Observing enzyme ternary transition state analogue complexes by $^{19}$F NMR spectroscopy

Anna Ampaw, a Madison Carroll, a Jill von Velsen, b Debabrata Bhattachari, c Alejandro Cohen, d Matthew W. Bowler, b David L. Jakeman a,c*

a. Department of Chemistry, Dalhousie University, Halifax, NS, Canada B3H 4R2.
b. European Molecular Biology Laboratory, Grenoble Outstation, 71 avenue des Martyrs, CS 90181 F-38042 Grenoble, France
c. College of Pharmacy, Dalhousie University, Halifax, NS, Canada B3H 4R2. E-mail: david.jakeman@dal.ca
d. Proteomics and Mass Spectrometry Core Facility, Life Sciences Research Institute, Dalhousie University, Halifax, NS, Canada B3H 4R2.

Fig. S1. Crystal structure of the $\beta$PGM-MgF$_3$-$\beta$G1CP TSA complex (4C4R). The cap domain housing W24 is shown in magenta and the core domain housing W216 is shown in green. Both Trp residues are highlighted with sticks. Asp8 residue responsible for the phosphorylation of the substrate lies on the core domain and $\beta$G1CP ligand is positioned in the active site at the interface of both domains. The atoms of MgF$_3$ are represented with spheres.
**Fig. S2.** IC₅₀ plots of βG1CP (circle and solid line) and βG1CF₃P (diamond and dashed line) as inhibitors of βPGM.

**Fig. S3.** 1% agarose gel of pET-22b_pgmB and pET-22b_pgmB-His digested with SacI and PstI. (A) 1 kb DNA ladder; (B) pET-22b_pgmB digested with PstI and SacI; (C) pET-22b_pgmB-His digested with PstI and SacI;
Fig. S4. 1% agarose gel of pET-22b_pgmB-His and pET-22b_pgmB-His_W24F plasmid digested with BglI and BstEII. (A) 1 kb DNA ladder; (B), (C), (D) pET-22b_pgmB_His_W24F digested with BglI and BstEII; (E) pET-22b_pgmB digested with BglI and BstEII.

Fig. S5. 1% agarose gel of pET-22b_pgmB-His_W216F plasmid digested with PstI. (A) 1 kb DNA ladder; (B) and (C) pET-22b_pgmB_His digested with PstI; (D) and (E) pET-22b_pgmB_His_W24F digested with PstI.
Fig. S6. SDS-PAGE analysis of \( \beta \text{PGM} \) purified from \textit{E. coli BL21(DE3)}. (A) NEB Biolabs broad range prestained protein marker (7-175 kDa); (B) unlabeled \( \beta \text{PGM} \); (C) 5FW\( \beta \text{PGM} \) expressed with 5-fluorotryptophan (method B) (D) W24F, 5FW\( \beta \text{PGM} \) expressed with 5-fluorotryptophan (method B).

Fig. S7. SDS-PAGE analysis of 5FW\( \beta \text{PGM} \) W216F purified from \textit{E. coli BL21(DE3)}. (A) NEB Biolabs broad range prestained protein marker (11-245 kDa); (B) and (C) 5FW\( \beta \text{PGM} \) W216F (MW= 25 kDa).
Fig. S8. Overlaid $^1$H-$^{15}$N HSQC of $^{15}$N-labeled βPGM (500 µM) (blue), $^{15}$N-labeled 5FWβPGM from method A (500 µM) (red) and method B (500 µM) (green). Both $^{15}$N-labeled 5FWβPGM spectra (red and green) show the absence of cross peak resonances for the side chain Trp residues (black box).
Fig. S9. $^1$H-$^{15}$N HSQC of $^{15}$N-labeled βPGM (500 µM). Cross peaks for W24 and W216 side chain residues are outlined (black box). Enzyme solution contains 5 mM MgCl$_2$ and 10% D$_2$O in 50 mM HEPES pH 7.2.
Fig. S10. $^1$H-$^{15}$N HSQC of $^{15}$N-labeled 5FWβPGM from method A (500 µM). Enzyme solution contains 5 mM MgCl$_2$ and 10% D$_2$O in 50 mM HEPES pH 7.2.
Fig. S11. $^1$H-15N HSQC of 15N-labeled 5FWβPGM from method A (500 μM). Enzyme solution contains 5 mM MgCl$_2$ and 10% D$_2$O in 50 mM HEPES pH 7.2.
**Fig. S12.** $^{19}$F NMR spectra of step 1 and step 2 TSA complexes with $^{19}$F- labeled wild-type βPGM. Red boxes (■) are apo-$^{5}$FW24 resonances, blue asterisks are TSA $^{5}$FW24 resonances (★), black circles (•) are free MgF$_3$ (-155 ppm), blue inverse triangles (▼) are TSA ligand resonances, and red triangles (▲) are free ligand resonances. Resonances at -119 are from free F$^{-}$. Assignment of MgF$_3$ and AlF$_4$ resonances were previously determined by chemical shift.$^{3}$ (A) $^{5}$FWβPGM; (B) $^{5}$FWβPGM-MgF$_3$-G6P TSA2 complex; (C) $^{5}$FWβPGM-AlF$_4$-G6P TSA2 complex; (D) $^{5}$FWβPGM-MgF$_3$-βG1CP TSA1 complex; (E) $^{5}$FWβPGM-AlF$_4$-βG1CP TSA1 complex; (F) $^{5}$FWβPGM-MgF$_3$-βG1CF$_3$P TSA1 complex; (G) $^{5}$FWβPGM-AlF$_4$-βG1CF$_3$P TSA1 complex; Sample A contains 0.5 mM enzyme, 5 mM MgCl$_2$, 10 mM NH$_4$F, and 10% D$_2$O in 50 mM Hepes pH 7.2. Samples B-G contain 1 mM enzyme, 5 mM substrate, 5 mM MgCl$_2$, 10 mM NH$_4$F, and 10% D$_2$O in 50 mM Hepes pH 7.2. Samples C, E, and G also contain 1 mM of AlCl$_3$. 
Fig. S13. $^{19}$F NMR spectra of the wild-type 5FWβPGM, the 5FWβPGM W216F mutant, and its TSA1 and TSA2 complexes. Blue asterisks (**) are TSA 5FW24 resonances, red squares (**) are apo 5FW24 resonances, black circles (*) are free MgF$_3$, blue inverse triangles (▼) are complexed G1C$_3$FP ligands, red triangles (▲) are free G1C$_3$FP. (A) wild-type 5FW24 5FW216 βPGM; (B) W216F 5FWβPGM; (C) W216F 5FWβPGM with NH$_4$F and MgCl$_2$; (D) W216F 5FWβPGM-MgF$_3$-G6P TSA2 complex; (E) W216F 5FWβPGM-MgF$_3$-G1CP TSA1 complex; (F) W216F 5FWβPGM-MgF$_3$-G1C$_3$FP TSA1 complex; (G) W216F 5FWβPGM-AlF$_4$-G6P TSA2 complex; (H) W216F 5FWβPGM-AlF$_4$-G1CP TSA1 complex; (I) W216F 5FWβPGM-AlF$_4$-G1C$_3$FP TSA2 complex.
LC-MS/MS Data

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis of $^{19}$F-labeled βPGM was performed for method A (5-fluoroindole incorporation) and method B (5-fluorotryptophan incorporation) βPGM. Gel bands from SDS-PAGE containing $^{19}$F-βPGM were excised from the gel and rinsed with deionized water. Excised gel bands were processed using Investigator ProGest automated system (Genomic Solutions). Samples were reduced using 100μL of 10 mM dithiothreitol (DTT), alkylated with 100μL of 55 mM iodoacetamide and finally digested with Lys-C. The peptides were extracted from the gel pieces by three 20 min incubations with a solution (30μL) containing acetonitrile (50%) and formic acid (5%) in LC-MS-grade water with gentle agitation. The extracts were pooled and dried using a vacuum concentrator (Speed Vac Concentrator, SPD 111 V-230, Thermo Electron Corp.) and finally resuspended in LC-MS-grade water (15μL) containing acetonitrile (3%) and formic acid (0.5%). LC-MS/MS was performed using a nano flow liquid chromatography system (Ultimate3000RSLnanno, ThermoScientific) interfaced to a hybrid ion trap-orbitrap high resolution tandem mass spectrometer (VelosPro, ThermoScientific) operated in data-dependent acquisition (DDA) mode. 1μL of each sample was injected onto a capillary column (50cm x 75μm PicoTip/PicoFrit Self packed column with Jupiter C18 4μ chromatographic media, Phenomenex) at a flow rate of 300 nL min$^{-1}$. Samples were electro-sprayed at 1.2 kV using a dynamic nanospray probe. Chromatographic separation was carried out using 90 min linear gradients (Mobile Phase A: 0.1% formic acid in MS-grade water, mobile phase B: 0.1% formic acid in MS-grade acetonitrile,) from 3% B to 35% B over 60 min, then increasing to 95% B over 5 min. MS/MS spectra were acquired using both collision induced dissociation (CID) and higher-energy collisional dissociation (HCD) for the top 15 peaks in the survey 30 000 resolution MS scan. The raw files were acquired (Xcalibur, ThermoFisher) and exported to Proteome Discoverer 2.0 (ThermoFisher) software for peptide and protein identification using SequestHT search algorithm. Extracted ion chromatographs (XICs) were obtained using XCalibur software (ThermoFisher) by screening for specific peptide masses using 50 ppm extraction windows. The intensity of the XIC of peptide fragments containing W24 and 5FW24 (AVLFDDLGVITDTAEYHFRAWK) were compared quantitatively, as well as the peptide fragments containing W216 and 5FW216 (EVWLLQK).

For method A, 5FW24 and W24 fragments (Fig. S14) suggested a 5-fluoroindole incorporation of approximately 70%, while 5FW216 and W216 fragments (Fig. S15) suggested a 5-fluoroindole incorporation of approximately 50%. However, in the XIC for the W216 peptide, a peak is seen at 35.50 min, the retention time at which the 5FW216 peptide elutes. We hypothesized that this could be a result of the loss of water from the 5FW216 peptide, therefore exhibiting a mass equal to the W216 peptide. Conclusively, this hypothesis would explain the discrepancy between the values observed. For method B, 5FW24 and W24 fragments (Fig. S16) suggested a 5-fluorotryptophan incorporation of approximately 85%.
5FW216 and W216 fragments (Fig. S17) suggested a similar 5-fluorotryptophan incorporation of approximately 80%, however, similar to method A, a minor peak is seen at 35.03 min, the retention time at which the 5FW216 peptide elutes. Again, this justifies the discrepancy in the values.

Fig. S14. Selective XIC of method A 19F-βPGM showing peptides containing FW24 (A) and W24 (B). The peak intensity of (B) is approximately 30% the intensity of (A) implying that 70% of the peptide exists in the fluorinated form (FW24). XIC is inclusive of m/z for +2, +3, +4, and +5 charges of peptide and contains a 50 ppm mass window.
Fig. S15. Selective XIC of method A $^{19}$F-βPGM showing peptides containing FW216 (A) and W216 (B). The peak intensity of (B) is approximately the same intensity as (A) implying that the peptide containing FW24 is present in the same ratio as the peptide containing W24. XIC is inclusive of m/z for +2 charge of peptide and contains a 100 ppm mass window.
Fig. S16. Selective XIC of method B $^{19}$F-βPGM showing peptides containing FW24 (A) and W24 (B). The peak intensity of (B) is approximately 15% the intensity of (A) implying that 85% of the peptide exists in the fluorinated form (FW24). XIC is inclusive of m/z for +2, +3, +4, and +5 charges of peptide and contains a 50 ppm mass window.
Fig. S17. Selective XIC of method B $^{19}$F-$\beta$PGM showing peptides containing FW216 (A) and W216 (B). The peak intensity of (B) is approximately 20% the intensity of (A) implying that 80% of the peptide exists in the fluorinated form (FW24). XIC is inclusive of m/z for +2 charge of peptide and contains a 100 ppm mass window.
Fig. S18. $^{19}$F NMR of 5FWβPGM in 100% D$_2$O (A) and 10% D$_2$O (B). Samples contain 0.5 mM 5FWβPGM, 5 mM MgCl$_2$ and 10 mM NH$_4$F. Sample A contains 50 mM Hepes buffer pH 7.2.

Among the various $^{19}$F NMR spectra that were recorded, we also noticed that the F' (~-119 ppm) and MgF$_x$ (~-155 ppm) resonances were not perfectly aligned. We hypothesized that these differences may be a result of solvent induced isotope shifts (SIIS) due to slight variations in the amount of D$_2$O in each sample. This was confirmed by a $^{19}$F NMR spectra of 5FWβPGM in 100% D$_2$O compared to the $^{19}$F NMR spectra in 10% D$_2$O (SI, Fig. S16). The $^{19}$F NMR spectra exhibited a 2.8 ppm change in chemical shift upfield for the F' resonance and a 1.7 ppm change in chemical shift upfield for the MgF$_x$ resonance, while the chemical shifts for 5FW24 and 5FW216 only exhibited minor changes (0.1-0.2 ppm). The proportions of the two resonances observed for 5FW216 were also not affected by the increase in D$_2$O in the sample suggesting that a second resonance observed for 5FW216 is not a result of SIIS.
|                         | Apo-protein   | 5FW8PGM-MgF\textsubscript{3}-G6P complex | 5FW8PGM-MgF\textsubscript{3}-G6P complex |
|-------------------------|---------------|-----------------------------------------|-----------------------------------------|
| **PDB code**            | 5OLW          | 5OLX                                    | 5OLY                                    |
| **X-ray source**        | MASSIF-1 (ESRF, France) | MASSIF-1 (ESRF, France) | MASSIF-1 (ESRF, France) |
| **Number of images**    | 887           | 2660                                    | 2280                                    |
| **Oscillation step (deg)** | 0.15          | 0.05                                    | 0.05                                    |
| **Wavelength (Å)**      | 0.966         | 0.966                                   | 0.966                                   |
| **Exposure per frame (s)** | 0.196         | 0.02                                    | 0.02                                    |
| **Indexing and Scaling** |               |                                         |                                         |
| **Space group**         | $P_{2_1}2_12_1$ | $P_{2_1}2_12_1$ | $P_{2_1}$  |
| **Unit cell dimensions a, b, c (Å)** | a = 61.0, b = 73.9, c = 120.4 | a = 37.5, b = 54.2, c = 104.2 | a = 57.7, b = 36.7, c = 107.8 |
| **Resolution Range (Å)** | 47.22 - 2.28  | 52.1 - 1.38                            | 44.47 - 2.01                            |
| **Wilson B-factor (Å\textsuperscript{2})** | 28.52         | 9.14                                    | 11.52                                   |
| **Completeness (%)**    | 99.7          | 98.0                                    | 91.9                                    |
| **I over Sigma (I)**    | 8.9           | 6.7                                     | 10.9                                    |
| **Multiplicity**        | 3.9           | 4.8                                     | 2.8                                     |
| **Refinement**          |               |                                         |                                         |
| **Resolution Limits (Å)** | 20.0-2.28     | 20.0-1.38                              | 20.0-2.01                               |
| **R\textsubscript{work}/R\textsubscript{free} (%)** | 20.25/26.59  | 15.03/19.29                            | 21.25/25.82                             |
| **Bond lengths (Å)**    | 0.013         | 0.008                                   | 0.013                                   |
| **Bond Angles (deg)**   | 1.337         | 1.068                                   | 1.292                                   |
| **Average B-factor (Å\textsuperscript{2})** | 46.98         | 13.95                                  | 17.60                                   |
| **Ramachandran favored (%)** | 95.0          | 97.8                                    | 97.0                                    |
| **Ramachandran outliers (%)** | 1.3          | 0.0                                     | 0.0                                     |
| **Clashscore**          | 7.90          | 2.00                                    | 5.21                                    |
Figure S19. Alternative conformations of 5FW216 observed in the monoclinic crystal form. Top: The two molecules on the asymmetric unit are superimposed showing the alternative orientations. The protein adopts exactly the same conformation but can accommodate the two positions of 5FW216. Bottom: Electron density (blue mesh contoured at 1.5σ) for the 2 conformations observed for 5FW216 in the monoclinic crystal form of the closed conformation structure of βPGM. The two molecules in the asymmetric unit show different conformations of the labelled amino acid.
Figure S20. Alternative conformation of 5FW24 in the cap domain of the open structure. Top: The open structure of 5FWβPGM (green) is shown aligned to the open unlabeled structure (cyan). An alternative conformation is observed for the 5FW24, not seen in the unlabeled protein. The occupancy of the alternative conformation is very low (10%). Bottom: Simulated annealing omit map (green mesh contoured at 2.5σ) showing the density for 5FW24 in the open apo-structure of βPGM. The second conformation is at an occupancy of 10%