A0A494SJC1_ENTCA/1-216
tr|A0A505INM1|A0A505INM1_9ENTE/1-216
tr|F0EFQ2|F0EFQ2_ENTCA/1-216
tr|A0A109Y4W3|A0A109Y4W3_ENTGA/1-216
tr|A0A291DM58|A0A291DM58_9ENTE/1-216
tr|A0A939CBC2|A0A939CBC2_ENTSX/1-150
tr|A0A2A7WPM4|A0A2A7WPM4_9BACI/1-218
tr|A0A2C0ZT10|A0A2C0ZT10_9BACI/1-220
tr|A0A4R0ZFU9|A0A4R0ZFU9_9BACL/1-216
tr|A0A8B4X3V4|A0A8B4X3V4_9BACL/1-216
tr|A0A248L9G7|A0A248L9G7_9BACL/1-220
tr|A0A369KH43|A0A369KH43_9BACL/1-220
tr|A0A7H0V8I8|A0A7H0V8I8_9BACL/1-220
tr|A0A419DMF6|A0A419DMF6_9BACL/1-220
tr|A0A7L5ZWY8|A0A7L5ZWY8_9RHOB/1-221
tr|A0A4V2NPW8|A0A4V2NPW8_9BACL/1-221
tr|A0A4V2NNL4|A0A4V2NNL4_9BACL/1-221
tr|A0A8B4X8H9|A0A8B4X8H9_9BACL/1-221
tr|A0A4R0Z8M4|A0A4R0Z8M4_9BACL/1-221
tr|A0A4R0ZUN1|A0A4R0ZUN1_9BACL/1-221
tr|A0A559K782|A0A559K782_9BACL/1-222
tr|A0A2A7WNR3|A0A2A7WNR3_9BACI/1-223
tr|A0A2C0ZWG1|A0A2C0ZWG1_9BACI/1-223
tr|A0A3R9N8S7|A0A3R9N8S7_9BACI/1-223
tr|A0A940IK21|A0A940IK21_9BACI/1-224
tr|A0A2G5RU54|A0A2G5RU54_9BACI/1-229
tr|A0A4S3L4M3|A0A4S3L4M3_9BACI/1-229
tr|A0A9E8RVT5|A0A9E8RVT5_9BACI/1-222
tr|A0A9E8RY60|A0A9E8RY60_9BACI/1-222
tr|C6J6W3|C6J6W3_9BACL/1-228
tr|A0A3P3TW86|A0A3P3TW86_9BACL/1-225
tr|A0A090ZEJ6|A0A090ZEJ6_PAEMA/1-221
tr|A0A6N8EZL5|A0A6N8EZL5_PAEMA/1-225
tr|A0A934J769|A0A934J769_9BACL/1-223
tr|A0A378XCL9|A0A378XCL9_PANTH/1-223
tr|A0A3A3GNA5|A0A3A3GNA5_PANTH/1-223
tr|A0A2W4GKD1|A0A2W4GKD1_9BACL/1-223
tr|A0A4V3B4R4|A0A4V3B4R4_9BACL/1-223
tr|A0A268EYU1|A0A268EYU1_9BACL/1-225
tr|A0A269W238|A0A269W238_9BACL/1-225
tr|A0A328W8W0|A0A328W8W0_PAELA/1-227
tr|A0A2A5LJV2|A0A2A5LJV2_PAELA/1-224
tr|A0A385TV84|A0A385TV84_PAELA/1-224
tr|F3MDU9|F3MDU9_9BACL/1-224
tr|A0A5C6VWC5|A0A5C6VWC5_9BACI/1-222
tr|A0A2C1K6Q4|A0A2C1K6Q4_9BACI/1-221
tr|A0A941FT46|A0A941FT46_9BACI/1-221
tr|A0A2S7N226|A0A2S7N226_9BACI/1-221
tr|A0A5J6XUF3|A0A5J6XUF3_9BACI/1-222
tr|A0A2W1L5A7|A0A2W1L5A7_9BACL/1-223
tr|A0A2N5G3S9|A0A2N5G3S9_9BACI/1-224
tr|A0A3R9MIK5|A0A3R9MIK5_9BACI/1-224
tr|A0A2N5GSU3|A0A2N5GSU3_9BACI/1-224
tr|A0A2T6G7D7|A0A2T6G7D7_9BACL/1-221
tr|A0A8J3AQN2|A0A8J3AQN2_9BACI/1-224
tr|A0A2N5GZT7|A0A2N5GZT7_9BACI/1-222
tr|A0A3C2CUI5|A0A3C2CUI5_BACSP/1-222
tr|A0A3R9E6G6|A0A3R9E6G6_9BACI/1-222
tr|A0A846TCF8|A0A846TCF8_9BACI/1-222
tr|A0A944GX86|A0A944GX86_9BACI/1-222
tr|A0A4R5HFE8|A0A4R5HFE8_9ALTE/1-218
tr|A0A2S2E5N7|A0A2S2E5N7_9ALTE/1-218
tr|A0A356BHN5|A0A356BHN5_9GAMM/1-215
tr|A0A2G4YVP5|A0A2G4YVP5_9PROT/1-215
tr|A0A972FR59|A0A972FR59_9GAMM/1-223
tr|A0A4R5VY47|A0A4R5VY47_9BURK/1-217
tr|A0A317IG90|A0A317IG90_9BACT/1-217
tr|A0A7X3G5A5|A0A7X3G5A5_9BURK/1-214
tr|A0A9E3JSB1|A0A9E3JSB1_9BURK/1-214
tr|A0A4Y9ST60|A0A4Y9ST60_9BURK/1-213
tr|A0A7Y2JW57|A0A7Y2JW57_9BURK/1-213
tr|A0A924AX92|A0A924AX92_9BURK/1-214
tr|A0A7G5ZJC5|A0A7G5ZJC5_9BURK/1-213
tr|A0A7L9U7Z4|A0A7L9U7Z4_9BURK/1-213
tr|A0A430HPB8|A0A430HPB8_9BURK/1-213
tr|A0A2D2DF33|A0A2D2DF33_9BURK/1-214
tr|A0A418XG46|A0A418XG46_9BURK/1-213
tr|A0A848HEI4|A0A848HEI4_9BURK/1-213
tr|A0A6L6QP00|A0A6L6QP00_9BURK/1-223
tr|A0A6L6Q9V0|A0A6L6Q9V0_9BURK/1-218
tr|A0A7X2IJB6|A0A7X2IJB6_9BURK/1-217
tr|W0V9S6|W0V9S6_9BURK/1-216
tr|A0A6A7N0Z0|A0A6A7N0Z0_9BURK/1-213
tr|A0A843SDE2|A0A843SDE2_9BURK/1-213
tr|A0A845GN86|A0A845GN86_9BURK/1-213
tr|A0A845G2E3|A0A845G2E3_9BURK/1-213
tr|A0A845HIW2|A0A845HIW2_9BURK/1-213
tr|A0A4Y9SJ79|A0A4Y9SJ79_9BURK/1-213
tr|A0A6L6PIS1|A0A6L6PIS1_9BURK/1-213
tr|A0A7X4GTC5|A0A7X4GTC5_9BURK/1-213
tr|A0A6L8KXG3|A0A6L8KXG3_9BURK/1-213
tr|A0A6L5QCK2|A0A6L5QCK2_9BURK/1-213

tr|A0A844D5Z9|A0A844D5Z9_9BURK/1-213
tr|A0A7X4GZG7|A0A7X4GZG7_9BURK/1-213
tr|A0A7X4KQV1|A0A7X4KQV1_9BURK/1-212
tr|A0A6I3SWV3|A0A6I3SWV3_9BURK/1-215
tr|A0A4P7BJE3|A0A4P7BJE3_9BURK/1-214
tr|A0A7U5Y1Y7|A0A7U5Y1Y7_9BURK/1-214
tr|A0A2R4CH24|A0A2R4CH24_9BURK/1-214
tr|A0A7Y6QK23|A0A7Y6QK23_9BURK/1-213
tr|A0A562PPF3|A0A562PPF3_9BURK/1-214
tr|A0A411X8D8|A0A411X8D8_9BURK/1-240
tr|A0A4V0Z4K1|A0A4V0Z4K1_9BURK/1-211
tr|A0A6I3XFE5|A0A6I3XFE5_9BURK/1-211
tr|A0A919Y3Z7|A0A919Y3Z7_9BACL/1-222
tr|A0A920CFR8|A0A920CFR8_9BACL/1-222
tr|A0A3P3U8Y7|A0A3P3U8Y7_9BACL/1-223
tr|A0A5D0CY64|A0A5D0CY64_9BACL/1-223
tr|A0A3M7TQ88|A0A3M7TQ88_9BACI/1-221
tr|A0A7G5C485|A0A7G5C485_9BACL/1-223
tr|A0A916W6Y2|A0A916W6Y2_9BACI/1-230
tr|A0A495A103|A0A495A103_9BACI/1-222
tr|A0A0D0EPC3|A0A0D0EPC3_9BACI/1-223
tr|A0A090J2K9|A0A090J2K9_9BACI/1-223
tr|A0A267TNJ7|A0A267TNJ7_9BACI/1-223
tr|A0A6N8FN11|A0A6N8FN11_9BACI/1-224
tr|A0A365KWN9|A0A365KWN9_9BACL/1-225
tr|A0A7Z2NGU6|A0A7Z2NGU6_9BACL/1-225
tr|A0A5C5SZX3|A0A5C5SZX3_9BACL/1-227
tr|A0A5C5TMX1|A0A5C5TMX1_9BACL/1-227
tr|A0A8T4FXA4|A0A8T4FXA4_9ARCH/1-214
tr|A0A385NJN7|A0A385NJN7_9BACI/1-226
tr|A0A5B0WH95|A0A5B0WH95_9BACL/1-230
tr|A0A5R9FYI8|A0A5R9FYI8_9BACL/1-226
tr|A0A9J6Z922|A0A9J6Z922_9BACL/1-225
tr|A0A8J8SFC7|A0A8J8SFC7_9FIRM/1-216
tr|A0A841YEW2|A0A841YEW2_9LIST/1-214
tr|A0A841ZAG7|A0A841ZAG7_9LIST/1-219
tr|A0A4U0FD26|A0A4U0FD26_9BACL/1-221
tr|A0A174GVY9|A0A174GVY9_9CLOT/1-220
tr|A0A3E3DMT7|A0A3E3DMT7_9CLOT/1-220
tr|A0A3A9SV55|A0A3A9SV55_9FIRM/1-218
tr|A0A355SIP2|A0A355SIP2_9CLOT/1-214
tr|A0A7C6KWV4|A0A7C6KWV4_THESZ/1-219
tr|A0A231VMQ5|A0A231VMQ5_THETR/1-219
tr|A0A6I5A6G2|A0A6I5A6G2_9BACI/1-219
tr|A0A4U2PZ95|A0A4U2PZ95_9BACL/1-218
tr|A0A7H0Y9N3|A0A7H0Y9N3_9BACL/1-221
tr|A0A2K1EKK7|A0A2K1EKK7_9BACL/1-221
tr|A0A8I1LRV1|A0A8I1LRV1_PAEPO/1-221
tr|A0A2N9Z7W2|A0A2N9Z7W2_PAEPO/1-219
tr|A0A2S6NXQ1|A0A2S6NXQ1_9BACL/1-219
tr|A0A3G8R3M7|A0A3G8R3M7_9BACL/1-221
tr|E3ECE7|E3ECE7_PAEPS/1-221
tr|A0A544WX53|A0A544WX53_9BACL/1-221
tr|A0A7X5N9Z2|A0A7X5N9Z2_PAEPO/1-221
tr|A0A5B8J215|A0A5B8J215_PAEPO/1-221
tr|A0A7Y8V6G9|A0A7Y8V6G9_BACSP/1-227
tr|A0A5J4J9N0|A0A5J4J9N0_9BACI/1-227
tr|A0A4V2WNT3|A0A4V2WNT3_9BACL/1-211
tr|A0A7X8IA92|A0A7X8IA92_9FIRM/1-214
tr|A0A7C6RPR9|A0A7C6RPR9_9FIRM/1-215
tr|A0A7C8LG40|A0A7C8LG40_9FIRM/1-215
tr|A0A7X8JQG6|A0A7X8JQG6_9FIRM/1-223
tr|A0A8J8MAD4|A0A8J8MAD4_9FIRM/1-217
tr|A0A2N2AZ87|A0A2N2AZ87_9FIRM/1-218
tr|A0A2N2EJW4|A0A2N2EJW4_9FIRM/1-218
tr|A0A7Y8V382|A0A7Y8V382_BACSP/1-222
tr|A0A415J3V6|A0A415J3V6_BACLI/1-226
tr|A0A7G7UV25|A0A7G7UV25_9BACI/1-226
tr|A0A6G7JAM1|A0A6G7JAM1_9BACI/1-226
tr|A0A6H3BAI2|A0A6H3BAI2_9BACI/1-229
tr|A0A8B5NL04|A0A8B5NL04_BACIU/1-224
tr|A0A6H0H7B8|A0A6H0H7B8_BACIU/1-226
tr|A0A6M3ZJJ7|A0A6M3ZJJ7_BACSU/1-226
tr|A0A8E0S7V4|A0A8E0S7V4_9BACI/1-224
tr|A0A8I1WC61|A0A8I1WC61_BACIU/1-224
tr|A0A6A8FQC0|A0A6A8FQC0_BACIU/1-224
tr|A0A809FTG2|A0A809FTG2_BACIU/1-224
tr|A0A6H0WLV2|A0A6H0WLV2_9BACI/1-223
tr|A0A5D4N8X7|A0A5D4N8X7_BACIU/1-226
tr|G4P132|G4P132_BACS4/1-224
tr|A0A2M8SV51|A0A2M8SV51_9BACI/1-224
tr|A0A3A5I083|A0A3A5I083_BACIU/1-224
tr|A0A0J6HAZ7|A0A0J6HAZ7_9BACI/1-228
tr|A0A1C3S9Q2|A0A1C3S9Q2_9BACI/1-227
tr|A0A5M8RHD3|A0A5M8RHD3_9BACI/1-229
tr|A0A6I7FGW1|A0A6I7FGW1_9BACI/1-221
tr|A0A410DAM6|A0A410DAM6_9BACL/1-220
tr|A0A4Q4IGM8|A0A4Q4IGM8_9BACL/1-220
tr|A0A4Z0GJV0|A0A4Z0GJV0_9BACL/1-219
tr|A0A511ASB5|A0A511ASB5_9LACT/1-220
tr|A0A5R9C516|A0A5R9C516_9LACT/1-220
tr|A0A511H1T7|A0A511H1T7_9LACT/1-220
tr|A0A4Y9J5L2|A0A4Y9J5L2_9LACT/1-217
tr|A0A6M0LAL1|A0A6M0LAL1_9LACT/1-219
tr|A0A7D7G6A6|A0A7D7G6A6_9LACT/1-219
tr|A0A5J6SMN0|A0A5J6SMN0_9BACI/1-227
tr|A0A7H8QFA3|A0A7H8QFA3_9BACL/1-231
tr|A0A3E0K0R9|A0A3E0K0R9_9BACI/1-229
tr|A0A953HZB8|A0A953HZB8_9BACI/1-229
tr|A0A494YTP8|A0A494YTP8_9BACI/1-224
tr|A0A511W697|A0A511W697_9BACI/1-229
tr|A0A7X9KYH6|A0A7X9KYH6_9FIRM/1-225
tr|A0A5C5E644|A0A5C5E644_9LACT/1-221
tr|A0A5C5E5M3|A0A5C5E5M3_9LACT/1-224
tr|A0A847D8P3|A0A847D8P3_9LACT/1-221
tr|A0A6M0YJI3|A0A6M0YJI3_CLOBO/1-215
tr|A0A1Z1SZP3|A0A1Z1SZP3_PROMI/1-214
tr|B4EUM4|B4EUM4_PROMH/1-214
tr|A0A6N8HKD6|A0A6N8HKD6_9BACI/1-223
tr|A0A371J5L0|A0A371J5L0_9FIRM/1-222
tr|A0A031WID8|A0A031WID8_CLODI/1-226
tr|A0A125V6Y2|A0A125V6Y2_CLODI/1-226
tr|Q185X7|Q185X7_CLOD6/1-226
tr|A0A0H3N4M0|A0A0H3N4M0_CLODC/1-226
tr|A0A921SZH2|A0A921SZH2_9FIRM/1-219
tr|A0A943AQN6|A0A943AQN6_9FIRM/1-219
tr|A0A090Y774|A0A090Y774_PAEMA/1-224
tr|A0A3P3U240|A0A3P3U240_9BACL/1-221
tr|A0A3S9V3A1|A0A3S9V3A1_9BACL/1-220
tr|A0A850EMZ4|A0A850EMZ4_9BACL/1-222
tr|A0A1S2LIJ3|A0A1S2LIJ3_9BACI/1-223
tr|A0A417YIL0|A0A417YIL0_9BACI/1-226
tr|A0A6F9XTC2|A0A6F9XTC2_9LACO/1-223

tr|A0A6F9XNZ0|A0A6F9XNZ0_9LACO/1-223
tr|A0A226RGF0|A0A226RGF0_9LACO/1-223
tr|A0A222W1X5|A0A222W1X5_9LACO/1-223
tr|A0A956SN91|A0A956SN91_9LACT/1-223
tr|A0A1D2L6G6|A0A1D2L6G6_BROTH/1-222
tr|A0A2R8A497|A0A2R8A497_CARDV/1-223
tr|A0A5F0MG50|A0A5F0MG50_CARDV/1-223
tr|A0A7Z8CZF0|A0A7Z8CZF0_CARDV/1-223
tr|A0A5F0MN92|A0A5F0MN92_CARDV/1-223
tr|A0A7I0FF23|A0A7I0FF23_CARDV/1-223
tr|A0A3C0SN96|A0A3C0SN96_CLOSP/1-217
tr|E0E247|E0E247_9FIRM/1-213
tr|A0A929VNH4|A0A929VNH4_9FIRM/1-213
tr|A0A135YWA8|A0A135YWA8_9FIRM/1-213
tr|D3MS55|D3MS55_9FIRM/1-213
tr|A0A943ICM4|A0A943ICM4_9FIRM/1-213
tr|A0A6N7XHW8|A0A6N7XHW8_9FIRM/1-213
tr|A0A3Q9I7Q4|A0A3Q9I7Q4_9BACL/1-225
tr|A0A7X2Z4I0|A0A7X2Z4I0_9BACL/1-222
tr|A0A544VH15|A0A544VH15_9BACL/1-227
tr|A0A974PI66|A0A974PI66_9BACL/1-218
tr|A0A358GQ99|A0A358GQ99_PAESP/1-232
tr|A0A7M2ADR2|A0A7M2ADR2_9BACL/1-233
tr|A0A329QHP7|A0A329QHP7_9BACL/1-229
tr|A0A2G7LRA7|A0A2G7LRA7_9BACL/1-225
tr|A0A510TT71|A0A510TT71_9LACO/1-223
tr|A0A351DMZ3|A0A351DMZ3_9LACO/1-223
tr|A0A5P8Q6V6|A0A5P8Q6V6_9LACO/1-223
tr|A0A419AYN1|A0A419AYN1_PECCA/1-223
tr|A0A125W3B5|A0A125W3B5_ENTFL/1-216
tr|A0A429Z7Z7|A0A429Z7Z7_9ENTE/1-215
tr|A0A2N2BKX0|A0A2N2BKX0_9FIRM/1-224
tr|A0A925LG24|A0A925LG24_9FIRM/1-218
tr|D5Q3A9|D5Q3A9_CLODI/1-218
tr|A0A371J728|A0A371J728_9FIRM/1-224
tr|A0A7X8HV95|A0A7X8HV95_9FIRM/1-216
tr|A0A8E3VRQ7|A0A8E3VRQ7_9GAMM/1-215
tr|A0A2X2BX19|A0A2X2BX19_PROMI/1-106

7.13 Y19 sidechain morph animation

Movie S1 – Animation showing domain reorientation and burial of Y19. A morph (generated in PyMOL) between the average coordinates from the ff15ipq-spceb-ecc-2whe simulation (2-3 μs) and the NAC I complex (PDB: 2WF9). The protein surface is shown in grey. A cavity forms in the hinge as the domain rotates, resulting from the displacement of Y19.

8 Bibliography

1. Garcia-Viloca, M., Gao, J., Karplus, M. & Truhlar, D. G. How Enzymes Work: Analysis by Modern Rate Theory and Computer Simulations. *Science* **303**, 186–195 (2004).
2. Warshel, A. *et al.* Electrostatic basis for enzyme catalysis. *Chem Rev* **106**, 3210–3235 (2006).
3. Kamerlin, S. C. L., Sharma, P. K., Chu, Z. T. & Warshel, A. Ketosteroid isomerase provides further support for the idea that enzymes work by electrostatic preorganization. *Proc Natl Acad Sci U S A* **107**, 4075–4080 (2010).
4. Warshel, A., Sharma, P. K., Chu, Z. T. & Åqvist, J. Electrostatic contributions to binding of transition state analogues can be very different from the corresponding contributions to catalysis: Phenolates binding to the oxyanion hole of ketosteroid isomerase. *Biochemistry* **46**, 1466–1476 (2007).
5. Feierberg, I. & Åqvist, J. The catalytic power of ketosteroid isomerase investigated by computer simulation. *Biochemistry* **41**, 15728–15735 (2002).
6. Lameira, J., Bora, R. P., Chu, Z. T. & Warshel, A. Methyltransferases do not work by compression, cratic, or desolvation effects, but by electrostatic preorganization. *Proteins: Structure, Function, and Bioinformatics* **83**, 318–330 (2015).
7. Warshel, A., Strajbl, M., Villa, J. & Florian, J. Remarkable rate enhancement of orotidine 5'-monophosphate decarboxylase is due to transition-state stabilization rather than to ground-state destabilization. *Biochemistry* **39**, 14728–14738 (2000).
8. Cui, Q. & Karplus, M. Triosephosphate isomerase: A theoretical comparison of alternative pathways. *J Am Chem Soc* **123**, 2284–2290 (2001).
9. Liu, H., Zhang, Y. & Yang, W. How is the active site of enolase organized to catalyze two different reaction steps? *J Am Chem Soc* **122**, 6560–6570 (2000).
10. Alhambra, C., Wu, L., Zhang, Z. Y. & Gao, J. Walden-inversion-enforced transition-state stabilization in a protein tyrosine phosphatase. *J Am Chem Soc* **120**, 3858–3866 (1998).
11. Menger, F. M. & Nome, F. Interaction vs Preorganization in Enzyme Catalysis. A Dispute That Calls for Resolution. *ACS Chem Biol* **14**, 1386–1392 (2019).

12. Souza, B. S. *et al.* Transforming a Stable Amide into a Highly Reactive One: Capturing the Essence of Enzymatic Catalysis. *Angewandte Chemie International Edition* **56**, 5345–5348 (2017).

13. Hutchins, R. O. & Rua, L. Neighboring Group Assistance in Azabicyclic Derivatives. Tremendous Rate Accelerations in 2-Aza-6-halobicyclo[2.2.2]- and 6-Aza-4-halobicyclo[3.2.1]octanes. *Journal of Organic Chemistry* **40**, 2567–2568 (1975).

14. Liz, D. G. *et al.* Supramolecular phosphate transfer catalysis by pillar[5]arene. *Chemical Communications* **52**, 3167–3170 (2016).

15. Jindal, G. & Warshel, A. Misunderstanding the preorganization concept can lead to confusions about the origin of enzyme catalysis. *Proteins: Structure, Function, and Bioinformatics* **85**, 2157–2161 (2017).

16. Bruice, T. C. A view at the millennium: The efficiency of enzymatic catalysis. *Acc Chem Res* **35**, 139–148 (2002).

17. Griffin, J. L. *et al.* Near attack conformers dominate β-phosphoglucomutase complexes where geometry and charge distribution reflect those of substrate. *Proc Natl Acad Sci U S A* **109**, 6910–6915 (2012).

18. Johnson, L. A. *et al.* Van der Waals Contact between Nucleophile and Transferring Phosphorus Is Insufficient to Achieve Enzyme Transition-State Architecture. *ACS Catal* **8**, 8140–8153 (2018).

19. Hur, S. & Bruice, T. C. Comparison of formation of reactive conformers (NACs) for the Claisen rearrangement of chorismate to prephenate in water and in the E. coli mutase: The efficiency of the enzyme catalysis. *J Am Chem Soc* **125**, 5964–5972 (2003).

20. Hur, S. & Bruice, T. C. The near attack conformation approach to the study of the chorismate to prephenate reaction. *Proc Natl Acad Sci U S A* **100**, 12015–12020 (2003).

21. Shurki, A., Štrajbl, M., Villà, J. & Warshel, A. How much do enzymes really gain by restraining their reacting fragments? *J Am Chem Soc* **124**, 4097–4107 (2002).

22. Štrajbl, M., Shurki, A., Kato, M. & Warshel, A. Apparent NAC effect in chorismate mutase reflects electrostatic transition state stabilization. *J Am Chem Soc* **125**, 10228–10237 (2003).

23. Jencks, W. P. Binding Energy, Specificity, and Enzymic Catalysis: The Circe Effect. *Adv Enzymol Relat Areas Mol Biol* **43**, 219–410 (2006).

24. Blake, C. C. F. *et al.* Structure of Hen Egg-White Lysozyme: A Three-dimensional Fourier Synthesis at 2 Å Resolution. *Nature 1965 206:4986* **206**, 757–761 (1965).

25. Blake, C. C. *et al.* Crystallographic studies of the activity of hen egg-white lysozyme. *Proc R Soc Lond B Biol Sci* **167**, 378–388 (1967).

26. Yoshioka, Y., Hasegawa, K., Matsuura, Y., Katsube, Y. & Kubota, M. Crystal structures of a mutant maltotetraose-forming exo-amylase cocrystallized with maltopentaose. *J Mol Biol* **271**, 619–628 (1997).

27. Khanjin, N. A., Snyder, J. P. & Menger, F. M. Mechanism of chorismate mutase: Contribution of conformational restriction to catalysis in the Claisen rearrangement. *J Am Chem Soc* **121**, 11831–11846 (1999).

28. Stockwell, G. R. & Thornton, J. M. Conformational Diversity of Ligands Bound to Proteins. *J Mol Biol* **356**, 928–944 (2006).

29. Andrews, L. D., Fenn, T. D. & Herschlag, D. Ground State Destabilization by Anionic Nucleophiles Contributes to the Activity of Phosphoryl Transfer Enzymes. *PLoS Biol* **11**, 1–18 (2013).

30. Ruben, E. A. *et al.* Ground state destabilization from a positioned general base in the Ketosteroid isomerase active site. *Biochemistry* **52**, 1074–1081 (2013).

31. Begley, T. P., Appleby, T. C. & Ealick, S. E. The structural basis for the remarkable catalytic proficiency of orotidine 5'-monophosphate decarboxylase. *Curr Opin Struct Biol* **10**, 711–718 (2000).

32. D'Ordine, R. L., Tonge, P. J., Carey, P. R. & Anderson, V. E. Electronic Rearrangement Induced by Substrate Analog Binding to the Enoyl-CoA Hydratase Active Site: Evidence for Substrate Activation. *Biochemistry* **33**, 12635–12643 (1994).

33. D'Ordine, R. L., Pawlak, J., Bahnson, B. J. & Anderson, V. E. Polarization of cinnamoyl-CoA substrates bound to enoyl-CoA hydratase: Correlation of 13C NMR with quantum mechanical calculations and calculation of electronic strain energy. *Biochemistry* **41**, 2630–2640 (2002).

34. Anderson, V. E. Quantifying energetic contributions to ground state destabilization. *Arch Biochem Biophys* **433**, 27–33 (2005).

35. Warshel, A. & Florián, J. Computer simulations of enzyme catalysis: Finding out what has been optimized by evolution. *Proceedings of the National Academy of Sciences* **95**, 5950–5955 (1998).

36. Warshel, A. Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *Journal of Biological Chemistry* **273**, 27035–27038 (1998).

37. Concluding Remarks. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **317**, 449–451 (1986).

38. Richard, J. P., Amyes, T. L., Goryanova, B. & Zhai, X. Enzyme architecture: on the importance of being in a protein cage. *Curr Opin Chem Biol* **21**, 1–10 (2014).

39. Karp, D. A. *et al.* High Apparent Dielectric Constant Inside a Protein Reflects Structural Reorganization Coupled to the Ionization of an Internal Asp. *Biophys J* **92**, 2041–2053 (2007).

40. Kukic, P. *et al.* Protein dielectric constants determined from NMR chemical shift perturbations. *J Am Chem Soc* **135**, 16968–16976 (2013).

41. Warshel, A., Aqvist, J. & Creighton, S. Enzymes work by solvation substitution rather than by desolvation. *Proceedings of the National Academy of Sciences* **86**, 5820–5824 (1989).

42. Lameira, J., Bora, R. P., Chu, Z. T. & Warshel, A. Methyltransferases do not work by compression, cratic, or desolvation effects, but by electrostatic preorganization. *Proteins: Structure, Function, and Bioinformatics* **83**, 318–330 (2015).

43. Araújo, E., Lima, A. H. & Lameira, J. Catalysis by solvation rather than the desolvation effect: exploring the catalytic efficiency of SAM-dependent chlorinase. *Physical Chemistry Chemical Physics* **19**, 21350–21356 (2017).

44. Henzler-Wildman, K. A. *et al.* Intrinsic motions along an enzymatic reaction trajectory. *Nature 2007 450:7171* **450**, 838–844 (2007).

45. Henzler-Wildman, K. A. *et al.* A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature 2007 450:7171* **450**, 913–916 (2007).

46. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature 2007 450:7172* **450**, 964–972 (2007).

47. Eisenmesser, E. Z. *et al.* Intrinsic dynamics of an enzyme underlies catalysis. *Nature 2005 438:7064* **438**, 117–121 (2005).

48. Eisenmesser, E. Z., Bosco, D. A., Akke, M. & Kern, D. Enzyme dynamics during catalysis. *Science* **295**, 1520–1523 (2002).

49. Warshel, A. & Bora, R. P. Perspective: Defining and quantifying the role of dynamics in enzyme catalysis. *J Chem Phys* **144**, 1–17 (2016).

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

51. Koonin, E. v. & Tatusove, R. L. Computer Analysis of Bacterial Haloacid Dehalogenases Defines a Large Superfamily of Hydrolases with Diverse Specificity: Application of an Iterative Approach to Database Search. *J Mol Biol* **244**, 125–132 (1994).

52. 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* **41**, 8351–8359 (2002).

53. 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 lactus. *Acta Crystallographica Section D* **58**, 324–326 (2002).

54. Levander, F., Andersson, U. & Rådström, P. Physiological Role of β-Phosphoglucomutase in Lactococcus lactis. *Appl Environ Microbiol* **67**, 4546–4553 (2001).

55. Jin, Y., Richards, N. G., Waltho, J. P. & Blackburn, G. M. Metal Fluorides as Analogues for Studies on Phosphoryl Transfer Enzymes. *Angewandte Chemie International Edition* **56**, 4110–4128 (2017).

56. Baxter, N. J. *et al.* Anionic charge is prioritized over geometry in aluminum and magnesium fluoride transition state analogs of phosphoryl transfer enzymes. *J Am Chem Soc* **130**, 3952–3958 (2008).

57. Cliff, M. J. *et al.* Transition state analogue structures of human phosphoglycerate kinase establish the importance of charge balance in catalysis. *J Am Chem Soc* **132**, 6507–6516 (2010).

58. Allen, K. N. & Dunaway-Mariano, D. Response to Comment on ‘The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction’. *Science* **301**, 1184–1184 (2003).

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

60. Tremblay, L. W., Zhang, G., Dai, J., Dunaway-Mariano, D. & Allen, K. N. Chemical confirmation of a pentavalent phosphorane in complex with β-phosphoglucomutase. *J Am Chem Soc* **127**, 5298–5299 (2005).

61. Baxter, N. J. *et al.* A Trojan horse transition state analogue generated by MgF3- formation in an enzyme active site. *Proc Natl Acad Sci U S A* **103**, 14732–14737 (2006).

62. Blackburn, G. M., Williams, N. H., Gamblin, S. J. & Smerdon, S. J. Comment on 'The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction'. *Science* **301**, (2003).

63. Webster, C. E. High-energy intermediate of stable transition state analogue: Theoretical perspective of the active site and mechanism of β-phosphoglucomutase. *J Am Chem Soc* **126**, 6840–6841 (2004).

64. Baxter, N. J. *et al.* Atomic details of near-transition state conformers for enzyme phosphoryl transfer revealed by MgF3- rather than by phosphoranes. *Proc Natl Acad Sci U S A* **107**, 4555–4560 (2010).

65. Jin, Y. *et al.* 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* **111**, 12384–12389 (2014).

66. Barrozo, A. *et al.* Computer simulations of the catalytic mechanism of wild-type and mutant β-phosphoglucomutase. *Org Biomol Chem* **16**, 2060–2073 (2018).

67. Robertson, A. J. *et al.* The Relationship between Enzyme Conformational Change, Proton Transfer, and Phosphoryl Transfer in β-Phosphoglucomutase. *ACS Catal* **11**, 12840–12849 (2021).

68. Cruz-Navarrete, F. A., Baxter, N. J., Wood, H. P., Hounslow, A. M. & Waltho, J. P. 1H, 15N and 13C backbone resonance assignments of the P146A variant of β-phosphoglucomutase from Lactococcus lactis in its substrate-free form. *Biomol NMR Assign* **13**, 349–356 (2019).

69. Wood, H. P. *et al.* Allomorphy as a mechanism of post-translational control of enzyme activity. *Nat Commun* **11**, 5538 (2020).

70. Lee, R. A., Razaz, M. & Hayward, S. The DynDom database of protein domain motions. *Bioinformatics* **19**, 1290–1291 (2003).

71. Qi, G. & Hayward, S. Database of ligand-induced domain movements in enzymes. *BMC Struct Biol* **9**, 1–9 (2009).

72. Qi, G., Lee, R. & Hayward, S. Structural bioinformatics A comprehensive and non-redundant database of protein domain movements. *Bioinformatics* **21**, 2832–2838 (2005).

73. Matthews, B. W. & Remington, S. J. The Three Dimensional Structure of the Lysozyme from Bacteriophage T4. *Proceedings of the National Academy of Sciences* **71**, 4178–4182 (1974).

74. Zhang, X. J., Wozniak, J. A. & Matthews, B. W. Protein Flexibility and Adaptability Seen in 25 Crystal Forms of T4 Lysozyme. *J Mol Biol* **250**, 527–552 (1995).

75. Nauli, S. *et al.* Polymer-driven crystallization. *Protein Science* **16**, 2542–2551 (2007).

76. Goto, N. K., Skrynnikov, N. R., Dahlquist, F. W. & Kay, L. E. What is the average conformation of bacteriophage T4 lysozyme in solution? a domain orientation study using dipolar couplings measured by solution NMR. *J Mol Biol* **308**, 745–764 (2001).

77. Chen, J. L. *et al.* Analysis of the solution conformations of T4 lysozyme by paramagnetic NMR spectroscopy. *Physical Chemistry Chemical Physics* **18**, 5850–5859 (2016).

78. Chen, Y., Hu, D., Vorpagel, E. R. & Lu, H. P. Probing single-molecule T4 lysozyme conformational dynamics by intramolecular fluorescence energy transfer. *Journal of Physical Chemistry B* **107**, 7947–7956 (2003).

79. Choi, Y. *et al.* Single-molecule dynamics of lysozyme processing distinguishes linear and cross-linked peptidoglycan substrates. *J Am Chem Soc* **134**, 2032–2035 (2012).

80. Gondeau, C. *et al.* Molecular basis for the lack of enantioselectivity of human 3-phosphoglycerate kinase. *Nucleic Acids Res* **36**, 3620–3629 (2008).

81. Zerrad, L. *et al.* A spring-loaded release mechanism regulates domain movement and catalysis in phosphoglycerate kinase. *Journal of Biological Chemistry* **286**, 14040–14048 (2011).

82. Palmai, Z. *et al.* Substrate binding modifies the hinge bending characteristics of human 3-phosphoglycerate kinase: A molecular dynamics study. *Proteins: Structure, Function, and Bioinformatics* **77**, 319–329 (2009).

83. Inoue, R. *et al.* Large Domain Fluctuations on 50-ns Timescale Enable Catalytic Activity in Phosphoglycerate Kinase. *Biophys J* **99**, 2309–2317 (2010).

84. Gabba, M. *et al.* Conformational State Distributions and Catalytically Relevant Dynamics of a Hinge-Bending Enzyme Studied by Single-Molecule FRET and a Coarse-Grained Simulation. *Biophys J* **107**, 1913–1923 (2014).

85. Nakasako, M., Fujisawa, T., Adachi, S. I., Kudo, T. & Higuchi, S. Large-scale domain movements and hydration structure changes in the active-site cleft of unligated glutamate dehydrogenase from Thermococcus profundus studied by cryogenic X-ray crystal structure analysis and small-angle X-ray scattering. *Biochemistry* **40**, 3069–3079 (2001).

86. Oroguchi, T. & Nakasako, M. Changes in hydration structure are necessary for collective motions of a multi-domain protein. *Sci Rep* **6**, 1–14 (2016).

87. Oide, M., Kato, T., Oroguchi, T. & Nakasako, M. Energy landscape of domain motion in glutamate dehydrogenase deduced from cryo-electron microscopy. *FEBS J* **287**, 3472–3493 (2020).

88. Shukla, S. *et al.* Differential Substrate Recognition by Maltose Binding Proteins Influenced by Structure and Dynamics. *Biochemistry* **57**, 5864–5876 (2018).

89. Evenäs, J. *et al.* Ligand-induced structural changes to maltodextrin-binding protein as studied by solution NMR spectroscopy. *J Mol Biol* **309**, 961–974 (2001).

90. Tang, C., Schwieters, C. D. & Clore, G. M. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature 2007 449:7165* **449**, 1078–1082 (2007).

91. Bucher, D., Grant, B. J., Markwick, P. R. & McCammon, J. A. Accessing a Hidden Conformation of the Maltose Binding Protein Using Accelerated Molecular Dynamics. *PLoS Comput Biol* **7**, e1002034 (2011).

92. Kim, E. *et al.* A single-molecule dissection of ligand binding to a protein with intrinsic dynamics. *Nat Chem Biol* **9**, 313–318 (2013).

93. Kooshapur, H., Ma, J., Tjandra, N. & Bermejo, G. A. The Open Conformation Satisfies Multiple NMR Experiments on Apo Glutamine-Binding Protein: Overcoming Pitfalls in the Study of Interdomain Dynamics. *Angew Chem Int Ed Engl* **58**, 16899 (2019).

94. Bermejo, G. A., Strub, M. P., Ho, C. & Tjandra, N. Ligand-free open-closed transitions of periplasmic binding proteins: The case of glutamine-binding protein. *Biochemistry* **49**, 1893–1902 (2010).

95. Flocco, M. M. & Mowbray, S. L. The 1.9 A x-ray structure of a closed unliganded form of the periplasmic glucose/galactose receptor from Salmonella typhimurium. *Journal of Biological Chemistry* **269**, 8931–8936 (1994).

96. Oswald, C. *et al.* Crystal Structures of the Choline/Acetylcholine Substrate-binding Protein ChoX from Sinorhizobium meliloti in the Liganded and Unliganded-Closed States. *Journal of Biological Chemistry* **283**, 32848–32859 (2008).

97. Krewulak, K. D., Shepherd, C. M. & Vogel, H. J. Molecular dynamics simulations of the periplasmic ferric-hydroxamate binding protein FhuD. *BioMetals* **18**, 375–386 (2005).

98. Kandt, C., Xu, Z. & Tieleman, D. P. Opening and closing motions in the periplasmic vitamin B12 binding protein BtuF. *Biochemistry* **45**, 13284–13292 (2006).

99. Liu, M. *et al.* Molecular dynamics simulations of the bacterial periplasmic heme binding proteins ShuT and PhuT. *Biophys Chem* **138**, 42–49 (2008).

100. Bowler, M. W. Conformational dynamics in phosphoglycerate kinase, an open and shut case? *FEBS Lett* **587**, 1878–1883 (2013).

101. Smolin, N., Biehl, R., Kneller, G. R., Richter, D. & Smith, J. C. Functional Domain Motions in Proteins on the ∼1–100 ns Timescale: Comparison of Neutron Spin-Echo Spectroscopy of Phosphoglycerate Kinase with Molecular-Dynamics Simulation. *Biophys J* **102**, 1108–1117 (2012).

102. Chandler, D. *Introduction to Modern Statistcal Mechanics* . (Oxford University Press, 1987).

103. Lindahl, E. Molecular dynamics simulations. *Methods in Molecular Biology* **1215**, 3–26 (2015).

104. Lindahl, E. R. Molecular dynamics simulations. *Methods in Molecular Biology* **443**, 3–23 (2008).

105. Eastman, P. *et al.* OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol* **13**, (2017).

106. Van Gunsteren, W. E. & Berendsen, H. J. C. Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. *angewandte chemie international edition* **29**, 992–1023 (1990).

107. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J Chem Theory Comput* **11**, 1864–1874 (2015).

108. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J Chem Phys* **103**, 8577–8593 (1995).

109. Riniker, S. Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview. *J. Chem. Inf. Model.* **58**, 565–578 (2018).

110. D.A. Case *et al. Amber 2021*. (2021).

111. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput* **11**, 3713 (2015).

112. Hornak, V. *et al.* Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins* **65**, 712 (2006).

113. Tian, C. *et al.* Ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J Chem Theory Comput* **16**, 528–552 (2020).

114. Koes, D. R. & Vries, J. K. Evaluating amber force fields using computed NMR chemical shifts. *Proteins* **85**, 1944–1956 (2017).

115. Escalante, D. E. & Ferguson, D. M. Structural modeling and analysis of the SARS-CoV-2 cell entry inhibitor camostat bound to the trypsin-like protease TMPRSS2. *Medicinal Chemistry Research* **30**, 399–409 (2021).

116. Jing, Z. *et al.* Thermodynamics of ion binding and occupancy in potassium channels. *Chem Sci* **12**, 8920–8930 (2021).

117. Duboué-Dijon, E., Javanainen, M., Delcroix, P., Jungwirth, P. & Martinez-Seara, H. A practical guide to biologically relevant molecular simulations with charge scaling for electronic polarization. *J Chem Phys* **153**, 050901 (2020).

118. Timr, Š., Kadlec, J., Srb, P., Ollila, O. H. S. & Jungwirth, P. Calcium Sensing by Recoverin: Effect of Protein Conformation on Ion Affinity. *Journal of Physical Chemistry Letters* **9**, 1613–1619 (2018).

119. Mason, P. E., Jungwirth, P. & Duboué-Dijon, E. Quantifying the Strength of a Salt Bridge by Neutron Scattering and Molecular Dynamics. *Journal of Physical Chemistry Letters* **10**, 3254–3259 (2019).

120. Salomon-Ferrer, R., Götz, A. W., Poole, D., le Grand, S. & Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J Chem Theory Comput* **9**, 3878–3888 (2013).

121. Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O. & Shaw, D. E. Long-timescale molecular dynamics simulations of protein structure and function. *Curr Opin Struct Biol* **19**, 120–127 (2009).

122. Herrera-nieto, pablo, pérez, A. & de fabritiis, G. characterization of partially ordered states in the intrinsically disordered n-terminal domain of p53 using millisecond molecular dynamics simulations. **10**, 12402 (2020).

123. Bussi, G. & Laio, A. Using metadynamics to explore complex free-energy landscapes. *Nature Reviews Physics* **2**, 200–212 (2020).

124. Sicard, F. & Senet, P. Reconstructing the free-energy landscape of Met-enkephalin using dihedral principal component analysis and well-tempered metadynamics. *J Chem Phys* **138**, 235101 (2013).

125. Kimanius, D., Pettersson, I., Schluckebier, G., Lindahl, E. & Andersson, M. SAXS-Guided Metadynamics. *J Chem Theory Comput* **11**, 3491–3498 (2015).

126. Pluhařová, E., Fischer, H. E., Mason, P. E. & Jungwirth, P. Hydration of the chloride ion in concentrated aqueous solutions using neutron scattering and molecular dynamics. *Mol Phys* **112**, 1230–1240 (2014).

127. Duboué-Dijon, E., Mason, P. E., Fischer, H. E. & Jungwirth, P. Hydration and Ion Pairing in Aqueous Mg2+ and Zn2+ Solutions: Force-Field Description Aided by Neutron Scattering Experiments and Ab Initio Molecular Dynamics Simulations. *Journal of Physical Chemistry B* **122**, 3296–3306 (2018).

128. Kohagen, M., Mason, P. E. & Jungwirth, P. Accounting for Electronic Polarization Effects in Aqueous Sodium Chloride via Molecular Dynamics Aided by Neutron Scattering. *Journal of Physical Chemistry B* **120**, 1454–1460 (2016).

129. Debiec, K. T. *et al.* Further along the Road Less Traveled: AMBER ff15ipq, an Original Protein Force Field Built on a Self-Consistent Physical Model. *J Chem Theory Comput* **12**, 3926–3947 (2016).

130. Putnam, D. K., Lowe, E. W. & Meiler, J. Reconstruction of SAXS Profiles from Protein Structures. *Comput Struct Biotechnol J* **8**, (2013).

131. Knight, C. J. & Hub, J. S. WAXSiS: a web server for the calculation of SAXS/WAXS curves based on explicit-solvent molecular dynamics. *Nucleic Acids Res* **43**, 225–230 (2015).

132. García de la Torre, J. & Hernández Cifre, J. G. Hydrodynamic Properties of Biomacromolecules and Macromolecular Complexes: Concepts and Methods. A Tutorial Mini-review. *J Mol Biol* **432**, 2930–2948 (2020).

133. Ortega, A., Amorós, D. & García De La Torre, J. Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophys J* **101**, 892–898 (2011).

134. Behlke, J. & Ristau, O. A new approximate whole boundary solution of the Lamm differential equation for the analysis of sedimentation velocity experiments. *Biophys Chem* **95**, 59–68 (2002).

135. Lipari, G. & Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J Am Chem Soc* **104**, 4546–4559 (1982).

136. García de la Torre, J. & Hernández Cifre, J. G. Hydrodynamic Properties of Biomacromolecules and Macromolecular Complexes: Concepts and Methods. A Tutorial Mini-review. *J Mol Biol* **432**, 2930–2948 (2020).

137. Lakomek, N. A., Ying, J. & Bax, A. Measurement of 15N relaxation rates in perdeuterated proteins by TROSY-based methods. *J Biomol NMR* **53**, 209–221 (2012).

138. d'Auvergne, E. J. & Gooley, P. R. Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *J Biomol NMR* **40**, 107–119 (2008).

139. Robertson, A. J. *et al.* An Enzyme with High Catalytic Proficiency Utilizes Distal Site Substrate Binding Energy to Stabilize the Closed State but at the Expense of Substrate Inhibition. *ACS Catal* **12**, 3149–3164 (2022).

140. Grudinin, S., Garkavenko, M. & Kazennov, A. Pepsi-SAXS: an adaptive method for rapid and accurate computation of small-angle X-ray scattering profiles. *Acta Crystallogr D Struct Biol* **73**, 449–464 (2017).

141. Chen, P. C. & Hub, J. S. Validating Solution Ensembles from Molecular Dynamics Simulation by Wide-Angle X-ray Scattering Data. *Biophys J* **107**, 435–447 (2014).

142. Oelschlaeger, P., Klahn, M., Beard, W. A., Wilson, S. H. & Warshel, A. Magnesium-cationic Dummy Atom Molecules Enhance Representation of DNA Polymerase β in Molecular Dynamics Simulations: Improved Accuracy in Studies of Structural Features and Mutational Effects. *J Mol Biol* **366**, 687–701 (2007).

143. Saxena, A. & Sept, D. Multisite ion models that improve coordination and free energy calculations in molecular dynamics simulations. *J Chem Theory Comput* **9**, 3538–3542 (2013).

144. Rahnamoun, A. *et al.* A Polarizable Cationic Dummy Metal Ion Model. *J. Phys. Chem. Lett* **13**, 2023 (2022).

145. Springs, B. & Haake, P. Equilibrium constants for association of guanidinium and ammonium ions with oxyanions: The effect of changing basicity of the oxyanion. *Bioorg Chem* **6**, 181–190 (1977).

146. Debiec, K. T., Gronenborn, A. M. & Chong, L. T. Evaluating the strength of salt bridges: A comparison of current biomolecular force fields. *Journal of Physical Chemistry B* **118**, 6561–6569 (2014).

147. Tomlinson, J. H., Ullah, S., Hansen, P. E. & Williamson, M. P. Characterization of salt bridges to lysines in the protein G B1 domain. *J Am Chem Soc* **131**, 4674–4684 (2009).

148. Ahmed, M. C., Papaleo, E. & Lindorff-Larsen, K. How well do force fields capture the strength of salt bridges in proteins? *PeerJ* **2018**, e4967 (2018).

149. Fernandez, P. L., Nagorski, R. W., Cristobal, J. R., Amyes, T. L. & Richard, J. P. Phosphodianion activation of enzymes for catalysis of central metabolic reactions. *J Am Chem Soc* **143**, 2694–2698 (2021).

150. Amyes, T. L. & Richard, J. P. Enzymatic catalysis of proton transfer at carbon: Activation of triosephosphate isomerase by phosphite dianion. *Biochemistry* **46**, 5841–5854 (2007).

151. Amyes, T. L., Richard, J. P. & Tait, J. J. Activation of orotidine 5′-monophosphate decarboxylase by phosphite dianion: The whole substrate is the sum of two parts. *J Am Chem Soc* **127**, 15708–15709 (2005).

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

153. Ray, W. J. & Long, J. W. Thermodynamics and Mechanism of the PO3 Transfer Process in the Phosphoglucomutase Reaction. *Biochemistry* **15**, 3993–4006 (1976).

154. Richard, J. P. Enabling Role of Ligand-Driven Conformational Changes in Enzyme Evolution. *Biochemistry* **61**, 1533–1542 (2022).

155. Reyes, A. C., Koudelka, A. P., Amyes, T. L. & Richard, J. P. Enzyme architecture: Optimization of transition state stabilization from a cation-phosphodianion pair. *J Am Chem Soc* **137**, 5312–5315 (2015).

156. Goryanova, B., Goldman, L. M., Amyes, T. L., Gerlt, J. A. & Richard, J. P. Role of a guanidinium cation-phosphodianion pair in stabilizing the vinyl carbanion intermediate

of orotidine 5′-phosphate decarboxylase-catalyzed reactions. *Biochemistry* **52**, 7500–7511 (2013).

157. Go, M. K., Koudelka, A., Amyes, T. L. & Richard, J. P. Role of Lys-12 in catalysis by triosephosphate isomerase: A two-part substrate approach. *Biochemistry* **49**, 5377–5389 (2010).

158. Kühlbrandt, W. Structure and Mechanisms of F-Type ATP Synthases. *Annu Rev Biochem* **88**, 515–549 (2019).

159. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature 2021 596:7873* **596**, 583–589 (2021).

160. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).

161. Lin, Z. *et al.* Language models of protein sequences at the scale of evolution enable accurate structure prediction. doi:10.1101/2022.07.20.500902.

162. David, A., Islam, S., Tankhilevich, E. & Sternberg, M. J. E. The AlphaFold Database of Protein Structures: A Biologist's Guide. *J Mol Biol* **434**, 167336 (2022).

163. Urban, P. & Pompon, D. Confrontation of AlphaFold models with experimental structures enlightens conformational dynamics supporting CYP102A1 functions. *Scientific Reports 2022 12:1* **12**, 1–13 (2022).

164. Sedinkin, S. L., Burns, D., Shukla, D., Potoyan, D. A. & Venditti, V. Solution Structure Ensembles of the Open and Closed Forms of the ∼130 kDa Enzyme I via AlphaFold Modeling, Coarse Grained Simulations, and NMR. *J Am Chem Soc* **145**, 13347–13356 (2023).

165. Grothaus, I. L., Bussi, G. & Colombi Ciacchi, L. Exploration, Representation, and Rationalization of the Conformational Phase Space of N-Glycans. *J Chem Inf Model* **62**, 4992–5008 (2022).

166. Wood, H. P. *et al.* Enzymatic production of β-glucose 1,6-bisphosphate through manipulation of catalytic magnesium coordination. *Green Chemistry* **23**, 752–762 (2021).

167. Rodrigues, J. P. G. L. M., Teixeira, J. M. C., Trellet, M. & Bonvin, A. M. J. J. pdb-tools: a swiss army knife for molecular structures. *F1000Res* **7**, (2018).

168. Duboué-Dijon, E. *et al.* Binding of Divalent Cations to Insulin: Capillary Electrophoresis and Molecular Simulations. *Journal of Physical Chemistry B* **122**, 5640–5648 (2018).

169. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* **32**, 2319–2327 (2011).

170. Gowers, R. J. *et al.* MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proceedings of the 15th Python in Science Conference* 98–105 (2016).

171. Manalastas-Cantos, K. *et al.* ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *J Appl Crystallogr* **54**, 343–355 (2021).

172. Philo, J. S. An improved function for fitting sedimentation velocity data for low-molecular-weight solutes. *Biophys J* **72**, 435–444 (1997).

173. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606–1619 (2000).

174. d'Auvergne, E. J. & Gooley, P. R. 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**, 121–133 (2008).

175. Winter, G. Xia2: An expert system for macromolecular crystallography data reduction. *J Appl Crystallogr* **43**, 186–190 (2010).

176. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallographica Section D* **66**, 22–25 (2010).

177. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta crystallographica Section D* **53**, 240–255 (1997).

178. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D* **66**, 12–21 (2010).

179. The PyMOL Molecular Graphics System, Version 2.4.1 Schrödinger, LLC.

180. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

181. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods 2020 17:3* **17**, 261–272 (2020).

182. Golicnik, M. *et al.* Kinetic analysis of β-phosphoglucomutase and its inhibition by magnesium fluoride. *J Am Chem Soc* **131**, 1575–1588 (2009).

183. Madeira, F. *et al.* Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res* **50**, W276–W279 (2022).

184. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

185. Livingstone, C. D. & Barton, G. J. Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *Comput Appl Biosci* **9**, 745–756 (1993).