Supplementary Information

Precise Probing of Residue Roles by Post-translational β,γ-C,N Aza-Michael Mutagenesis in Enzyme Active Sites

Jitka Dadová, Kuan-Jung Wu, Patrick G. Isenegger, James C. Errey, Gonçalo J. L. Bernardes, Justin M. Chalker, Lluís Raich, Carme Rovira and Benjamin G. Davis*

Table of contents

1 Supplementary Figures

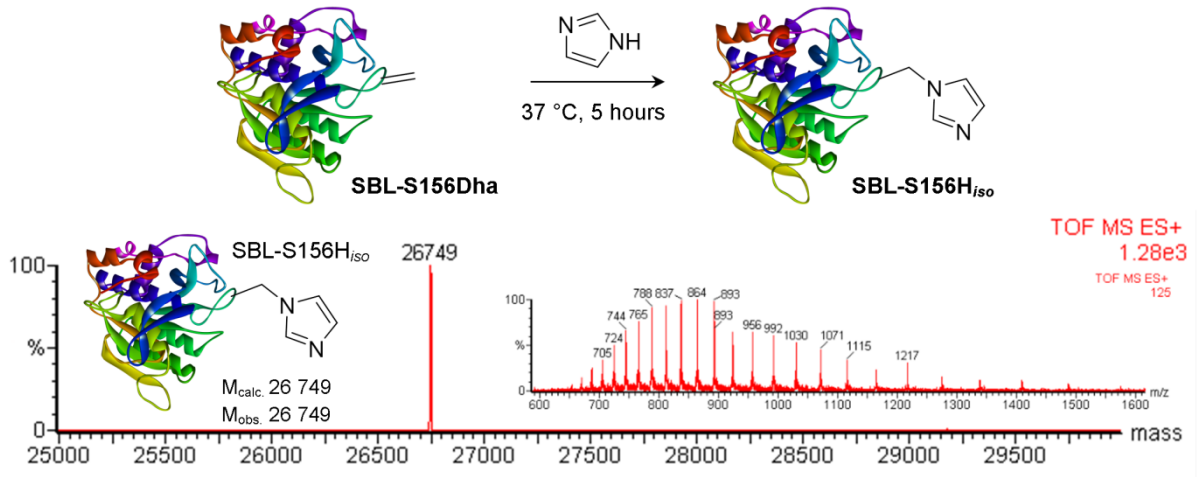


Figure S1. Aza-Michael addition of imidazole to SBL-S156Dha.

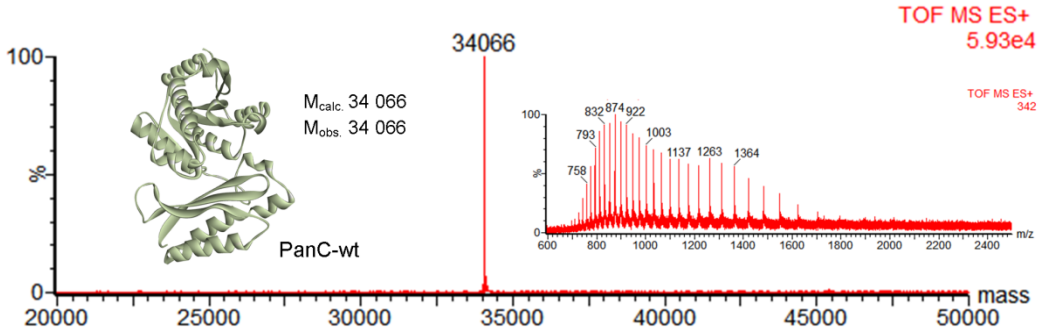


Figure S2. Mass spectrum of wild-type pantothenate synthetase (PanC-wt).

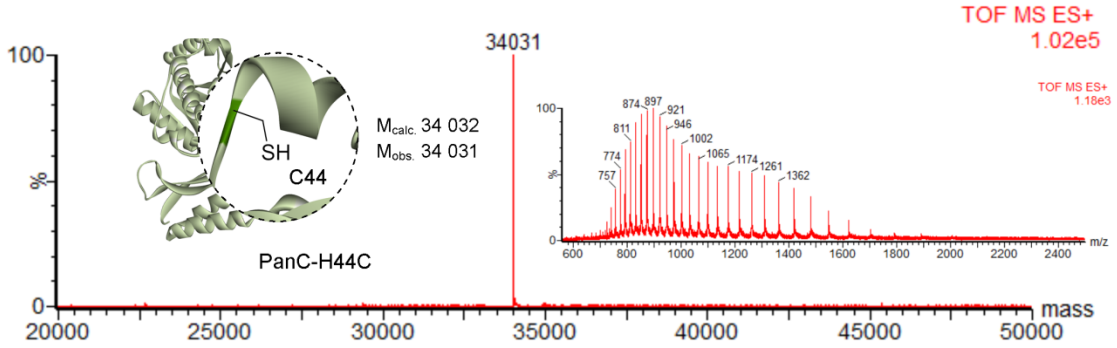


Figure S3. Mass spectrum of PanC-H44C mutant.

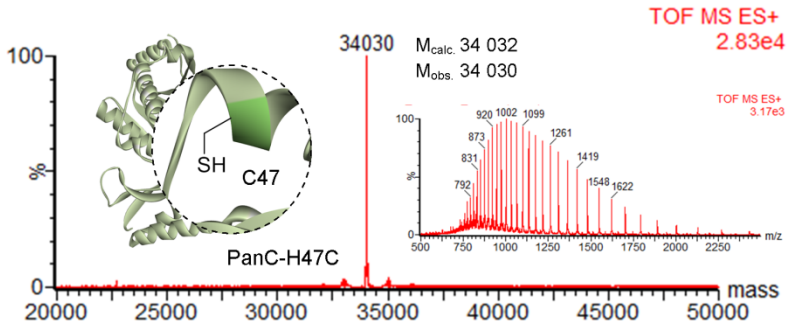


Figure S4. Mass spectrum of PanC-H47C mutant.

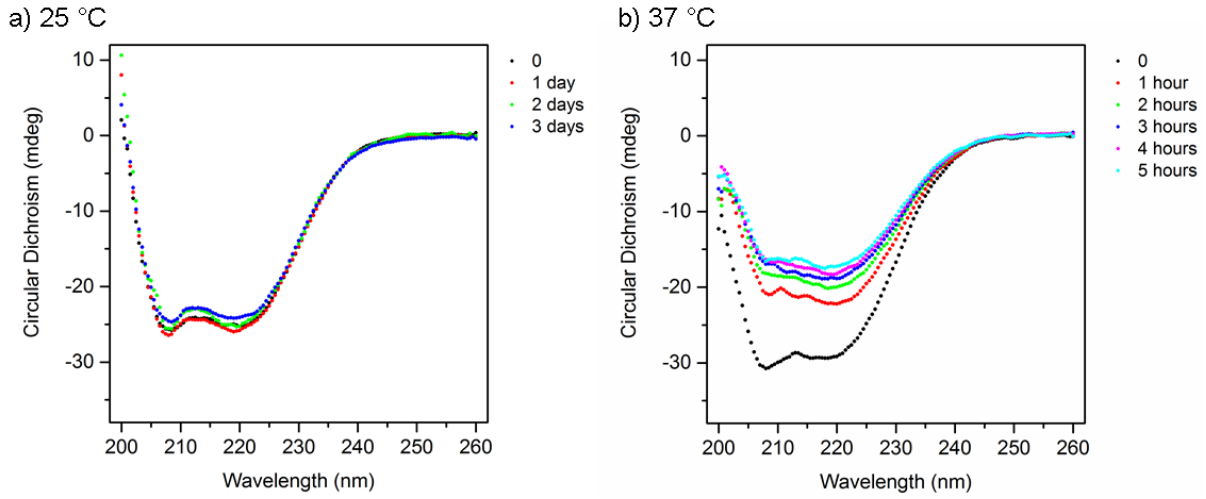


Figure S5. Stability of PanC at a) 25 °C, b) 37 °C studied by CD spectroscopy.

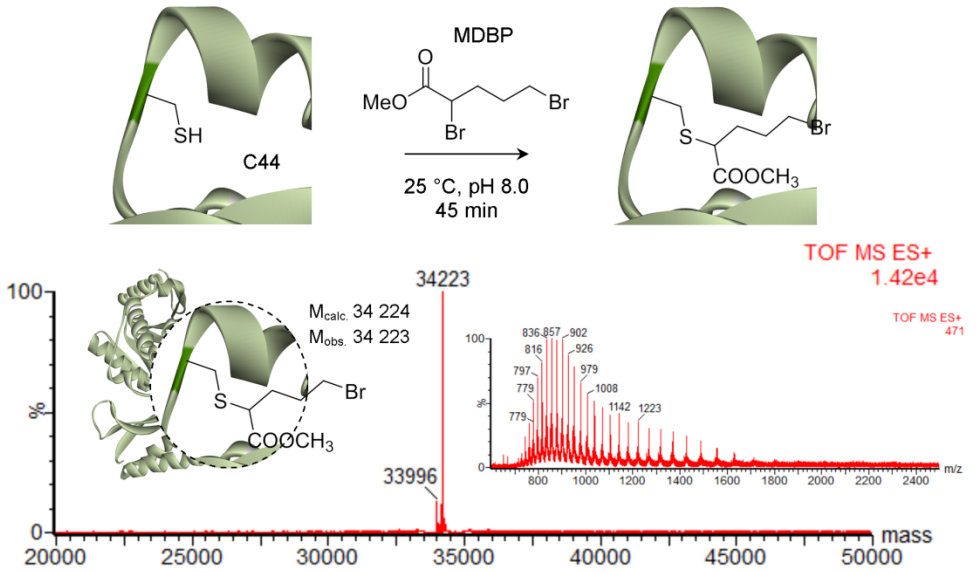


Figure S6. Alkylation of PanC-H44C with methyl 2,5-dibromopentanoate (MDBP).

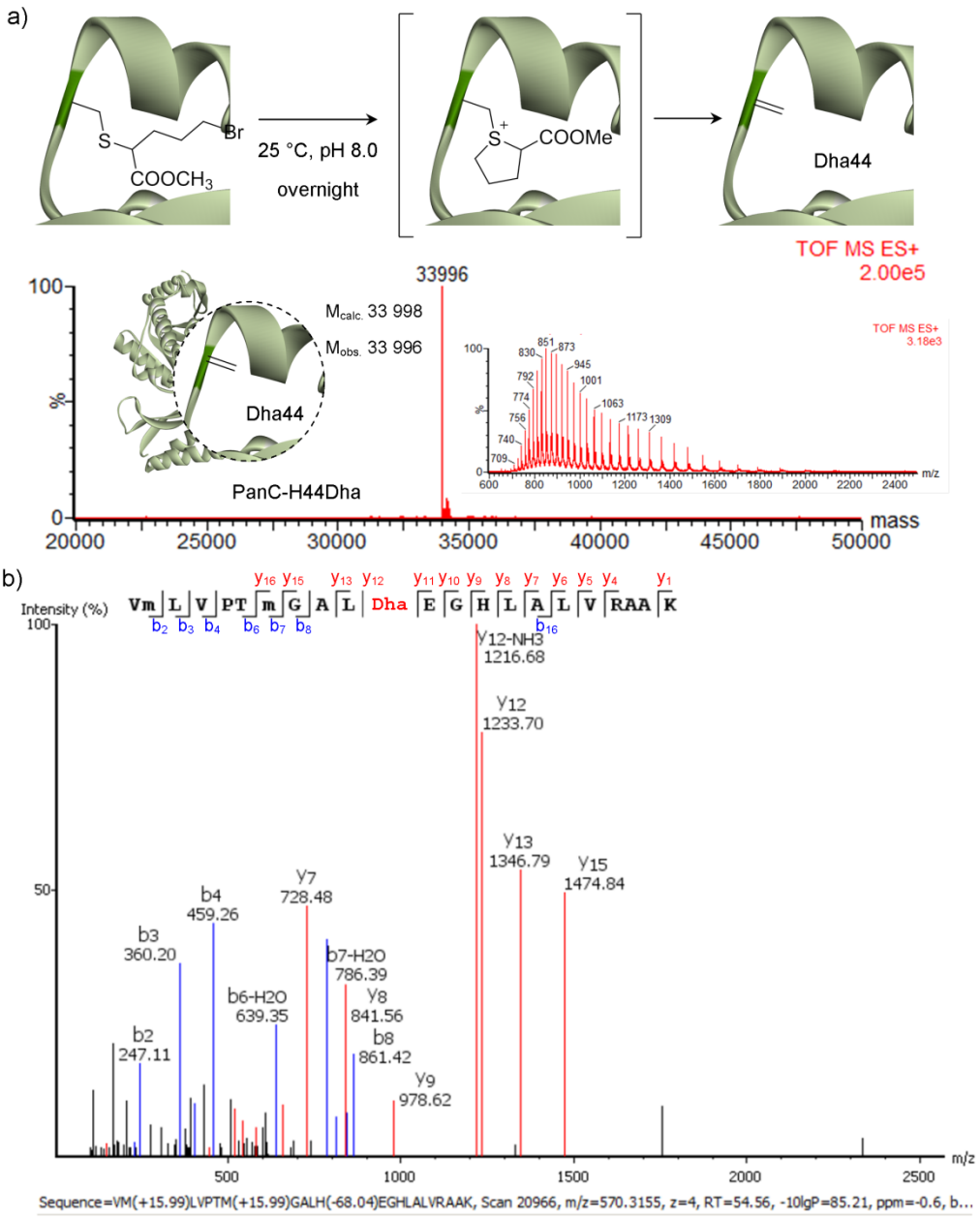


Figure S7. a) Formation of PanC-H44Dha; b) MS/MS analysis of peptide containing Dha44 (trypsin digest of PanC-H44Dha).

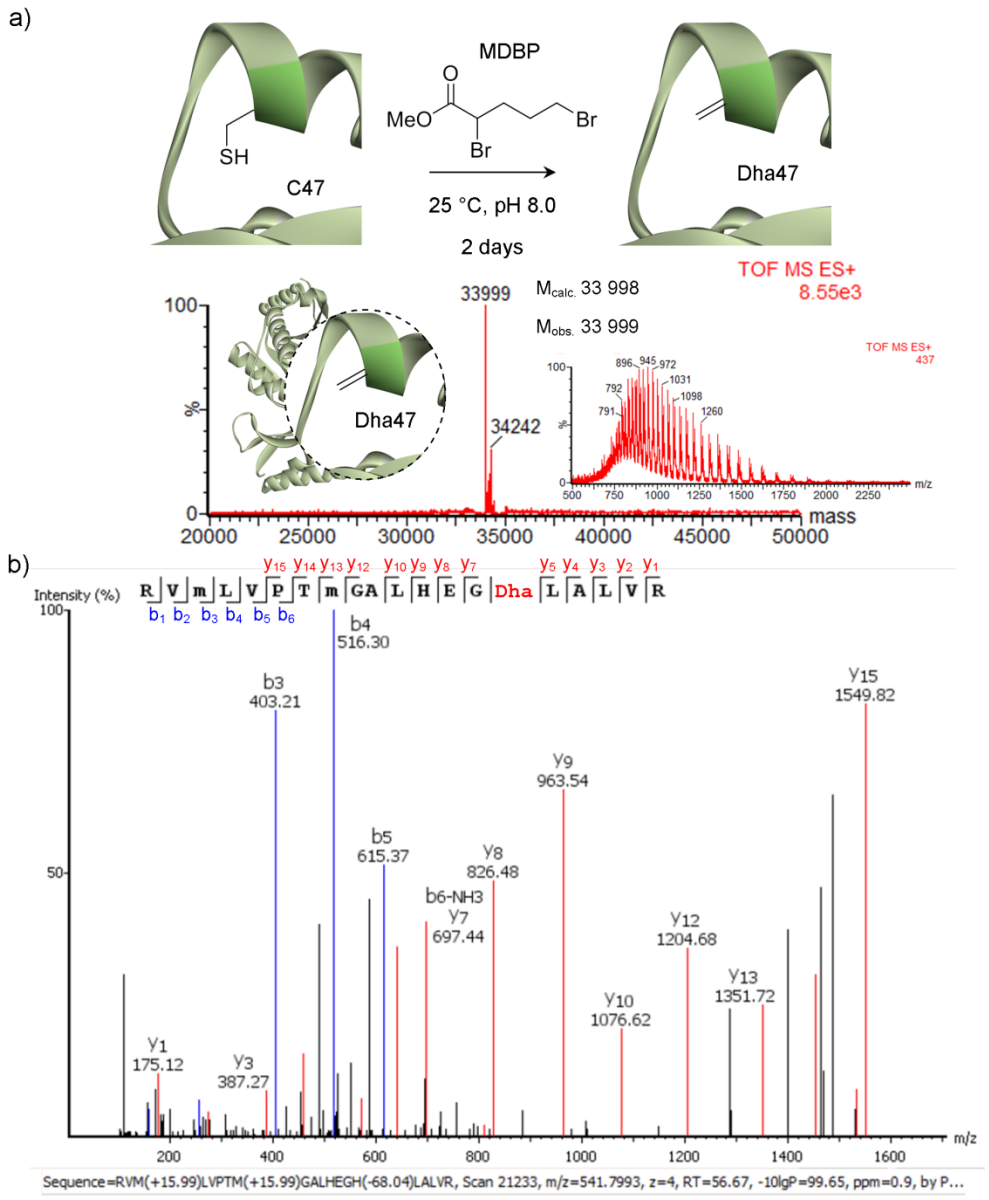


Figure S8. a) Mass spectrum of PanC-H47Dha; b) MS/MS analysis of peptide containing Dha47 (trypsin digest of PanC-H47Dha).

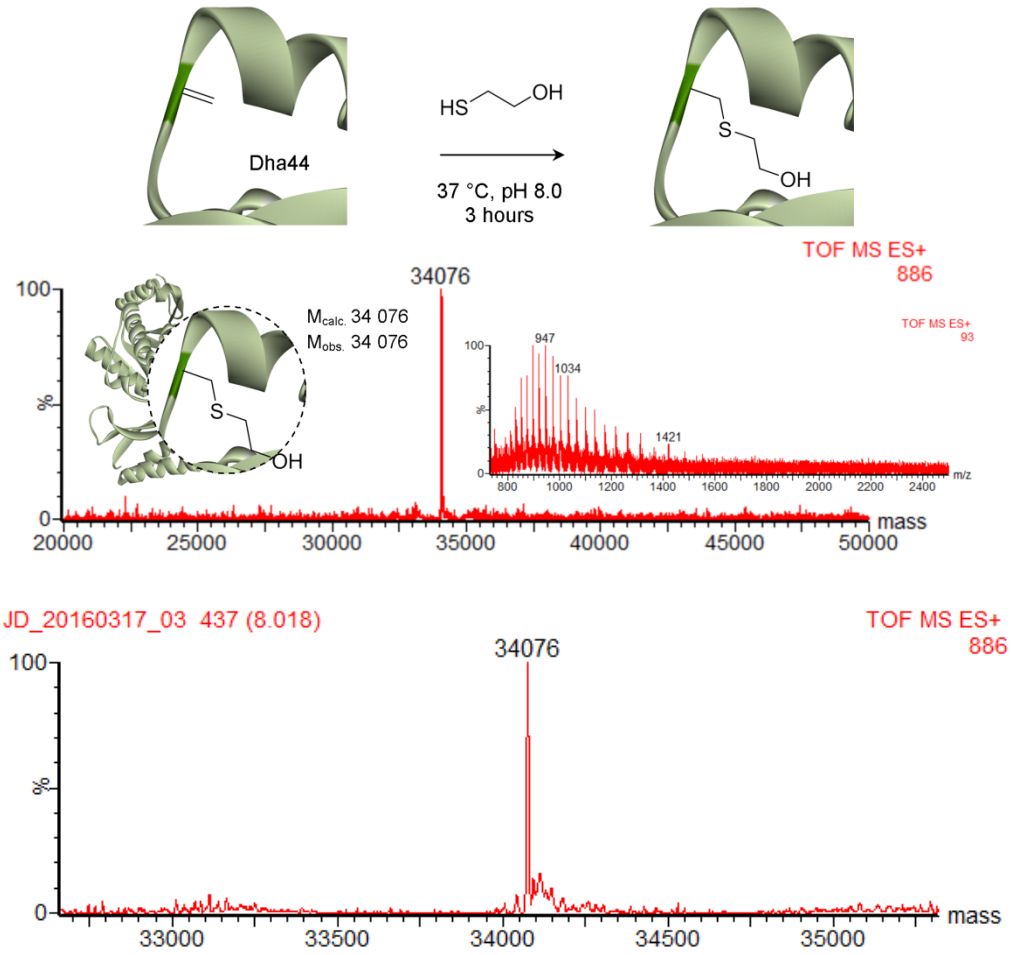


Figure S9. Control thia-Michael addition of *β*-mercaptoethanol to PanC-H44Dha.

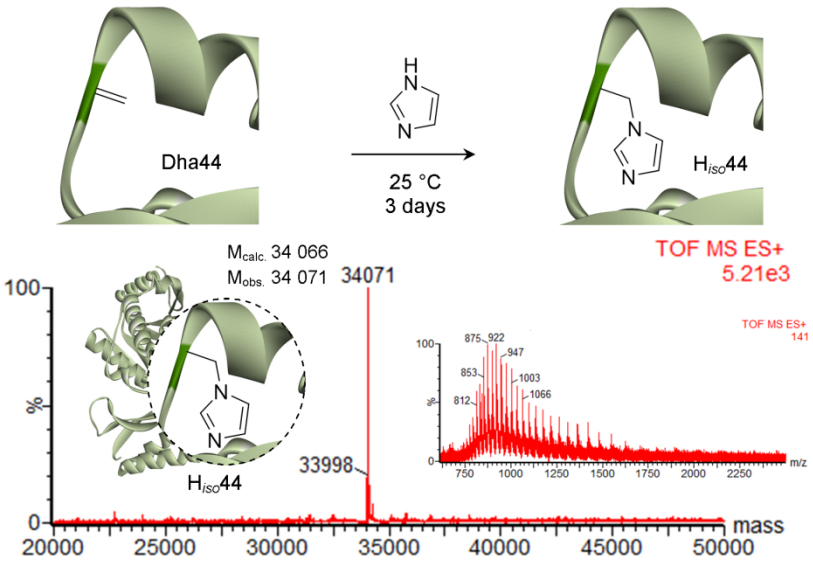


Figure S10. Aza-Michael addition of imidazole to PanC-H44Dha at 25 °C.

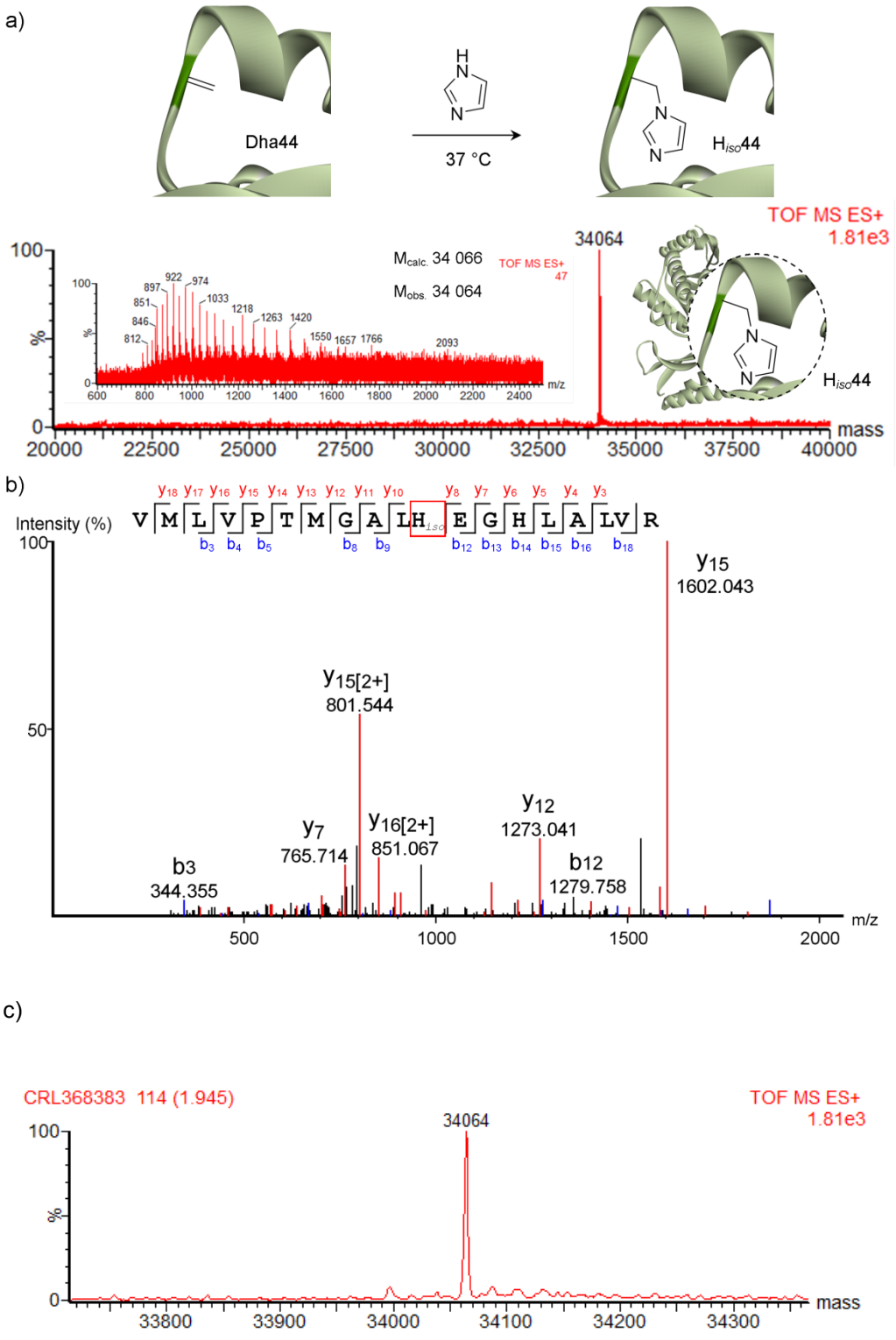


Figure S11. a) Aza-Michael addition of imidazole to PanC-H44Dha at 37 °C; b) MS/MS analysis of peptide containing H_{iso}44 (trypsin digest of PanC-H44H_{iso}); c) 'zoom' of a.

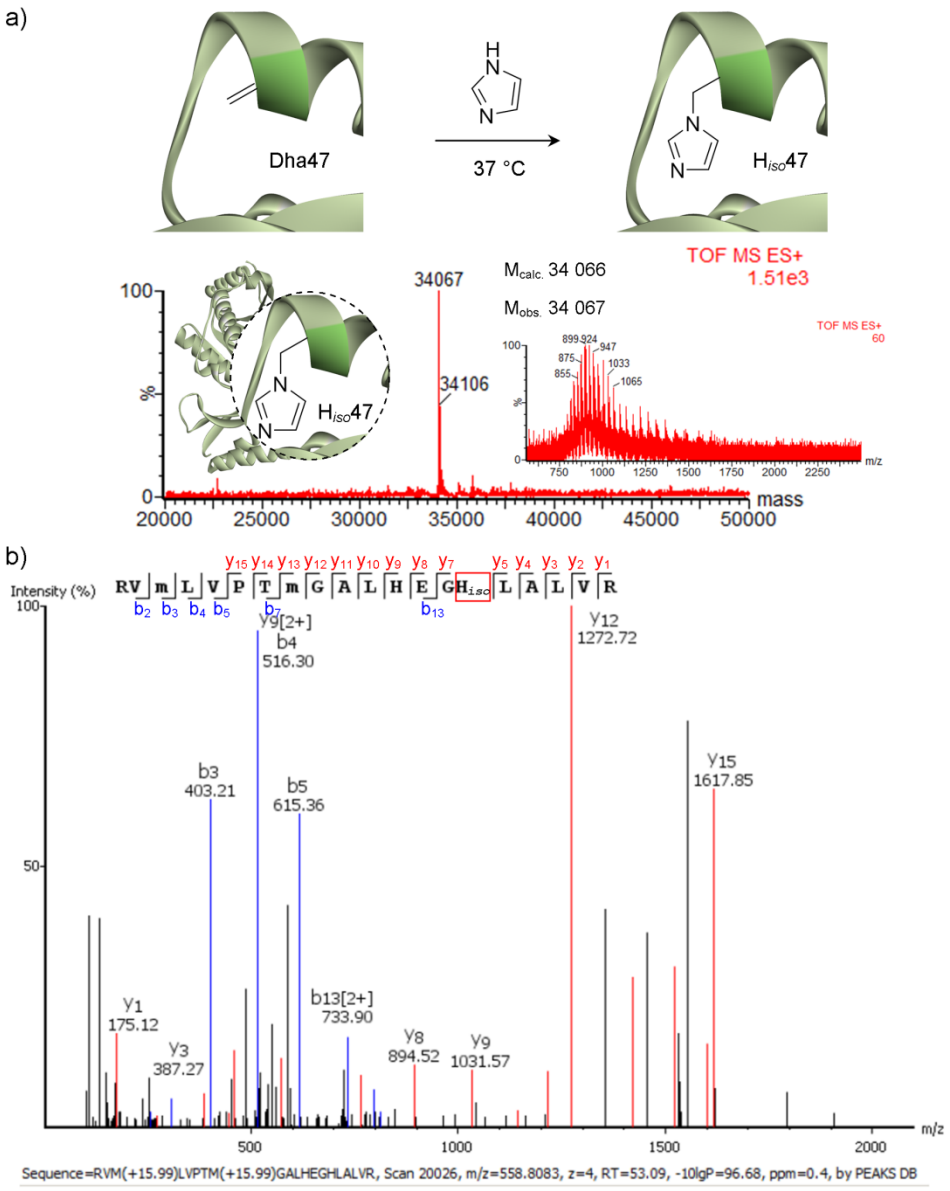


Figure S12. a) Aza-Michael addition of imidazole to PanC-H47Dha at 37 °C; b) MS/MS analysis of peptide containing $H_{iso}47$ (trypsin digest of PanC-H47H_{iso}).

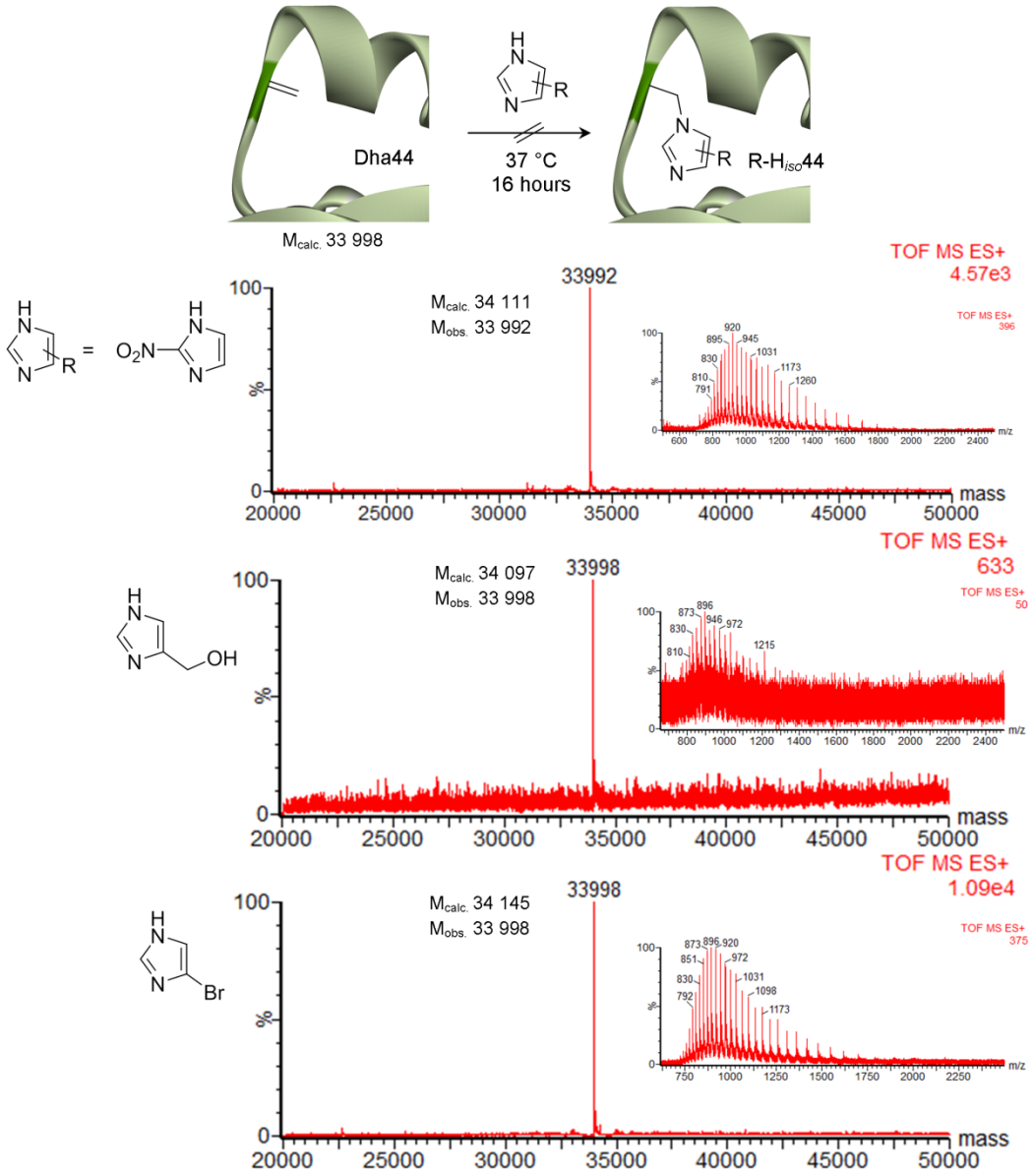


Figure S13. Aza-Michael addition of substituted imidazoles to PanC-H44Dha.

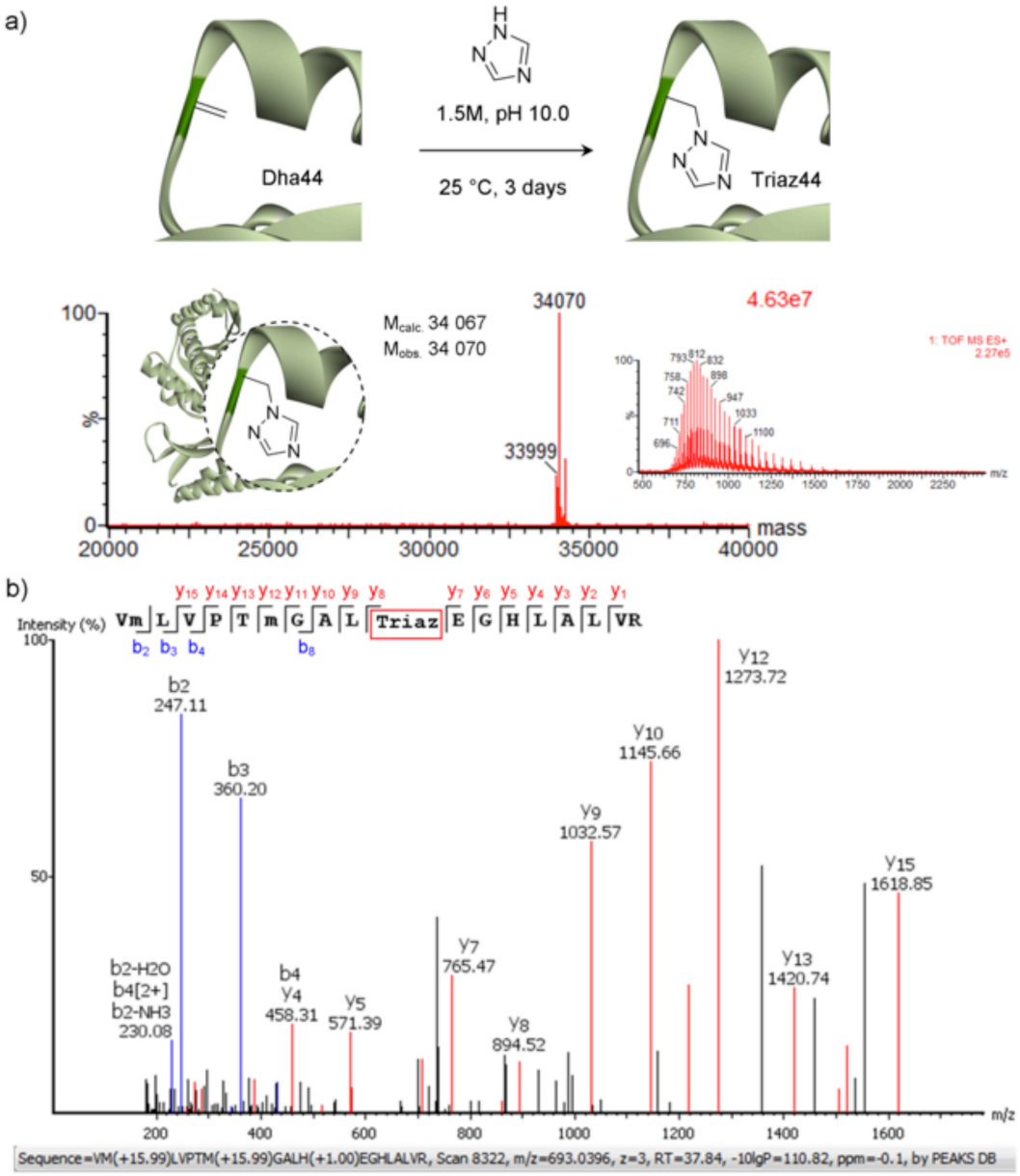


Figure S14. a) Aza-Michael addition of 1,2,4-triazole to PanC-H44Dha at 25 °C; b) MS/MS analysis of peptide containing azaHis$_{iso}$ = Triaz44 (tryptic digest of PanC-H44azaHis$_{iso}$ = Triaz).

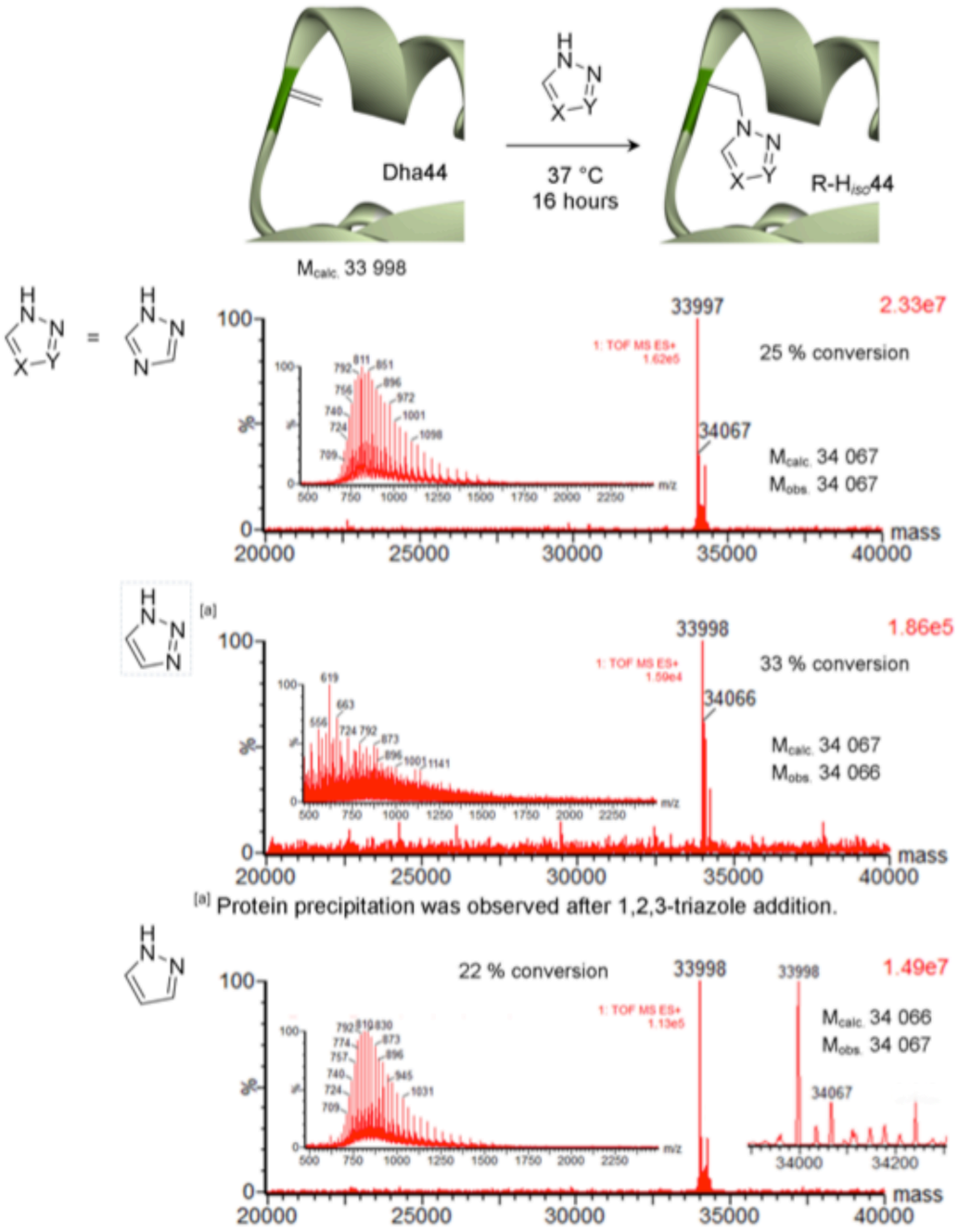


Figure S15. Aza-Michael addition of 1,2,4-triazole, 1,2,3-triazole and pyrazole to PanC-H44Dha at 37 °C, pH 8.

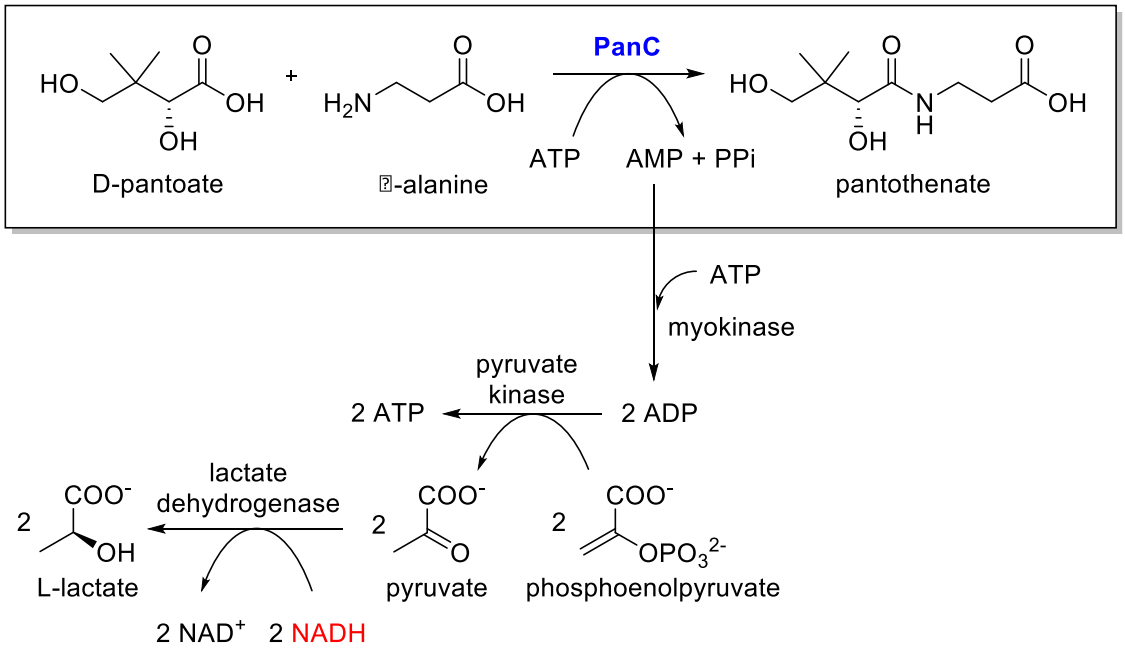


Figure S16. General scheme of NADH-coupled activity assay used for kinetic studies of PanC variants.

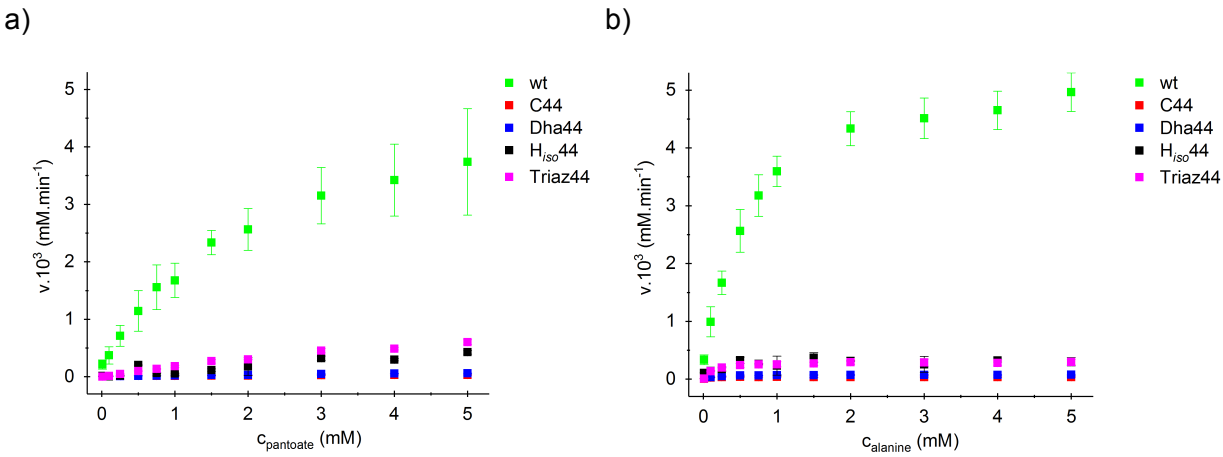
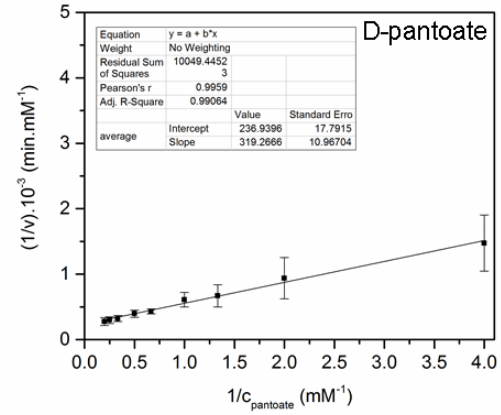
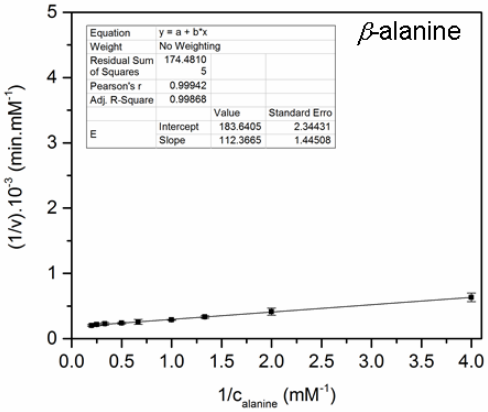


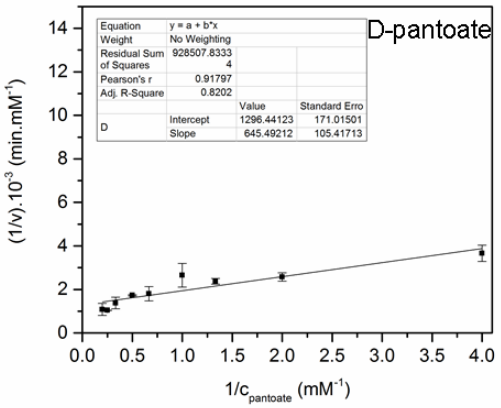
Figure S17. Michaelis-Menten kinetic of pantothenate formation catalysed by PanC variants for a) D-pantoate and b) *β*-alanine as a substrate.

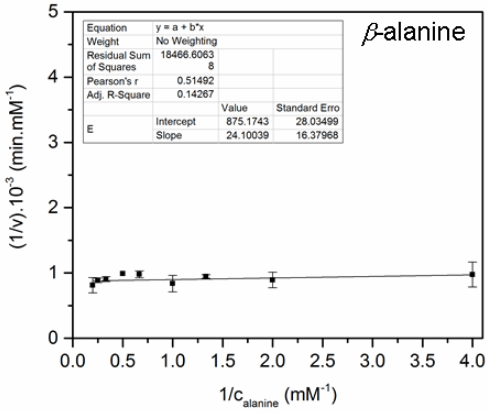
a) PanC-wt



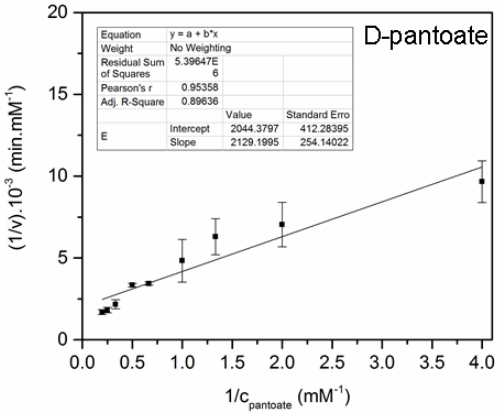


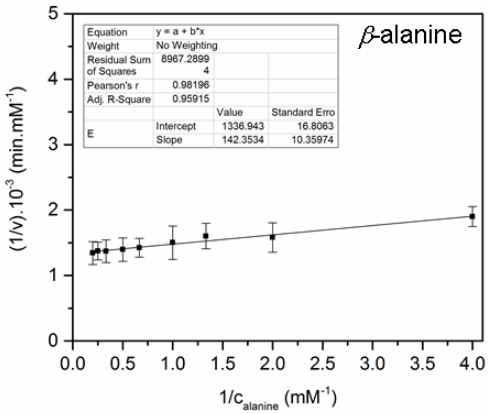
b) PanC-H44C



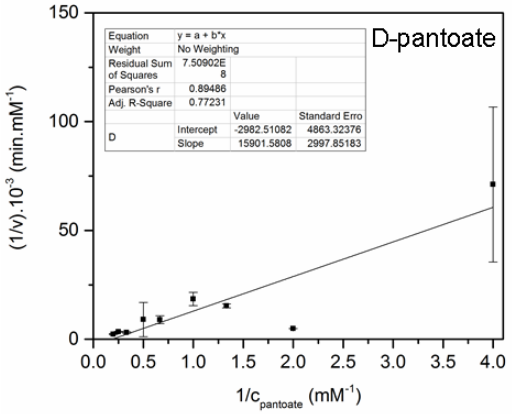


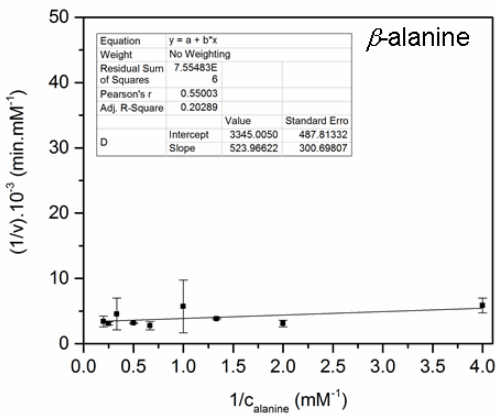
c) PanC-H44Dha





d) PanC-H44H$_{iso}$





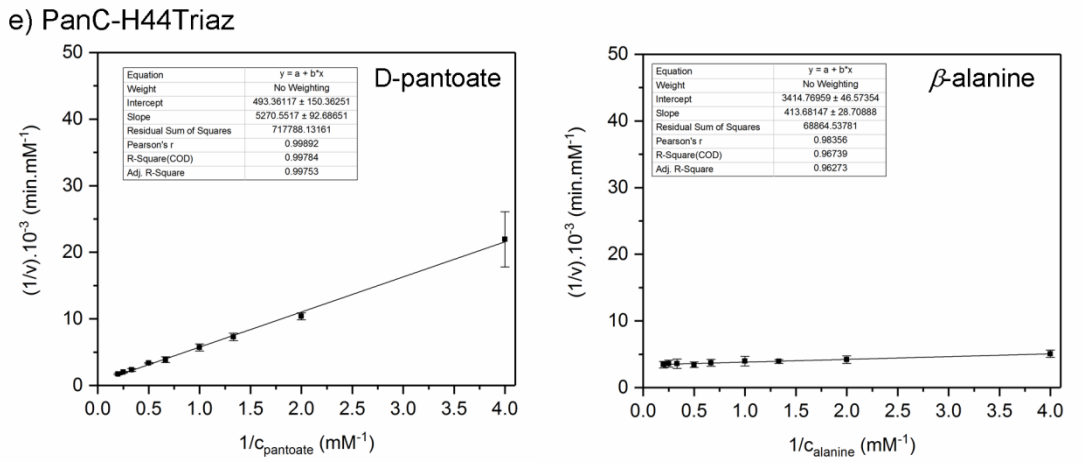


Figure S18. Lineweaver-Burk plots (1/*v* vs. 1/c) constructed from Michaelis-Menten kinetic data for D-pantoate and *β*-alanine. a) PanC-wt, b) PanC-H44C, c) PanC-H44Dha, d) PanC-H44H$_{iso}$, e) PanC-H44azaHis$_{iso}$ (=Triaz).

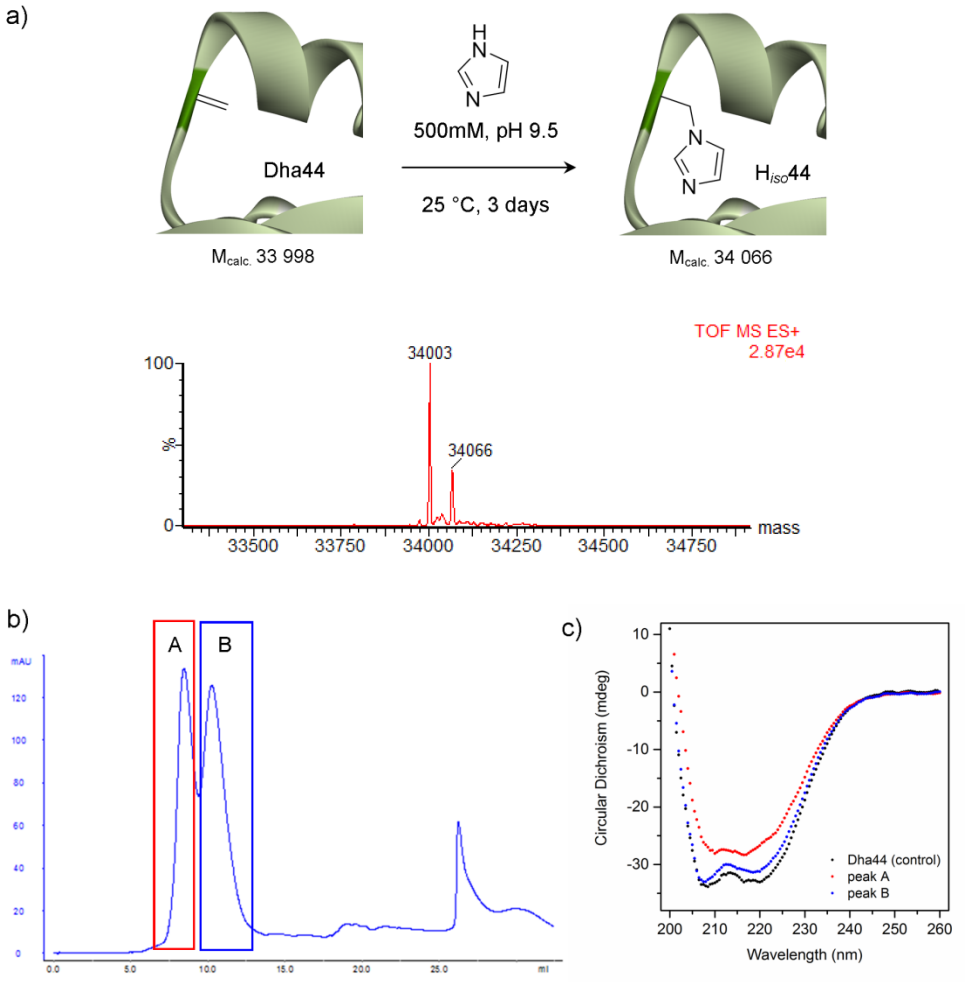


Figure S19. Preparation of PanC-H44H$_{iso}$ for enzyme activity assay. a) LC-MS analysis of the reaction mixture revealed ~15-20% conversion to PanC-H44H$_{iso}$. b) UV (280 nm) trace of the reaction mixture purification by gel filtration. c) CD spectra confirming correct folding of PanC that was used for activity assay (peak B).

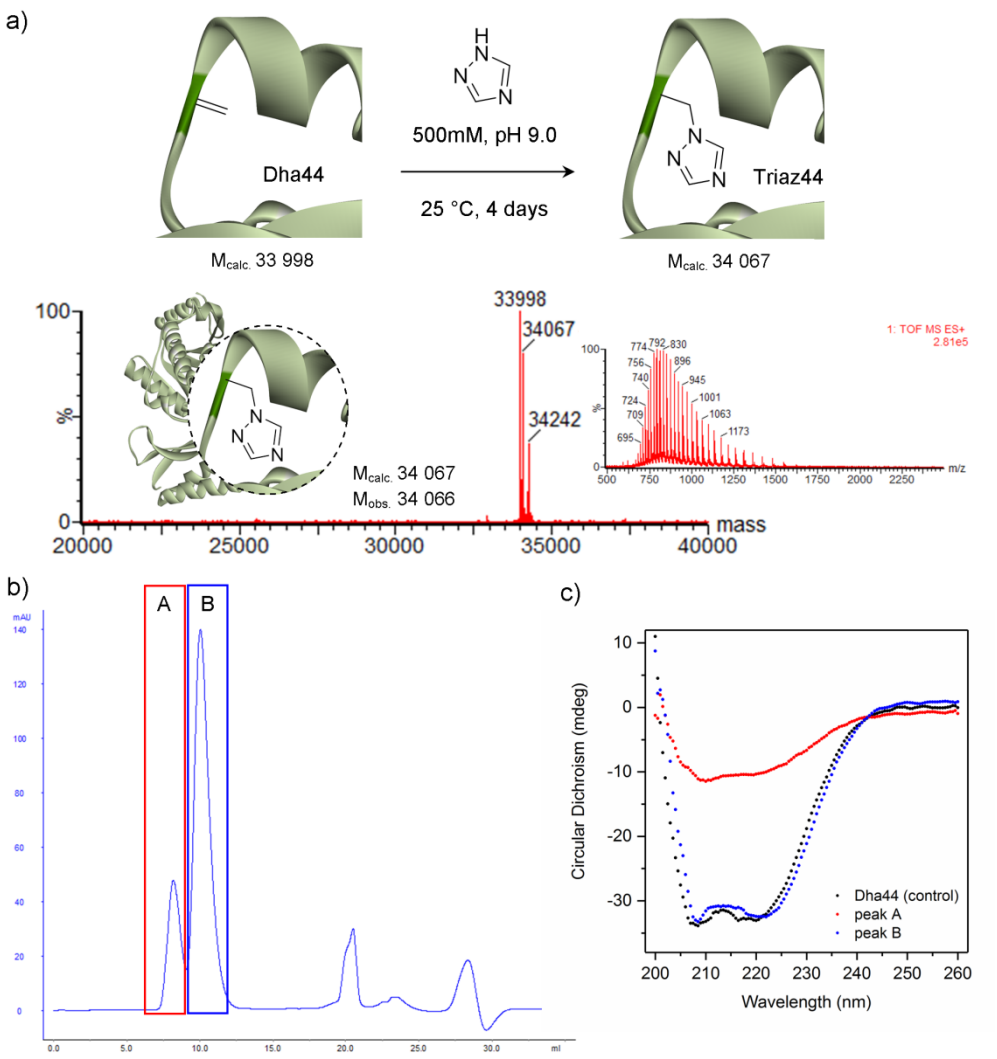


Figure S20. Preparation of PanC-H44azaHis$_{iso}$ (= H44Triaz) for enzyme activity assay. a) LC-MS analysis of the reaction mixture revealed 45% conversion to PanC-H44Triaz. b) UV (280 nm) trace of the reaction mixture purification by gel filtration. c) CD spectra confirming correct folding of PanC that was used for activity assay (peak B).

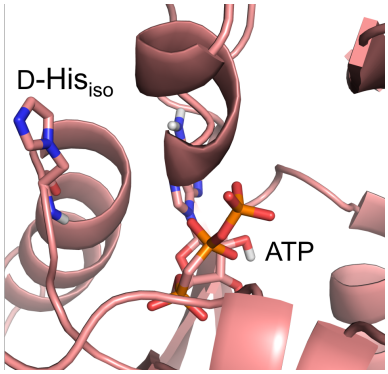


Figure S21 Minimized structure of PanC-44D-His$_{iso}$. PanC (PDB 1N2E) was virtually mutated with MOE2016.08 (Chemical Computing Group, Canada). Energy minimization was performed with the Amber10:EHT forcefield under default settings. Images were built with PyMOL (Schrödinger LLC, USA).

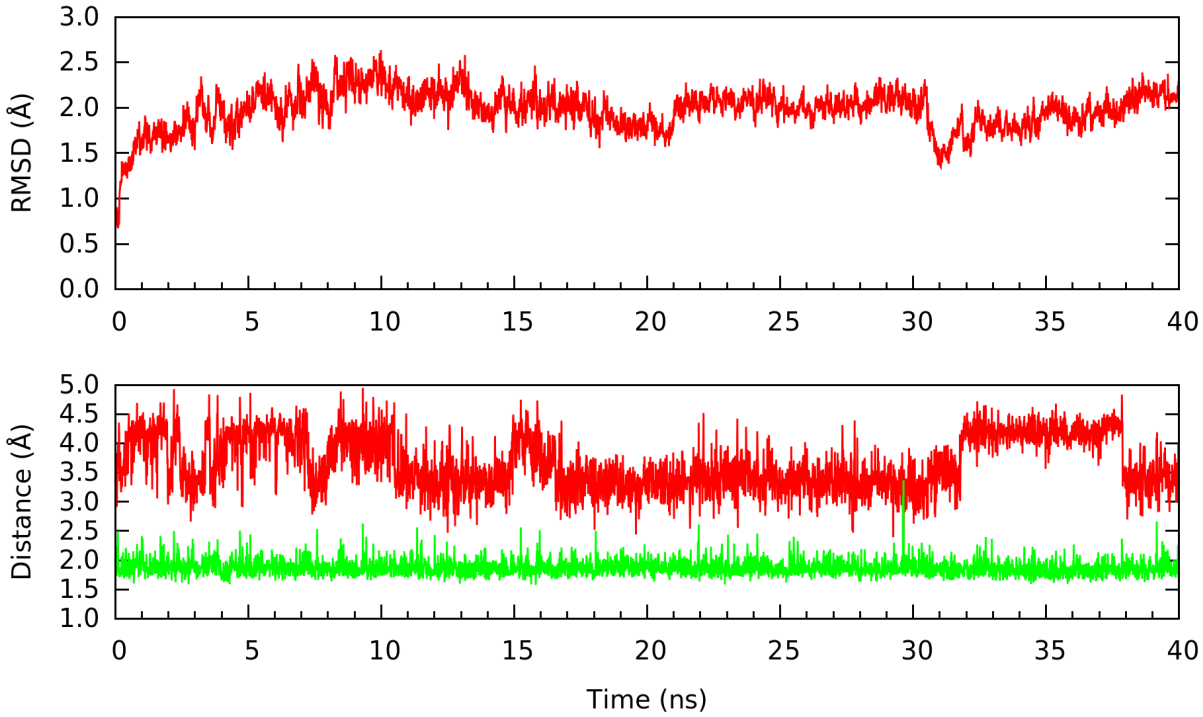


Figure S22. (Top panel) RMSD evolution during the MD simulation of the WT system. (Bottom panel) Hydrogen bond distances between the epsilon proton of His44 and either O1A (red) or O1B (green) atoms of the β-phosphate of ATP.

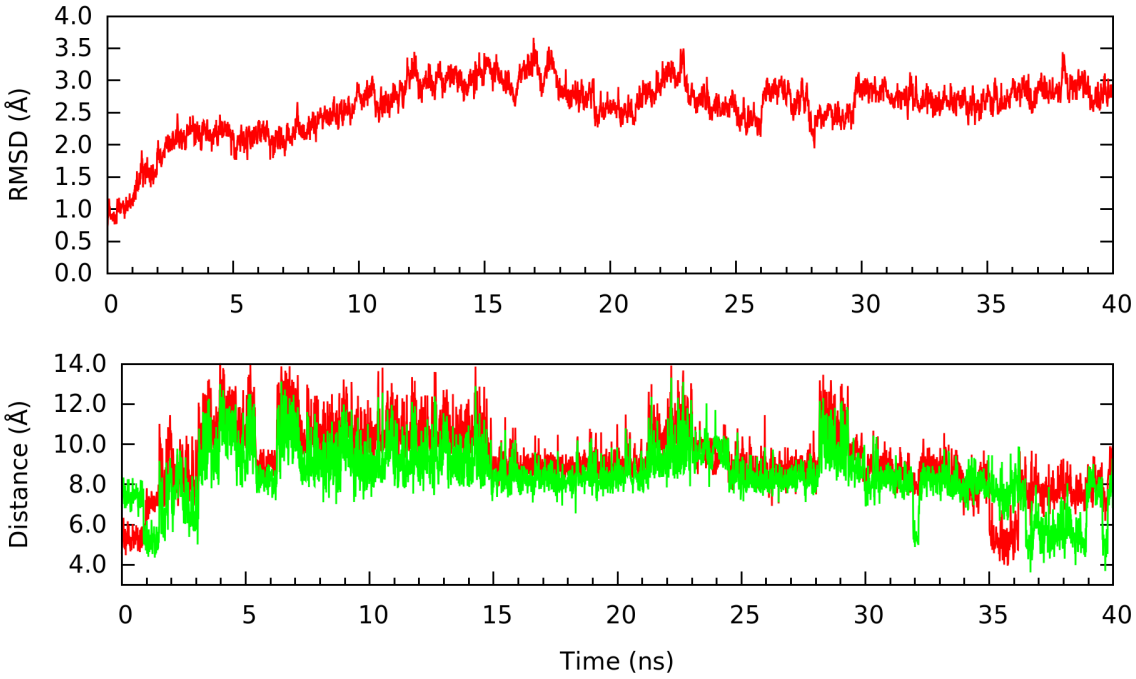


Figure S23. (Top panel) RMSD evolution during the MD simulation of the His44Cys mutant. (Bottom panel) Distance between the thiol proton of Cys44 and either the O1A (red) or O1B (green) atoms of the β-phosphate of ATP. Note the different scale of the distance axis in this case (no hydrogen bond) with respect to that of Figure S22 (hydrogen bond).

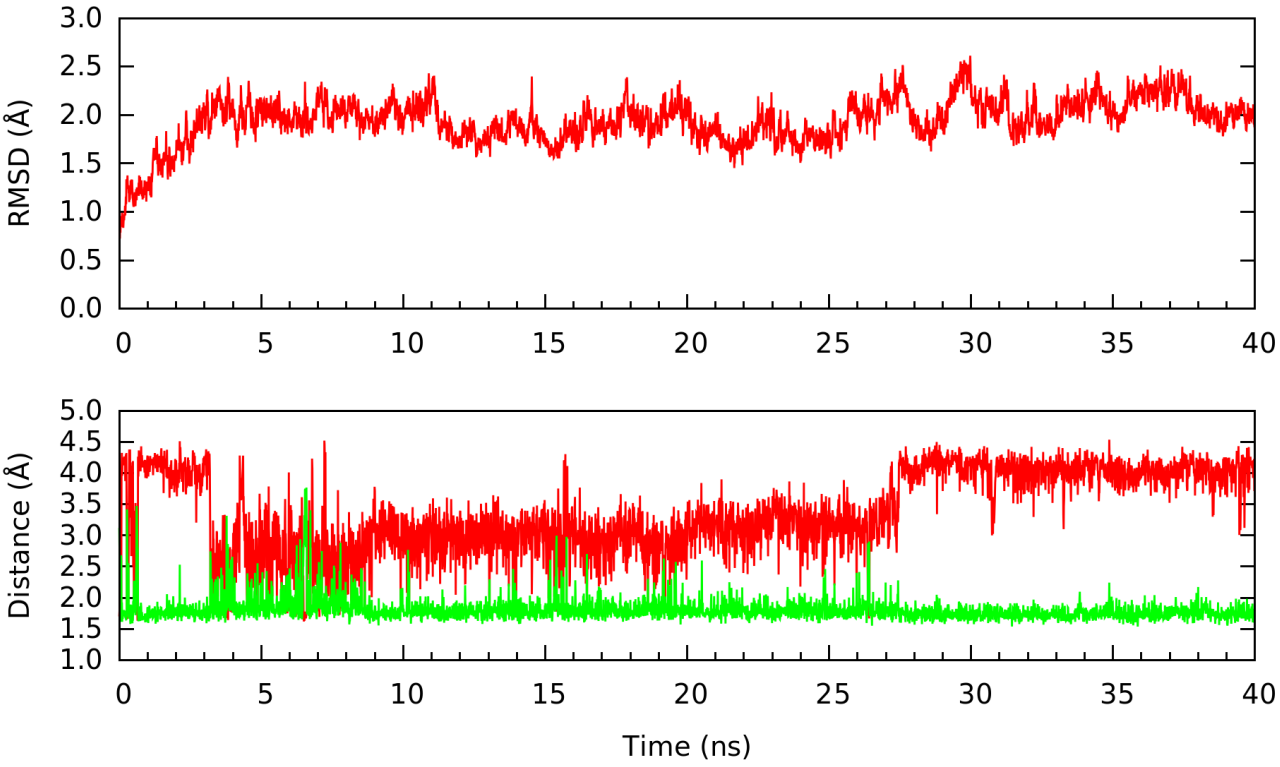


Figure S24. (Top panel) RMSD evolution during the MD simulation of the His44His$_{iso}$ mutant, considering a protonated iso-His44. (Bottom panel) Hydrogen bond distance between the proton of iso-His44 and either O1A (red) or O1B (green) atoms of the β-phosphate of ATP.

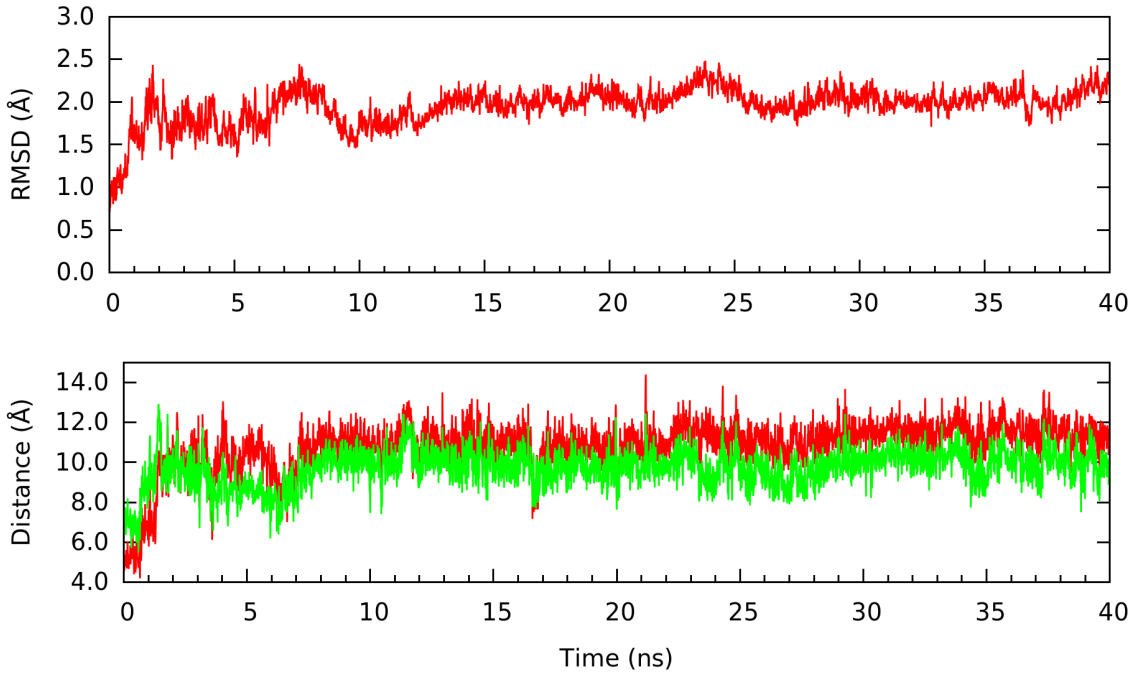


Figure S25. (Top panel) RMSD evolution during the MD simulation of the His44His$_{iso}$ mutant, considering a neutral iso-His44. (Bottom panel) Distance between the nitrogen atom of iso-His44 and either the O1A (red) and O1B (green) atoms of the β-phosphate of ATP. Note the different scale of the distance axis in this case (no hydrogen bond) with respect to that of Figure S24 (hydrogen bond).

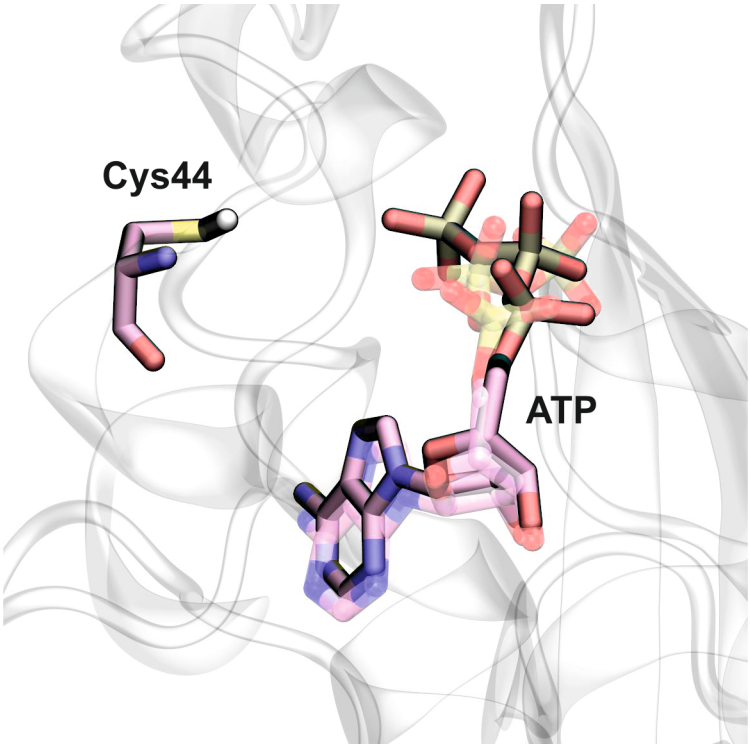


Figure S26. Conformational change of the ATP phosphate groups when there is no hydrogen bond interaction with the residue at the 44 position. The conformation of ATP before the conformational change is represented in transparent color.

2 Experimental

2.1 General remarks

Molecular graphics and analyses were performed with Discovery Studio Modeling Environment 4.5 (Dassault Systemes BIOVIA).

Protein expression and purification

All chemicals, primers, buffer and media reagents were purchased from common suppliers and used as delivered. Solutions were sterilized by autoclaving at 121 °C for 10–15 min. Solutions that could not be autoclaved were filtered through 0.2 μm or 0.45 μm Minisart® syringe filters (Sartorius AG). BL21(DE3)pLysS (Novagen) and XL10-Gold (Agilent) *E. coli* competent cells were handled as recommended by the manufacturer. BL21(DE3)pLysS were used for protein expression, XL10-Gold for cloning. Site-directed mutagenesis was performed using QuikChange XL Site-Directed Mutagenesis Kit (Agilent) following manufacturer's protocol. Plasmid DNA was isolated from overnight cell culture using QIAprep MiniPrep Kit (Qiagen). DNA and protein concentrations were determined using a NanoDrop 1000 spectrophotometer (Thermo Scientific) at 260 nm and 280 nm, respectively. The protein mass and extinction coefficient were calculated using ProtParam tool (ExPASy).[1] Plates were incubated at 37 °C in a Heraeus B6030 incubator. Shaking incubation of cell cultures was performed using either a New Brunswick Scientific Innova or a New Brunswick Scientific Model G25 incubator shaker. Microcentifugation was performed using an Eppendorf Centrifuge 5415R centrifuge. Low speed centrifugation of large volumes (>20 mL) was performed using a Beckman Coulter Allegra^{TM} X-12R centrifuge. High speed centrifugation of large volumes (>2mL) was performed using a Beckman Coulter Avanti^{TM} J-25 centrifuge. Fast protein liquid chromatography (FPLC) was done using an ÄKTA Purifier^{TM} system (Amersham Biosciences/GE Healthcare). Pre-packed FPLC Ni^{2+}-affinity columns (HisTrap HP 5 mL) were purchased from GE Healthcare. Dialysis was performed using Slide-A-lyser® 10 000 MWCO dialysis cassettes (Thermo Scientific). SDS-PAGE protein gel analysis was done using 4-12% NuPAGE® Bis-Tris pre-cast gels (Invitrogen) in NuPAGE® MES SDS running buffer (50mM MES, 50mM Tris, 0.1% SDS, 1mM EDTA, pH 7.3) at 200 V. The samples were prepared by mixing of protein sample with equal volume of SDS-PAGE loading buffer (10% SDS, 40% (v/v) glycerol, 200mM Tris pH 6.8, 0.04% bromophenol blue, 5% β-mercaptoethanol) and directly used for SDS-PAGE analysis. Gels

were stained with InstantBlue™ protein stain (C.B.S. Scientific) for 1 hour and destained with distilled water for 18 hours.

LC-MS analysis of intact proteins

LC-MS-TOF analysis of proteins was performed using following systems.

a) LCT Premier

Waters LCT Premier XE was coupled to a Waters 1525 Micro HPLC. RP18e (5x2mm) guard cartridge was run with eluents A: water/formic acid (100/0.1) and B: acetonitrile. The method was programmed as follows: initial 95% A (flow 0.3 mL/min), 95% A (0.5 min), 95% A to 40% A (1 min), 40% A (1.5 min, flow 0.3 mL/min and 1.5 min, flow 0.75 mL/min), 2% A (0.5 min, flow 0.75 mL/min). The electrospray source of LCT was operated with a capillary voltage of 3 kV and a cone voltage of 100 V. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 300 L hr-1. The protein mixture was diluted to 0.01-0.02 mg/mL with MilliQ water, 10 µL were injected.

b) LCT Classic

Waters Micromass LCT (ESI-TOF-MS) instrument was coupled to a Shimadzu HPLC system composed a Shimadzu LC-20AD SP, a DGU-20A5 degasser, a SIL-20Ac autosampler-cooler, SCL-10A VP system controller. Thermo Scientific, Proswift™ RP-2H (4.6x50 mm SS) column was run with eluents A: water/acetonitrile/formic acid (95:5:0.1), and B: acetonitrile/water/formic acid (95:5:0.1). The method was programmed as follows: 100% A (1 min, flow 0 to 0.4 mL/min), 100% A to 100% B (6 min), 100% B (4 min), 100% B to 100% A (1 min) and 100% A (5 min). The electrospray source of LCT was operated with a capillary voltage of 3.2 kV and a cone voltage of 25 V. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 L hr^{-1}. The protein mixture was diluted to 0.03-0.05 mg/mL with MilliQ water, 8-10 µL were injected.

Intact protein LC-MS data were analysed using MassLynx version 4.1 (Waters). For LCT Premier, the ion series was autocalibrated. For LCT Classic, the ion series were calibrated using a calibration curve constructed from the ion series of equine myoglobin, obtained with the same LC-MS settings. Total mass spectra were reconstructed from the ion series using the preinstalled MaxEnt algorithm. The deconvolution was set up as follows: output mass, ranges 20000:50000, resolution 1 Da/channel; damage model, uniform Gaussian width at half height

0.75 Da, minimum intensity ratios left 33% and right 33% and completion options, iterate to convergence.

LC-MS/MS analysis of proteins

Protein sample was digested with trypsin in solution and analysed on an Orbitrap Elite (Thermo Fisher Scientific) with the following instrument parameters: ESI nano-spray ion source; CID, CAD (y and b ions) fragmentation mode; FT-ICR/Orbitrap MS scan mode; linear ion trap MS/MS scan mode. Raw MS data were processed using PEAKS Studio software v8.0 (Bioinformatics Solutions Inc.) with the following search parameters: parent mass error tolerance 10.0 ppm; fragment mass error tolerance 0.05 Da; monoisotopic precursor mass search type, enzyme used trypsin; precursor options corrected; 1 missed cleavage; both non-specific cleavage; variable modification methionine oxidation (+15.99), deamidation (+0.98), histidine to dehydroalanine mutation (-68.04), histidine to triazole mutation (+1.00).

NADH-coupled enzyme activity assay

Absorbance at 340 nm was monitored using SPECTROstar^{Nano} microplate reader (BMG Labtech) in a 96-well F-bottom plate (Greiner Bio-One Cellstar). The length of the light path was determined (l = 0.325 cm) in a fixed volume (200 μl) using Lambert-Beers law equation and NADH extinction coefficient (6220 $cm^{-1}M^{-1}$ at 340 nm). Stock solution of D-pantoate (100mM) was prepared by hydrolysis of (*R*)-(–)-pantolactone in NaOH (150mM) and used directly for enzyme assay. Kinetic data analysis was done using MS Excel and OriginPro 2015 software.

2.2 Expression and purification of PanC variants

Plasmid preparation

The pET23b(+) plasmid with panC-wt gene (Appendix 5) was obtained from Prof. John S. Blanchard, Albert Einstein College of Medicine, Bronx, USA. Site-directed mutagenesis was performed using following primers according to manufacturer's protocol. Plasmid DNA was isolated from overnight cell culture using QIAprep MiniPrep Kit (Qiagen). The correct DNA sequence was confirmed by Sanger sequencing with T7F and T7R primers.

PanC-H44C

forward: 5'-CTATGGGTGCGCTGTGCGAAGGCCACCTCGC-3'
reverse: 5'-GCGAGGTGGCCTTCGCACAGCGCACCCATAG-3'

PanC-H47C

forward: 5'-GCGCTGCACGAAGGCTGCCTCGCGTTGGTGCGTG-3'
reverse: 5'-CACGCACCAACGCGAGGCAGCCTTCGTGCAGCGC-3'

Protein expression and purification

The plasmid (10 ng/μL, 1 μL) was added to BL21(DE3)pLysS competent cells (10 μL) and incubated on ice for 45 minutes. The tube was then heated at 42 ^{o}C in water bath for 45 seconds and immediately placed on ice for 2 minutes. SOC medium (90 μL, preheated to 42 ^{o}C) was added and the samples were incubated at 37 ^{o}C and 300 rpm for 1 hour. Transformed cells (100 μL) were plated onto LB agar plate with ampicillin (50 μg/mL) and chloramphenicol (34 μg/mL). The plates were incubated at 37 ^{o}C overnight. A single cell colony was picked to LB medium (4 mL) with ampicillin (50 μg/mL) and chloramphenicol (34 μg/mL) and shaken at 160 rpm and 37^{o}C overnight. LB medium (1 L) containing ampicillin (50 μg/mL) and chloramphenicol (34 μg/mL) was inoculated with overnight culture (4 mL) and incubated at 37 °C and 160 rpm until OD_{600} 0.5 was reached (≈ 3 hours). IPTG (1M, 100 μL) was added and PanC protein variant was expressed at 18 °C and 160 rpm overnight. Cells were harvested by centrifugation (4 °C, 8 000 × g, 10 min), flash-frozen with liquid nitrogen and stored at -80 °C until further protein purification.

Cell pellet (4 g) was re-suspended in buffer A (15 mL) containing cOmplete^{TM} protease inhibitor mix (EDTA free, Roche). Lysozyme (3 mg) and DNase I (1.5 mg) were added and the suspension was stirred vigorously at 4 °C for 2 hours. The lysate was centrifuged at 4 °C and 22 000 × g for 30 min. Supernatant was filtered through 0.2μm syringe filter and loaded

to HisTrap HP 5mL column (GE Healthcare, 3 mL/min, washed with 10 CV water and equilibrated with 10 CV buffer A) using 50mL superloop (GE Healthcare). Unbound proteins were washed from the column with buffer A (10 CV). Pantothenate synthetase was eluted by linear gradient of buffer B (0-100 % in 25 CV). Fractions were analysed by 4-12 % SDS PAGE with 1 × MES running buffer. Fractions containing protein were combined and dialyzed using dialysis cassette (30 mL, 10k MWCO, Pierce) against PanC storage buffer (2 × 4 L; 50mM Tris-HCl, 200mM NaCl, 1mM DTT, 0.05% (v/v) β-mercaptoethanol, pH 7.8) at 4 °C (1 × 2 hours, then overnight). Structure of PanC-wt and PanC-H44C were confirmed by LC/MS analysis. Glycerol was added to protein stock solutions to reach final concentration of 5 %, aliquots (100 μL) were flash-frozen in liquid nitrogen and stored at -80 °C.

Buffer A - 20mM Tris-HCl, 15mM imidazole, 500mM NaCl, 1mM DTT, 0.05% (v/v) β-mercaptoethanol, pH 7.8

Buffer B - 20mM Tris-HCl, 500mM imidazole, 500mM NaCl, 1mM DTT, 0.05% (v/v) β-mercaptoethanol, pH 7.8

Yields:

PanC-wt –174 mg per 1 L of cell culture

PanC-H44C –155 mg per 1 L of cell culture

PanC-H47C –64 mg per 1 L of cell culture

2.3 Chemical modifications of proteins

Chemical mutation of SBL-S156C to SBL-S156H$_{iso}$

Aza-Michael addition of imidazole to SBL-S156Dha

SBL-S156Dha was prepared using procedure adapted from the literature.[2] A 200 μl aliquot of SBL-S156Dha (0.7 mg/mL) was prepared in sodium phosphate buffer (pH 8.0). Imidazole (3.6 mg, 0.052 mmol) was added to the protein solution as a solid. The reaction was incubated at 37 °C for 5 hours. Complete conversion to SBL-S156H$_{iso}$ was confirmed by LC-MS ($M_{calc.}$ = 26 749; M_{found} = 26 749).

Chemical mutation of PanC in positions 44 and 47

PanC-H44C and PanC-H47C were buffer-exchanged from storage buffer to sodium phosphate buffer (50mM, pH 8.0) using PD MiniTrap^{TM} G-25 column (GE Healthcare) equilibrated with sodium phosphate buffer following gravity protocol recommended by the manufacturer.

Two-step formation of PanC-H44Dha and PanC-H47Dha

Freshly prepared stock solution of methyl 2,5-dibromopentanoate (MDBP; 1M in DMSO) was added (50 equiv. to protein) to PanC cysteine mutant (3.5 mg/mL) in sodium phosphate buffer (50mM, pH 8.0). The reaction mixture was shaken at 25 °C and 500 rpm. Complete alkylation of cysteine was confirmed by LC/MS (after 45 minutes for Cys44, 16 hours for Cys47). Excess of MDBP was immediately removed by gel filtration using PD SpinTrap^{TM} G-25 column (GE Healthcare) equilibrated with sodium phosphate buffer (50mM, pH 8.0). The resulting solutions of alkylated PanC were further incubated at 25 °C and 500 rpm yielding PanC modified with dehydroalanine (16 hours for PanC-H44Dha, $M_{calc.}$ = 33 998, M_{found} = 33 996; 30 hours for PanC-H47Dha, $M_{calc.}$ = 33 998, M_{found} = 33 999).

Control thia-Michael addition of β-mercaptoethanol to PanC-H44Dha

Freshly prepared solution of *β*-mercaptoethanol (2 μL, 10% v/v in 50mM phosphate buffer pH 8.0) was added to PanC-H44Dha (1 mg/mL) in phosphate buffer (50mM, pH 8.0). The reaction mixture was shaken at 37 °C and 500 rpm for 3 hours and then analysed by LC/MS ($M_{calc.}$ = 34 076; M_{found} = 34 076).

Aza-Michael addition of imidazole to PanC-H44Dha or PanC-H47Dha - formation of PanC-H44H$_{iso}$ and *PanC-H47H$_{iso}$*

Method A: Imidazole (102 mg, 1.5 mmol) was dissolved in PanC-H44Dha or PanC-H47Dha (1.5 mg/mL, 1 mL) solution in sodium phosphate buffer (50mM, pH 8.0). The reaction mixtures were shaken at 37 °C, 500 rpm overnight. Full conversions to PanC-H44H$_{iso}$ and PanC-H47H$_{iso}$ were confirmed by LC/MS.

Method B: Imidazole (3M, pH 9.5, 450 μL) was added to PanC-H44Dha or PanC-H47Dha (3.0 mg/mL; 450 μL) solution in sodium phosphate buffer (50mM, pH 8.0). The reaction mixtures were shaken at 25 °C, 500 rpm for 3 days and analysed by LC/MS.

Preparation of PanC-H44H$_{iso}$ for kinetic assay

Imidazole (1M, pH 9.5, 450 μL) was added to PanC-H44Dha (3.0 mg/mL; 450 μL) solution in sodium phosphate buffer (50mM, pH 8.0). The reaction mixture was shaken at 25 °C, 500 rpm for 3 days and analysed by LC/MS revealing 15% conversion to PanC-H44H$_{iso}$. The solution was centrifuged at 15 000 × g, 4 °C for 10 minutes prior gel filtration to remove undesired PanC aggregates. The purification was run using HiLoad Superdex 75 10/300 column (GE Healthcare) pre-equilibrated with gel filtration buffer (50mM Tris pH 8.0, 100mM NaCl, 5% glycerol, 0.2mM EDTA, 0.2mM DTT) at 4 °C using FPLC (for UV trace see Figure S19b). Structure of purified protein was studied by CD spectroscopy (Figure S19c). Fractions containing correctly folded PanC (peak B, PanC-H44H$_{iso}$ / PanC-H44Dha 15:85) were used for NADH coupled spectrophotometric assay. The kinetic parameters were measured against PanC-H44Dha background.

Aza-Michael addition of 1,2,4-triazole to PanC-H44Dha - formation of PanC-H44Triaz

1,2,4-Triazole (3M, pH 10.0, 500 μL) was added to PanC-H44Dha (3.4 mg/mL; 500 μL) solution in sodium phosphate buffer (50mM, pH 8.0). The reaction mixture was shaken at 25 °C, 500 rpm for 3 days and analysed by LC/MS.

Preparation of PanC-H44Triaz for enzyme kinetic assay

1,2,4-Triazole (1M, pH 9.5, 500 μL) was added to PanC-H44Dha (3.4 mg/mL; 500 μL) solution in sodium phosphate buffer (50mM, pH 8.0). The reaction mixture was shaken at 25 °C, 500 rpm for 4 days and analysed by LC/MS revealing 45% conversion to PanC-H44Triaz. The solution was centrifuged at 15 000 × g, 4 °C for 10 minutes prior gel

filtration to remove undesired PanC aggregates. The purification was run using HiLoad Superdex 75 10/300 column (GE Healthcare) pre-equilibrated with gel filtration buffer (50mM Tris pH 8.0, 100mM NaCl, 5% glycerol, 0.2mM EDTA, 0.2mM DTT) at 4 °C using FPLC (for UV trace see Figure S20b). Structure of purified protein was studied by CD spectroscopy (Figure S20c). Fractions containing correctly folded PanC (peak B, PanC-H44Triaz / PanC-H44Dha 45:55) were used for NADH coupled spectrophotometric assay. The kinetic parameters were measured against PanC-H44Dha background.

Attempted Aza-Michael addition of substituted imidazoles to PanC-H44Dha

4-Hydroxymethyl-imidazole (9.6 mg, 71 μmol), 4-bromoimidazole (8.1 mg, 55 μmol) or 2-nitroimidazole (11 mg, 97 μmol) were added to PanC-H44Dha (4.5 mg/mL; 50 μL) in sodium phosphate buffer (50mM, pH 8.0). The reaction mixtures were shaken at 37 °C and 500 rpm overnight and analysed by LC-MS.

Attepted Aza-Michael addition of pyrazole, 1,2,4- and 1,2,3-triazole to PanC-H44Dha

Pyrazole (27 mg, 0.39 mmol), 1,2,4-triazole (30 mg, 0.43 mmol) or 1,2,3-triazole (27 mg, 0.39 mmol) were added to PanC-H44Dha (1 mg/mL; 100 mL) in sodium phosphate buffer (50mM, pH 8.0). The reaction mixtures were shaken at 37 °C and 500 rpm overnight and analysed by LC-MS.

Protein characterization

CD spectroscopy

CD spectra were measured on a Chirascan device at 20 °C, 25 °C or 37 °C with a water bath to control the temperature. Experiments were performed in 200 µL of sample at protein concentration of 0.25 mg/mL in 1mm cuvette in sodium phosphate buffer (50mM, pH 8.0). Far-UV spectra were collected from 200 to 260 nm (0.5 nm step, 1 s per point) as an average from three scans and smoothed. Buffer solution was used as a blank.

LC-MS/MS analysis

DTT (2 µL, 200mM in 50mM NH_4HCO_3) was added to protein (15 µg) solution in 8M urea and 50mM NH_4HCO_3 (100 µL). The mixture was incubated at 56 °C for 25 min. After diluting with NH_4HCO_3 (325 µL, 50mM), protein was digested by trypsin (1.5 µL, 0.2 µg/µL, 1:50 w/w to protein) at 37 °C and 500 rpm overnight. The digestion was stopped by addition of 0.2% v/v formic acid (430 µL). Sample concentration was adjusted with MilliQ water to 400 fmol/µL. LC-MS/MS analysis and data processing were conducted as described in the general remarks section.

SBL Kinetic Parameter Determination

Michaelis-Menten constants were measured at 25 ˚C by non-linear correlation of initial rate data (Graphpad Prism) determined at eight or nine concentrations (0.125-5mM) of succinyl-AAPF-*p*NA substrate in 0.1 M Tris•HCl buffer containing 0.005% Tween 80, 1% DMSO, pH 8.6 using UV-vis spectroscopy (ε_{410} = 8880 $M^{-1}cm^{-1}$).

PanC Kinetic Parameter Determination

NADH-coupled spectrophotometric assay[3]

The reaction mixture (200 µl) containing 100mM Hepes (pH 7.8), 10mM $MgCl_2$, 10mM ATP, 5mM D-pantoate, 5mM β-alanine, 1mM potassium phosphoenolpyruvate, 200µM NADH, myokinase, pyruvate kinase and lactate dehydrogenase (18 U/mL each) was prepared to 96-well plate, shaken for 10 s and incubated at 25 °C for 10 minutes. Pantothenate formation was initiated by addition of PanC (≤ 10 µL, 0.35 µg PanC-wt, 9.9 µg PanC-H44C, 3.4 µg PanC-H44Dha, 0.45 µg PanC-H44H_{iso}, 2.7 µg PanC-H44Triaz). The linear decrease

of absorbance at 340 nm was measured at 25 °C every 60 s (10 s shaking (400 rpm) before each measurement). K_m for both D-pantoate and β-alanine were measured at concentration of one substrate varying from 0.01 to 5mM whilst other saturated and constant (5mM).

Data analysis

Reaction rate v of pantothenate synthetase is half of the NADH consumption rate and can be calculated from equation (1).

$$\frac{\Delta A}{\Delta t} = \varepsilon l \frac{\Delta[NADH]}{\Delta t} = 2\varepsilon l v \quad (1)$$

where A is absorbance at 340 nm, ε is NADH extinction coefficient (6220 $cm^{-1}M^{-1}$), c is concentration, l is path length (0.325 cm), and $\Delta A/\Delta t$ is a slope in absorbance vs. time plot for each substrate concentration.

Double reciprocal (Lineweaver-Burk) plots ($1/v$ vs. $1/c$) were constructed for D-pantoate and β-alanine and fitted to a linear function (Figure S18). K_m and V_{max} were calculated as follows

$$\frac{1}{v} = \frac{K_m}{V_{max}} . \frac{1}{c} + \frac{1}{V_{max}} \quad (2)$$

2.4 Structural Analyses and Computational Methods

The initial structure for the simulations was taken from the ternary complex of PanC with D-pantoate and AMPCPP (PDB 1N2E). The AMPCPP was manually reverted to ATP by replacing the carbon atom by oxygen. Four models were considered: (i) the wild type enzyme, (ii) the His44Cys mutant, and the $His44His_{iso}$ mutant, in its protonated (iii) and deprotonated (iv) forms. The first two models models were taken as reference of active and inactive systems, respectively.

All the mutants were generated with AMBER. Protonation states were selected according to protein environment, with all glutamates and aspartates deprotonated and neutral histidine residues. All histidines were considered as N_ε-protonated, except for His135, which was considered as N_δ-protonated in view of its interaction with the backbone carbonyl oxygen of Gln72. The D-pantoate substrate was taken as protonated. The total model, including eight sodium counterions (seven in the case the charged $His45His_{iso}$ variant is composed of 64,943 atoms (WT), including 20,190 solvent water molecules.

Classical molecular dynamics simulations were performed to equilibrate the ternary complex using the Amber11 software.[4] The FF99SB and TIP3P force-fields were selected to describe the protein and water solvent, respectively.[5,6] Parm99 parameters were used to describe the D-pantoate molecule and the H_{iso} residue (their charges were derived according to Amber protocols). ATP parameters were taken from Richard Bryce AMBER database (Manchester).[7,8]

MD simulations were carried out for each system in several steps. First, the system was minimized, holding the protein and D-pantoate substrate fixed. Then, the entire system (protein + substrate + water solvent + counterions) was allowed to relax dynamically. To gradually reach the desired temperature of 300 K in the MD simulations, weak spatial constraints were initially added to the protein and substrate, while the water molecules and sodium ions were allowed to move freely at 100 K. The constraints were then removed and the working temperature of 300K was reached after two more 100 K heatings in the NVT ensemble. Afterwards, the density was converged up to water density at 300 K in the NPT ensemble and the simulation was extended to 40 ns in the NVT ensemble when the system reached equilibrium according to the root mean squared deviation of enzyme backbone (Figures S19 to S22). Analysis of the trajectories was carried out using standard tools of AMBER and VMD.[9]

3 Appendices

Appendix 1 - ***Sequence of SBL-S156C***

(BPN' numbering, PDB code for wild-type SBL = 1GCI, ref.[10])

AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNGHGTHVAG
TIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEWAGNNGMHVANLSLGSPSPSATLEQA
VNSATSRGVLVVAASGN**C**GAGSISYPARYANAMAVGATDQNNNRASFSQYGAGLDIVAPGVNVQSTYP
GSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

$M_{calc.}$ = 26 715, ε = 26 930 $M^{-1}cm^{-1}$

Appendix 2 - ***Sequence of PanC-wt*** (C-terminal HisX6 tag)

PDB code 1N2E, ref.[11]

TIPAFHPGELNVYSAPGDVADVSRALRLTGRRVMLVPTMGALHEGHLALVRAAKRVPGSVVVVSIFVNPMQFGAG
EDLDAYPRTPDDDLAQLRAEGVEIAFTPTTAAMYPDGLRTTVQPGPLAAELEGGPRPTHFAGVLTVVLKLLQIVR
PDRVFFGEKDYQQLVLIRQLVADFNLDVAVVGVPTVREADGLAMSSRNRYLDPAQRAAAVALSAALTAAAHAATA
GAQAALDAARAVLDAAPGVAVDYLELRDIGLGPMPLNGSGRLLVAARLGTTRLLDNIAIEIGTFAGTDRPDGYRA
ILESHWRNKLAAALEHHHHHH

$M_{calc.}$ = 34 066, ε = 15 930 $M^{-1}cm^{-1}$

Appendix 3 - ***Sequence of PanC-H44C*** (C-terminal HisX6 tag)

TIPAFHPGELNVYSAPGDVADVSRALRLTGRRVMLVPTMGAL**C**EGHLALVRAAKRVPGSVVVVSIFVNPMQFGAG
EDLDAYPRTPDDDLAQLRAEGVEIAFTPTTAAMYPDGLRTTVQPGPLAAELEGGPRPTHFAGVLTVVLKLLQIVR
PDRVFFGEKDYQQLVLIRQLVADFNLDVAVVGVPTVREADGLAMSSRNRYLDPAQRAAAVALSAALTAAAHAATA
GAQAALDAARAVLDAAPGVAVDYLELRDIGLGPMPLNGSGRLLVAARLGTTRLLDNIAIEIGTFAGTDRPDGYRA
ILESHWRNKLAAALEHHHHHH

$M_{calc.}$ = 34 032, ε = 15 930 $M^{-1}cm^{-1}$

Appendix 4 - ***Sequence of PanC-H47C*** (C-terminal HisX6 tag)

TIPAFHPGELNVYSAPGDVADVSRALRLTGRRVMLVPTMGALHEG**C**LALVRAAKRVPGSVVVVSIFVNPMQFGAG
EDLDAYPRTPDDDLAQLRAEGVEIAFTPTTAAMYPDGLRTTVQPGPLAAELEGGPRPTHFAGVLTVVLKLLQIVR
PDRVFFGEKDYQQLVLIRQLVADFNLDVAVVGVPTVREADGLAMSSRNRYLDPAQRAAAVALSAALTAAAHAATA
GAQAALDAARAVLDAAPGVAVDYLELRDIGLGPMPLNGSGRLLVAARLGTTRLLDNIAIEIGTFAGTDRPDGYRA
ILESHWRNKLAAALEHHHHHH

$M_{calc.}$ = 34 032, ε = 15 930 $M^{-1}cm^{-1}$

Appendix 5 - *Sequence of panC gene coding PanC-wt*

(NCBI website, http://www.ncbi.nlm.nih.gov/gene/885459)

FASTA: >gi|57116681:c4045210-4044281 Mycobacterium tuberculosis H37Rv, complete genome

ATGACGATTCCTGCGTTCCATCCCGGTGAACTCAATGTGTACTCGGCACCGGGGGATGTCGCCGATGTCAGTCGC
GCACTGCGACTCACCGGCCGGCGAGTGATGTTGGTGCCTACTATGGGTGCGCTGCACGAAGGCCACCTCGCGTTG
GTGCGTGCGGCCAAGCGGGTGCCCGGATCGGTCGTCGTCGTGTCGATCTTCGTCAACCCGATGCAATTCGGTGCC
GGGGAAGATCTCGACGCCTATCCCCGCACCCCGGACGACGACCTGGCGCAACTGCGGGCCGAAGGCGTGGAAATC
GCTTTCACGCCAACTACCGCGGCGATGTATCCCGACGGCCTGCGCACCACCGTGCAACCCGGTCCGTTGGCCGCC
GAACTCGAGGGCGGCCCGCGGCCAACCCATTTCGCCGGCGTGCTGACGGTCGTGCTAAAGCTGCTGCAGATCGTG
CGCCCGGATCGGGTGTTCTTCGGTGAGAAGGACTACCAGCAGCTGGTGCTGATCCGGCAGCTGGTCGCGGACTTC
AACCTCGATGTCGCGGTGGTCGGCGTGCCGACCGTGCGCGAAGCCGACGGGCTGGCGATGTCGTCGCGCAACCGC
TACCTGGACCCGGCCCAGCGTGCGGCGGCCGTCGCGCTCTCGGCGGCGCTAACGGCCGCAGCGCATGCGGCAACG
GCTGGCGCGCAGGCCGCGCTGGATGCCGCCCGTGCGGTGCTCGACGCTGCACCCGGCGTGGCGGTCGACTACCTG
GAGCTGCGCGATATCGGGCTTGGCCCGATGCCGCTCAACGGTTCCGGTCGGCTGCTGGTTGCTGCCCGGCTTGGC
ACCACCAGGCTGCTGGACAACATTGCGATTGAAATCGGAACTTTCGCCGGCACCGACCGCCCGGACGGATACCGG
GCAATCCTCGAATCACATTGGAGAAACTGA

References

(1) Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J. C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F. *Methods Mol Biol* **1999**, *112*, 531.

(2) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chem. Sci.* **2011**, *2*, 1666.

(3) Zheng, R.; Blanchard, J. S. *Biochemistry* **2001**, *40*, 12904.

(4) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M. F.; Walker, R.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A. E.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. *AMBER 11* University of California, San Francisco, 2010.

(5) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(6) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.

(7) Walker, R. C.; de Souza, M. M.; Mercer, I. P.; Gould, I. R.; Klug, D. R. *J. Phys. Chem. B* **2002**, *106*, 11658.

(8) Pavelites, J. J.; Gao, J.; Bash, P. A.; Mackerell, A. D. *J. Comput. Chem.* **1997**, *18*, 221.

(9) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33.

(10) Kuhn, P.; Knapp, M.; Soltis, S. M.; Ganshaw, G.; Thoene, M.; Bott, R. *Biochemistry* **1998**, *37*, 13446.

(11) Wang, S.; Eisenberg, D. *Protein Science* **2003**, *12*, 1097.