– Supporting Information –

Optimization of the Interligand Overhauser Effect for fragment linking: application to inhibitor discovery against *Mycobacterium tuberculosis* pantothenate synthetase

Pawel Sledz,[1] H. Leonardo Silvestre,[2] Alvin W. Hung,[1] Alessio Ciulli,[1] Tom L. Blundell[2] and Chris Abell[1*]

[1] University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom
[2] Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, United Kingdom

Correspondence should be addressed to: Prof. Chris Abell, email: ca26@cam.ac.uk.

Contents:

1. Protein preparation and handling

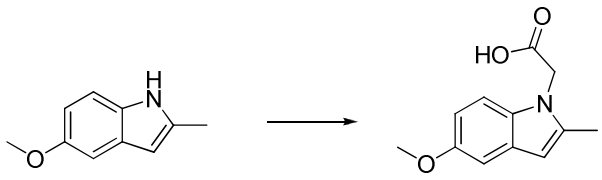
Mycobacterium tuberculosis pantothenate synthetase for biophysical and crystallization studies has been obtained as previously reported.[S1]

For perdeuteriated protein preparation, the following minimal medium was used: 6.86 g/L $(NH_4)_2SO_4$, 1.56 g/L KH_2PO_4, 6.48 g/L $Na_2HPO_4 \times 2\ H_2O$, 0.49 g/L diammonium hydrogen citrate, 0.25 g/L $MgSO_4 \times 7\ H_2O$, 5 g/L glycerol and 1 mL/L of 1000× trace elements stock solution (0.5 g/L $CaCl_2 \times 2\ H_2O$, 16.7 $FeCl_3 \times 6\ H_2O$, 0.18 g/L $CuSO_4 \times 5\ H_2O$, 0.11 g/L $MnSO_4 \times H_2O$, 0.18 g/L $CoCl_2 \times 6\ H_2O$ and 20.1 g/L EDTA). Protein expression was optimized in this medium prior to proceeding with perdeuteriation yielding ~40 mg of protein per liter of the culture. Residual H_2O was removed from salts by dissolving them in D_2O and lyophilizing; the procedure was repeated twice. Deuteriated glycerol solution was prepared in D_2O. All the solutions were autoclaved prior to use. Cells have been adapted to grow in deuteriated medium by growing subsequent seed cultures in rich LB medium, then in 0, 30, 50, 70 and 100% heavy water based M9 (12 h with exception of the 100% D_2O seed culture which was grown for 36 h). Subsequently, a 140 mL culture has been inoculated with 1:40 v/v of steady state seed culture, grown at 37 ^oC to the OD_{600} of 0.6 and induced with 0.5 mM IPTG overnight at 30 ^oC. Afterwards protein was treated in the same way as non-deuteriated protein. Final yield was 3.8 mg of perdeuteriated PS from 140ml of the culture (~28mg per litre). ESI-MS analysis confirmed ~95% deuteration level at non-exchangeable positions (mass of the deuteriated monomer of 39474 Da as compared with mass of 37430 Da for unmodified protein).

2. Chemicals and synthesis

All reagents and starting materials were used as obtained from commercial sources unless otherwise indicated. All reactions were performed under nitrogen. Column chromatography was performed using silica gel (230 – 400 mesh). ^{1}H NMR spectra were recorded on either a Bruker DPX-400 MHz or a Bruker DPX-500 MHz spectrometer. High-resolution mass spectrometry (HRMS) was carried out using a Micromass Quadrupole-Time of flight spectrometer. Liquid-chromatography mass spectrometry (LCMS) was carried out using an Alliance HT Waters 2395 Separations Module coupled to a photomultiplier detection system. In LCMS, the first eluent was 10 mM ammonium acetate and the second eluent 95% aqueous acetonitrile. Samples were run on a gradient from 0-100% over a period of 8 min. Infrared (IR) spectroscopy of solids were recorded on a Perkin Elmer Spectrum One FTIR spectrometer using attenuated transmittance reflectance.

2-(5-methoxy-2-methyl-1H-indol-1-yl)acetic acid (5)



4 **5**

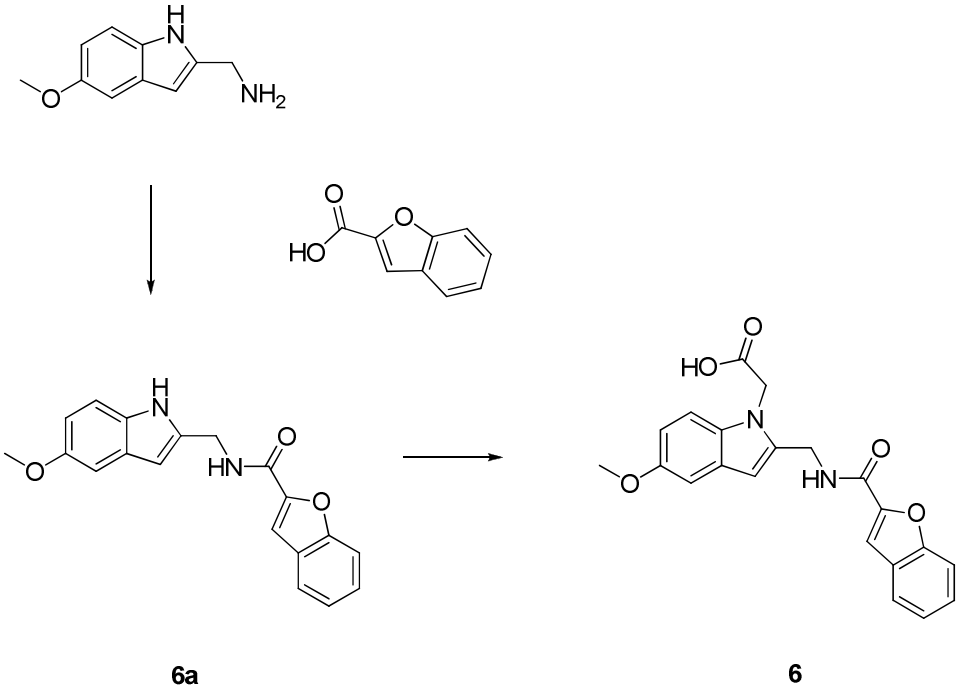
To a stirred solution of 5-methoxy-2-methylindole **4** (0.400 g, 2.48 mmol) in DMF (10 mL) was added potassium carbonate (0.680 g, 4.93 mmol) and *tert*-butylbromoacetate (0. 44 mL, 3.11 mmol). The mixture was stirred at 25 ^{o}C for 20 h and concentrated *in vacuo*. Water (15 mL) was added to the mixture and the aqueous phase extracted with ethyl acetate (3 x 15 mL). The combined extracts were dried over sodium sulfate and concentrated *in vacuo*. Purification by column chromatography (1:4 ethyl acetate/hexane) gave the *tert*-butyl protected indole as a solid (49%).

To a solution of the *tert*-butyl protected indole (80 mg, 0.37 mmol) in dichloromethane (3 mL) was added TFA (3 mL) dropwise. The mixture was stirred for 0.5 h at 25 ^{o}C and concentrated *in vacuo*. Saturated aqueous sodium bicarbonate (5 mL) was added to the

residual solid and washed with diethyl ether (2 x 5 mL). The aqueous phase was cooled to 0°C was acidified with 1M HCl until white precipitate forms. Collection of the white precipitate by filtration gave product **5** as a solid (35 mg, 55%).

^{1}H NMR (500 MHz, d^6-acetone): δ (ppm) 7.20 (d, J = 8.8 Hz, 1H), 6.96 (d, J = 2.4 Hz, 1H), 6.71 (dd, J = 2.4 Hz, J = 8.9 Hz, 1H), 6.16 (s, 1H), 4.92 (s, 2H), 3.77 (s, 3H), 2.37(s, 3H).
^{13}C NMR (125 MH, d^6-acetone): δ (ppm) 170.9, 155.6, 138.7, 134.0, 130.1, 111.2, 110.5, 103.0, 101.3, 56.2, 45.2, 12.9.
IR, vmax (ATR): 2901, 1702 cm^{-1}.
LCMS ([M – H$^+$]$^-$ 218.0), retention time 3.4 min.
HRMS ((ES) for $C_{12}H_{14}NO_3$ (MH^+) calcd: 220.0974, found: 220.0972.

2-(2-((Benzofuran-2-carboxamido)methyl)-5-methoxy-1*H*-indol-1yl)acetic acid (6)



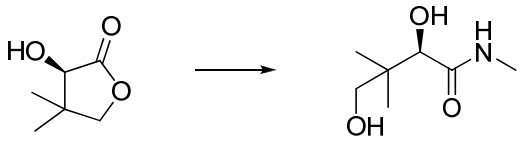
To a stirred solution of (5-methoxy-1H-indol-2-yl)methanamine[S2] (100 mg, 0.57 mmol) in dichloromethane (20 mL) was added EDCI (210 mg, 1.10 mmol), DMAP (5 mg, cat.), benzofuran-2-carboxylic acid (101 mg, 0.62 mmol) and triethylamine (0.24 mL, 1.70 mmol).The mixture was stirred for 20 h at 25 °C. Saturated aqueous sodium bicarbonate (15 mL) was added to the reaction mixture and the aqueous phase extracted with dichloromethane (3 x 15 mL). The combined extracts were dried over sodium sulfate and concentrated *in vacuo*. Purification by column chromatography (1:4 ethyl acetate/hexane) gave the N-((5-

methoxy-1H-indol-2-yl)methyl)benzofuran-2-carboxamide **6a** (150 mg, 83%) as a white solid.

To a stirred solution of the N-((5-methoxy-1H-indol-2-yl)methyl)benzofuran-2-carboxamide (60 mg, 0.19 mmol) was added bromomethylacetate (25 µl, 0.25 mmol), potassium carbonate (77 mg, 0.56 mmol) in dimethylformamide (4 mL). The mixture was stirred for 16 h at 25 ^{o}C. Saturated aqueous sodium bicarbonate (15 mL) was added to the reaction mixture and the aqueous phase extracted with dichloromethane (3 x 15 mL). The combined extracts were dried over sodium sulfate and concentrated *in vacuo*. Purification by column chromatography (1:4 ethyl acetate/hexane) gave the methyl ester protected indole amide (12 mg, 16%). To a solution of this methyl ester protected indole (0.012 g, 0.03 mmol) in THF (2 mL) was added 2 M aqueous lithium hydroxide (0.3 mL). The mixture was stirred for 2 h at 25 ^{o}C and concentrated *in vacuo*. Water (10 mL) was added to the residual solid, and the obtained aqueous solution was washed with ethyl acetate (2 x 10 mL). The aqueous phase was acidified with 1 M aqueous HCl until pH = 4 and extracted with ethyl acetate (3 x 15 mL). The combined extracts were dried over sodium sulfate and concentrated *in vacuo* to give **6** (10 mg, 86%) as a white solid.

^{1}H NMR (500 MHz, d^{6}-acetone): δ (ppm) 8.21 (m, 1H), 7.75 (d, J = 7.8 Hz, 1H), 7. 53 -7.49 (m, 2H), 7.44 (td, J = 7.2 Hz, J = 1.3 Hz, 1H), 7.32 (m, 1H), 7.26 (d, J = 9.1 Hz, 1H), 7.05 (d, J = 2.4 Hz, 1H), 6.80 (dd, J = 9.1 Hz, J =2.4 Hz, 1H), 6.48 (s, 1H), 5.12 (s, 2H), 4.80 (d, J = 5.8, 2H), 3.78 (s, 3H).
^{13}C NMR (125 MHz, d^{6}-acetone): δ (ppm) 171.1, 159.2, 156.0, 155.8, 150.6, 138.8, 134.4, 129.5, 128.9, 128.0, 124.9, 123.9, 112.9, 112.8, 111.2, 110.9, 103.5, 103.5, 56.2, 45.7, 36.1.
IR, ν_{max} (ATR): 2902, 1719, 1594 cm^{-1}.
LCMS (MH^+ 377.5), retention time 3.8 min.
HRMS (ES) for $C_{21}H_{18}N_2O_5Na$ (MNa^+) calcd: 401.1113, found: 401.1130.

2,4-Dihydroxy-3,3,*N*-trimethylbutyramide (3)



3

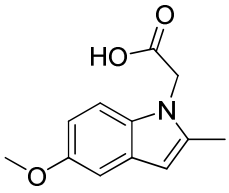
Compound **3** was prepared according to the literature procedure.[S3] To a stirred solution of R-pantolactone (1.00 g, 7.69 mmol) in methanol (3 ml) 40% aqueous solution of *N*-methylamine was added (0.8 ml, 0.32 g, 10.3 mmol). Reaction mixture was stirred at 25 ^{o}C for 3 h until

completion and then solvents and unreacted *N*-methylamine were removed *in vacuo* yielding pure product **3** (1.24 g, quantitative yield).

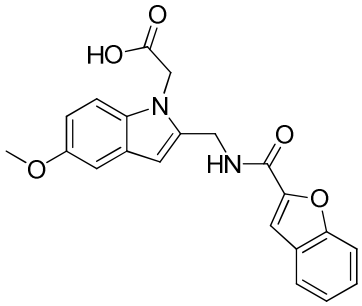
^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) 7.07 (s, 1H), 3,96 (s, 1H), 3.43 (d, J = 11.1 Hz, 1H), 3.39 (d, 11.1 Hz, 1H), 2.77 (d, J = 5.0 Hz, 3H), 0.91 (s, 3H), 0.85 (s, 3H).
^{13}C NMR (125 MHz, $CDCl_3$): δ (ppm) 174.5, 77.2, 70.9, 39.2, 25.7, 21.0, 20.3.
IR, ν_{max} (ATR): 3334.1 2961.9, 1539.2 cm^{-1}.
HRMS (ES) for $C_7H_{16}N_1O_3$ (MH^+) calcd: 162.1130, found: 162.1125.

3. ITC experiments

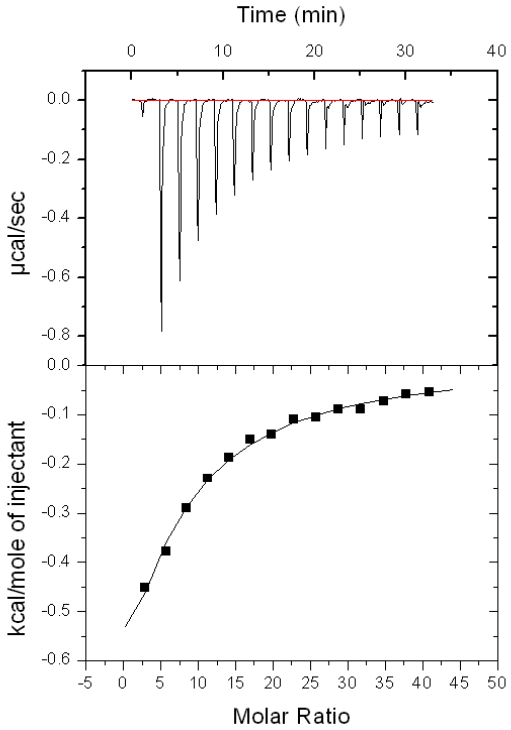
ITC experiments were performed using the ITC200 instrument from Microcal Inc. (GE Healthcare) at 25 ^{o}C. His_6-PS was buffer exchanged into 50 mM HEPES-HCl, pH 7.6 containing 50 mM NaCl and 5 mM $MgCl_2$, and loaded into the ITC cell at concentration of 60 µM with 10 % v/v DMSO solution. Ligands were dissolved in the same buffer/DMSO percentage. Typically, 16 injections of 2.4 µL each were performed over a period of 30 min. Data was fitted to single binding site model using the Origin software package provided by the manufacturer. ITC traces together with their interpretation for novel compounds discussed in the paper are presented below:

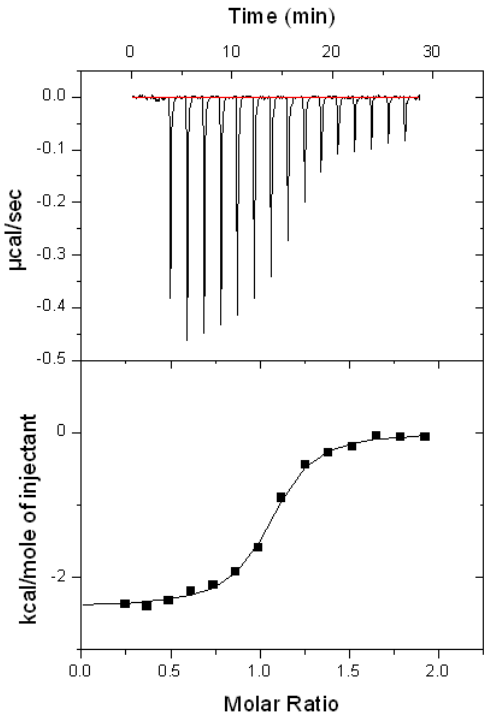


5



6





K_D	800 ± 30 µM	860 ± 100 nM
ΔH	- 10.7 ± 0.2 kcal/mol	-2.41 ± 0.03 kcal/mol
ΔS	-22.1 cal/(mol ^{o}C)	19.7 cal/(mol ^{o}C)

4. X-ray diffraction experiments

Crystals of *M.tb.* pantothenate synthetase were obtained and soaked with the ligand as previously described, then diffraction data was collected and refined as previously reported.[S1] Structure of ternary complex of ligands **1**, **2** and the protein has been previously reported and deposited in the Protein Data Bank with the accession code 3IMG.[S4] Structure of the complex of the protein with final linked inhibitor **6** has been deposited in the PDB with the accession code 3LE8. All derived data was indexed and scaled using iMosflm and Scala. Refinement was carried out using the graphical interface of the CCP4 suite, running Refmac 5.0.[S5] Model building was done using COOT.[S6] Data collection and refinement statistics for all the refined coordinate sets are presented in Table S1.

Table S1 - Crystallographic data collection and refinement statistics for compound 6

Data Collection	**6**
X-Ray Source	Diamond, I02
Space Group	$P2_1$
Cell Parameters ($\alpha=\gamma=90^0$)	
a	48.58
b	71.14
c	81.49
β	99.04
Resolution range, A (outer shell)	(38.63-1.70) (1.79-1.70)
Total no. of reflections	210501
No. of unique reflections	59906
Multiplicity	3.5
R merge % (outer shell)	60.0
Average I/s(I)	2.1
Completeness % (outer shell)	99.7
Mosaicity, ^{0}	0.92
Wilson B	20.9
Refinement	
Resolution range	80-1.70
R_{cryst}, %	17.2
R_{free}, %	22.4
Number of reflections	
working set	56868
test set	2843
Number of protein atoms	4330
Water molecules	583
Ligand	2
Model quality	
Average isotropic B value (protein)	25.5
Average Isotropic B value (waters)	40.6
Average isotropic B value (ligand)	25.5

5. NMR experiments

All NMR experiments were carried out using Bruker Avance 500MHz with TCI cryoprobe (Department of Chemistry, University of Cambridge). ^{1}H-^{1}H 2D NOESY experiments[S7] were performed typically with 24 scans for each of 512 points in indirect dimension, relaxation delay of 1.5 s with mixing times of typically 300 and 700 ms, which summed up to 8-11 h of experiment time.

Experiments were performed using 200 μL capillaries with 20 μM protein (perdeuteriated or wild type PtS) in 50 mM TRIS/HCl pH = 8.0, 50mM NaCl and 5 mM $MgCl_2$, with 10% of D_2O and 20 μM TSP (sodium 3-trimethylsilyltetradeuteriopropionate) as internal standard. Samples contained typically 1-2% of DMSO-d6 as a result of addition of chemical stock solutions in DMSO-d6.

Additional NMR experiments are shown in the Figures S1-S6. Positive peaks are shown in blue and negative in green.

6. Coupled enzymatic assay for pantothenate synthetase

Enzyme kinetic assay – Pantothenate synthetase activity was assayed by coupling the formation of AMP to the oxidation of NADH with myokinase, pyruvate kinase and lactate dehydrogenase as previously described.[S1,S4,S8] K_M of the substrate ATP was calculated to be 0.15 mM. K_I of the compound **6** was calculated to be 5.4 μM (Fig S7).

7. References

.

S1. Ciulli, A.; Scott, D.E.; Ando, M.; Reyes, F.; Saldanha, A.; Tuck, K.L.; Chirgadze, D.Y.; Blundell, T.L.; Abell, C. *ChemBioChem* **2008**, *9,* 2606-2611.

S2. Cruces, M. A.; Elorriaga, C.; Fernández-Alvarez, E.; Lopez, O. N. *Eur. J. Med. Chem* **1990**, *25*, 257-265.

S3. Barrios, I.; Camps, P.' Comes-Franchini, M.; Munoz-Torrero, D.; Ricci, A.; Sanchez, L. *Tetrahedron*, (**2003**), *59*, 1971-1979.

S4. Hung, A.; Silvestre, H.L.; Wen, S.; Ciulli, A.; Blundell, T.L.; Abell, C. *Angew. Chem. Int. Ed.* **2009**, *48*, 8452-8456.

S5. Murshudov, G.N.; Vagin, A.A.; Dodson, E.J. *Acta Cryst. D* (**1997**), *D53*, 240-255.

S6. Emsley, P.; Cowtan, K. *Acta Cryst.* (**2004**), *D60,* 2126-2132.

S7. Liu, M.; Mao, X.; He, C.; Huang, H.; Nicholson, J.K.; Lindon, J.C. *J. Magn. Reson.* (**1998**), *132*, 125-129.

S8. Zheng, R.; Blanchard, J. S. *Biochemistry* **2001**, *40*, 12904-12912.

S9. Feng, B.Y.; Shoichet, B.K. *Nature Protocols* (**2006**), *1:2*, 550-553.

8. Figures

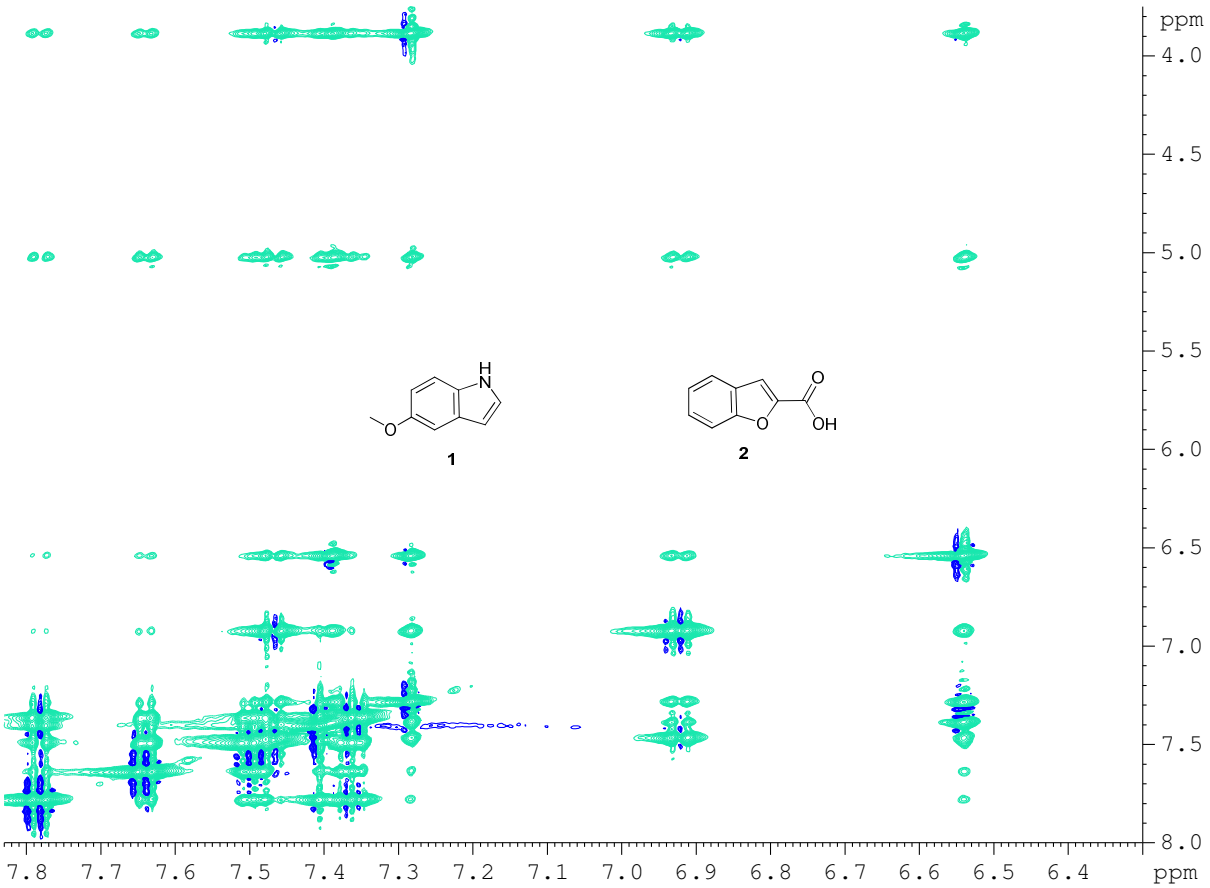


Figure S1. Initially obtained NOESY spectrum with compounds **1**, **2** and dPtS (negative peaks in green). ILOE peaks between ligands **1** and **2** were observed but failed to yield structural information as peaks between all possible combinations of protons were seen. This could indicate the spin diffusion present in the system or significant contribution from non-specific binding.

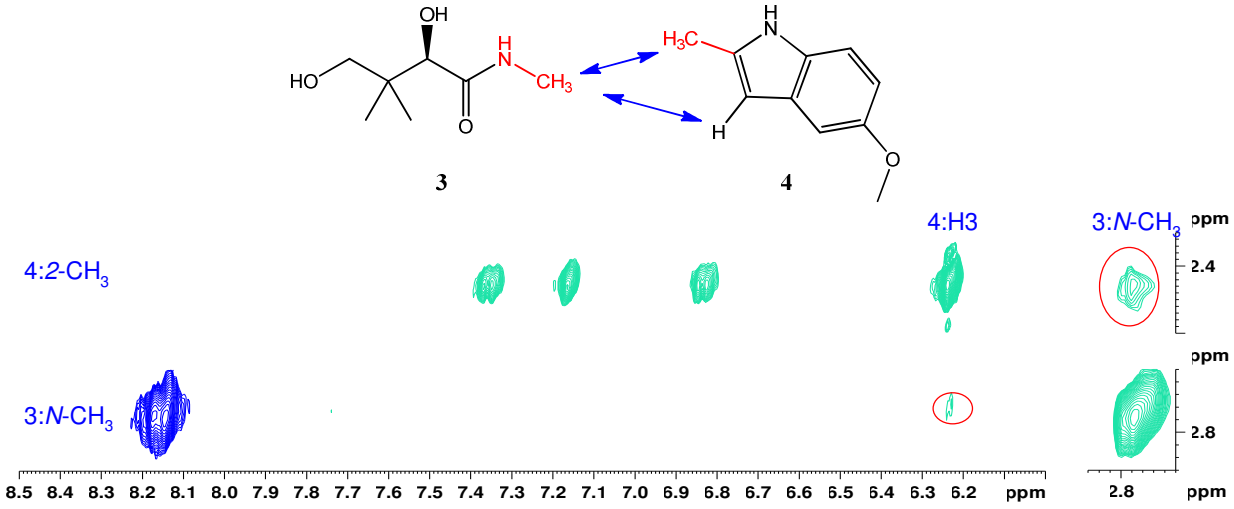


Figure S2. A fragment of 2D NOESY spectrum confirms the binding mode of **4**. ILOE peaks were observed between *N*-CH_3 of **3** and 2-CH_3 and H3 of compound **4** which unambiguously

confirmed its binding mode with 2-Me group facing the pantoate pocket (10 mM **3**, 3 mM **4** and 20 μM dPtS used in the experiment).

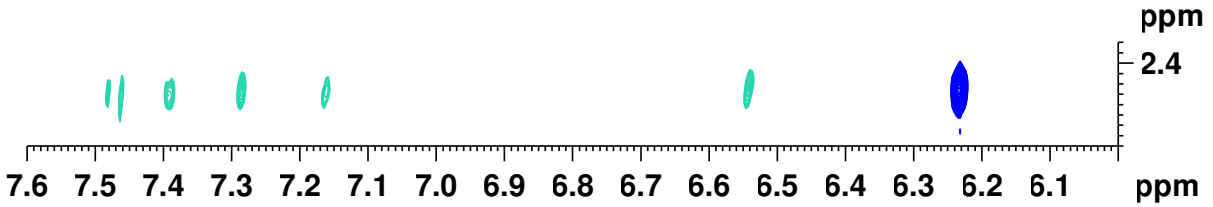


Figure S3. A fragment of 2D NOESY spectrum of **1** and **4** (both at 3 mM concentration) in the absence of the protein. Some intramolecular trNOEs are positive, the rest and all interligand peaks are negative. This shows ligand coaggregation in aqueous environment – interligand aggregates give rise to the negative NOE building up in the slow tumbling motion regime. It was suggested previously by Feng and Shoichet[S9] that such aggregates can be removed upon addition of a detergent eg. Triton X, however this was not successful in our case. After consultation with Prof. Shoichet an attempt to remove aggregates by extensive centrifugation was taken which led to disappearance of protein-independent ILOEs in our model system supporting the presence of aggregates.

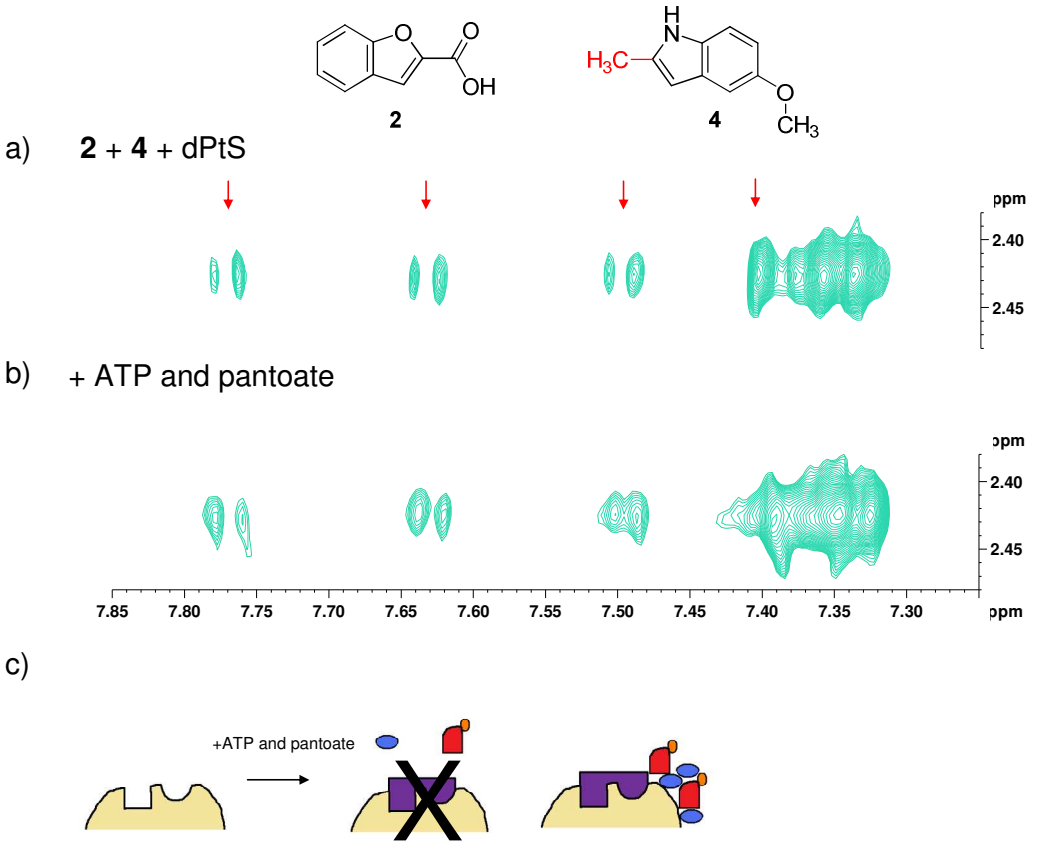


Figure S4. a) a fragment of the 2D NOESY spectrum with compounds **2**, **4** (at 3 mM concentration) and perdeuteriated pantothenate synthetase (20 μM), ILOE peaks marked with red arrows; b) as in a) but upon addition of 1 mM ATP and 1 mM pantoate – ILOE peaks are still observed and practically undiminished; c) scheme showing proposed interpretation - lack of displacement explains why ILOE did not provide the structural information in the case of **2** and **4** – it arose from nonspecific binding contributions.

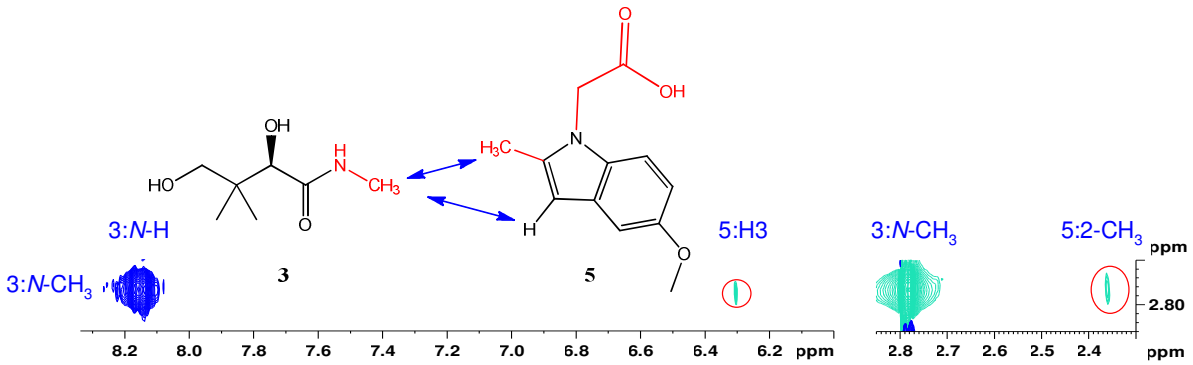


Figure S5. A fragment of 2D NOESY spectrum confirms the binding mode of **5**. ILOE peaks were observed between *N*-CH_3 of **3** and 2-CH_3 and H3 of compound **5** which unambiguously confirmed its binding mode with 2-Me group facing the pantoate pocket (10 mM **3**, 3 mM **5** and 20 μM dPtS used in the experiment).

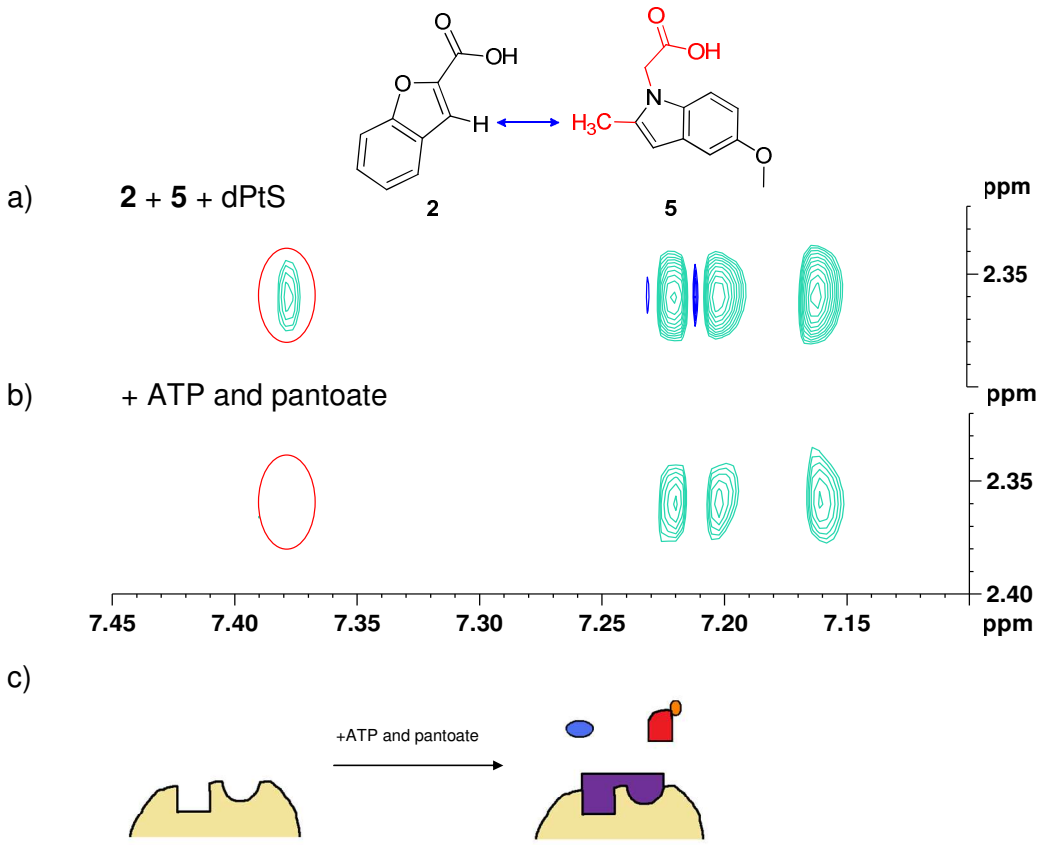


Figure S6. a) a fragment of the 2D NOESY spectrum with compounds **2**, **5** (at 3 mM concentration) and pantothenate synthetase (20 μM), ILOE peak marked with red circle; b) as in a) but upon addition of 1 mM ATP and 1 mM pantoate – ILOE peak disappears; c) scheme showing proposed interpretation - displacement of ILOE proves that it rose from specific binding contribution and is indeed a genuine ILOE peak as opposed to the data obtained with compounds **2** and **4**.

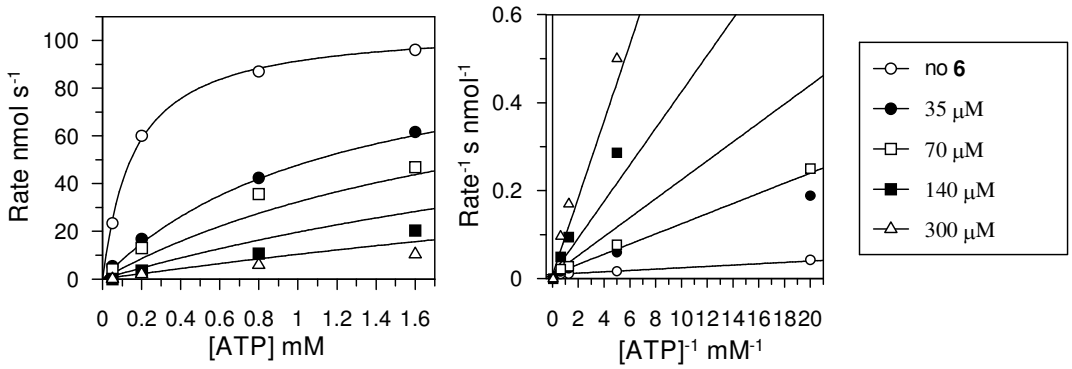


Figure S7. Michaelis-Menten and Lineweaver-Burk plots for the kinetics of the pantothenate synthesis inhibited by the presence of compound **6.** The data shows competitive inhibition with ATP.