Supplementary Material (ESI) for Chemical Communications

Supplementary Information

Palladium-Mediated Site-Selective Suzuki-Miyaura Protein Modification at Genetically Encoded Aryl Halides

Christopher D. Spicer and Benjamin G. Davis

Chemistry Research Laboratory, University of Oxford, 12 MansfieldRoad, Oxford, UK OX1 3TA. E-mail: Ben.Davis@chem.ox.ac.uk;

Fax: +44 (0)1865 275674; Tel: +44 (0)1865 275652

Table of contents

S1 General Considerations

S4 Chemical Syntheses

S5 Site Directed Mutagenesis

S7 Protein Expressions

S8 Initial Cross-Coupling Attempts on MBP

S10 Palladium Adducts on Np276

S11 Scavenging on Myoglobin

S13 Cross-Couplings on MBP

S14 Control reactions

General Considerations

Melting points (m.p.) were recorded on a Leica Galen III hot stage microscope equipped with a Testo 720 thermocouple probe. Proton nuclear magnetic resonance (1H NMR) spectra were recorded on a Bruker AV400 (400 MHz) spectrometer.. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard (1H NMR: D_2O = 4.79; DMSO-d6 = 2.50). Coupling constants (*J*) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet.

Supplementary Material (ESI) for Chemical Communications

Infrared (IR) spectra were recorded on a Bruker Tensor 27 Fourier Transform spectrophotometer using KBr discs. Absorption maxima (υmax) are reported in wavenumbers (cm-1).

Low resolution mass spectra (LRMS) were recorded on a Waters Micromass LCT Premier TOF spectrometer Waters using electrospray ionization (ESI).

Optical rotations were measured on a Perkin-Elmer 241 polarimeter with a path length of 1.0 dm and are reported with implied units of 10^{-1} deg cm^2 g^{-1}. Concentrations (c) are given in g/100 ml.

Thin layer chromatography (TLC) was carried out using Merck aluminium backed sheets coated with 60F254 silica gel. Visualization of the silica plates was achieved using a UV lamp (λmax = 254 nm), and/or ammonium molybdate (5% in 2M H_2SO_4), and/or potassium permanganate (5% $KMnO_4$ in 1M NaOH with 5% potassium carbonate).

All solvents were used as supplied (Analytical or HPLC grade), without prior purification. Deionized water was used for chemical reactions and Milli-Q purified water for protein manipulations. Reagents were purchased from Aldrich and used as supplied, unless otherwise indicated.

Liquid chromatography-mass spectrometry (LC-MS) was performed on a Micromass LCT (ESI-TOF-MS) coupled to an Agilent 1100 Series HPLC using a Phenomenex Jupiter C4 column (250 × 4.6 mm × 5μm). Water:acetonitrile, 95:5 (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid by volume, were used as the mobile phase at a flow rate of 1.0 mL min^{-1}. The gradient was programmed as follows:

Supplementary Material (ESI) for Chemical Communications

Time/ min	Solvent A/ %	Solvent B/ %
0	95	5
3	95	5
10	65	35
15	35	65
25	35	65
27	5	95
31	5	95
33	95	5
35	95	5

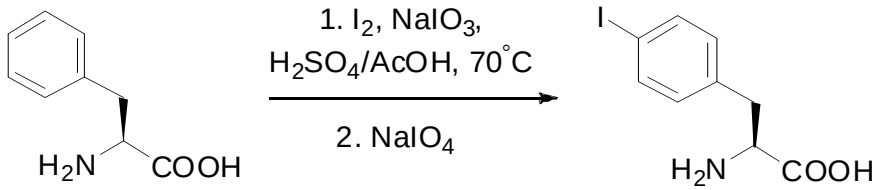
The electrospray source was operated with a capillary voltage of 3.2 kV and a cone voltage of 25 V. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 l hr^{-1}. Proteins typically elute between 13 and 17 minutes using this method. Spectra were calibrated using a calibration curve constructed from a minimum of 17 matched peaks from the multiply charged ion series of equine myoglobin obtained at a cone voltage of 25V. Total mass spectra were obtained from the ion series using the MaxEnt algorithm preinstalled on MassLynx software (v. 4.0 from Waters) according to manufacturer's instructions.

Plasmid pEVOL (pIPhe) was generously donated by the group of Prof. P. G. Schultz at The Scripps Institute. Plasmid pDB.His.MBP was obtained from the Berkeley Structural Genomics Centre through the PlasmID database. Equine myoglobin was purchased from Sigma-Aldrich. SDS-PAGE gels were run using pre-cast gels purchased from Invitrogen (NuPAGE 4-12% Bis-Tris gel).

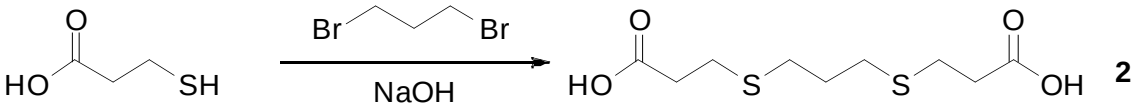
Chloramphenicol and kanamycin were both used at a concentration of 50 mg L^{-1}.

Supplementary Material (ESI) for Chemical Communications

Chemical Syntheses



Iodine (25.2 g, 100 mmol) and sodium iodate (9.9 g, 50 mmol) were added to a solution of L-phenylalanine (41.25 g, 250 mmol) in acetic acid (250 mL) and concentrated sulphuric acid (31 mL). The mixture was heated to 70 ºC for 18 hrs, and two portions of sodium periodate (2 × 1.05 g) were then added. After continuing to stir at 70ºC for ~30 min the mixture turned orange. The acetic acid was then removed *in vacuo*, and the crude mixture was diluted with water (400 mL), washed with ether (2 x 200 mL) and DCM (2 x 300 mL). To the aqueous layer was added 2M NaOH until a white precipitate was formed. The solid was collected by filtration and re-crystallised from boiling water:ethanol (160 mL:100 mL). The resulting crystals were collected and dried *in vacuo*, to yield *p*-Iodophenylalanine (**1**) as white crystals. A yield of 33.01 g, 113 mmol (57 %) was obtained. Spectroscopic data was consistent with that previously reported. ^{1}H NMR (400MHz, NaOD/D_2O): δ= 7.60 (2H, d, *J* = 8.7 Hz, *ortho*-H), 6.93 (2H, d, *J* = 8.7 Hz, *meta*-H), 3.37 (1H, t, *J* = 6.1 Hz, H_α), 2.82 (1H, dd, *J* = 14.5, 6.1 Hz, -CH_2Ar ABX system), 2.68 (1H, dd, *J* = 14.5, 6.1 Hz, -CH_2Ar ABX system); m.p- 260-263 ºC; $[\alpha]_D$= -20;

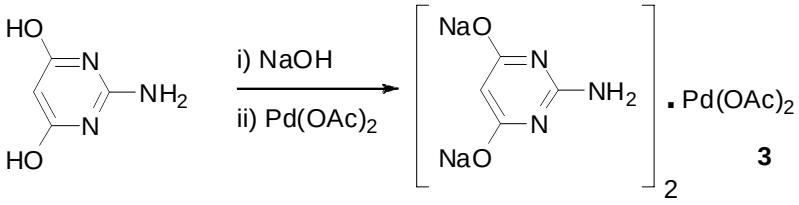


2

Sodium hydroxide (3.12 g, 78 mmol), was dissolved in methanol (70 mL), and 3-mercaptopropionic cid (3.6 mL, 41 mmol) was then added. 1,3-Dibromopropane (2 mL, 19 mmol) dissolved in methanol (10 mL) was then added drop-wise, and the mixture stirred at room temperature for 2 hours. It was then refluxed for 18 hrs, then allowed to cool to rt,

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011

resulting in the formation of a white ppt. The methanol was removed *in vacuo*, and the resultant sodium salt re-dissolved in water (80 mL). The product was then precipitated by addition of concentrated HCl to a pH of 4. The white solid was collected by filtration, and dried *in vacuo* for 24 hr. A yield of 3.78 g, 15 mmol (79%) was obtained. Spectroscopic data was consistent with that previously reported. ^{1}H NMR (400 MHz, DMSO): δ = 3.08 (4H, t, *J* =8.0 Hz), 2.99 (4H, t, *J* = 7.2 Hz), 2.91 (4H, t, *J* = 6.8 Hz), 2.16 (2H, quin., *J* = 7.2 Hz) ppm;



3

2-amino-4,6-dihydroxypyrimidine (13 mg, 0.1 mmol) was dissolved in NaOH (2 mL, 0.1 M) at 65 ºC. Palladium acetate (11 mg, 0.05 mmol) was then added and the solution stirred at 65 ºC for 30 min. The orange solution was then allowed to cool and diluted to 5 mL with distilled water to give a stock 0.01 M catalyst solution.

Site-Directed Mutagenesis

Site-directed mutagenesis was undertaken on plasmid pDB.His.MBP using a Stratagene QuikChange II Site-Directed Mutagenesis Kit, utilising the following primers, purchased from Sigma-Aldrich:

5'- CACGGTACCAAAACTTAGGAAGGTAAACTGG-3' (Forward Primer)

5'- CCAGTTTACCTTCCTAAGTTTTGGTACCGTG- 3' (Reverse Primer)

Resultant colonies, after transformation of the mutagenesis product into *E. coli* XL-1 Blue supercompetent cells, were sent for sequencing by Geneservice, using T7 forward and reverse

Supplementary Material (ESI) for Chemical Communications

primers. The desired mutation was shown to have been installed, to give the mutant plasmid pMBP.E13·.

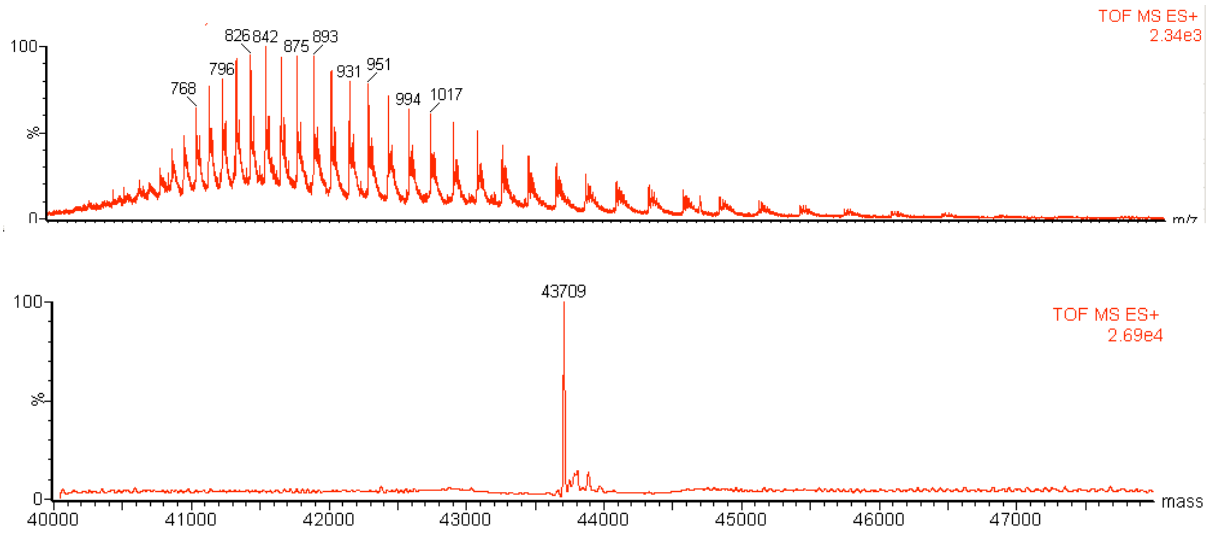
```
5′ATGGGCAGCAGCCATCATCATCATCACGGTACCAAAACTTAGGAAGGTAAACTGGTAATCTGG
ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAAT
TAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATG
GCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAA
ATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGG
CAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGA
ACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCG
CTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTT
CAAGTATGAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTC
TGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAA
GCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTG
GCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCGAAAGAGTTCCTCGAAAAC
TATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAA
GTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACCATGGAAAACGCCCAGAAAGGTG
AAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC
GCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTGGTACCGATTACGATATCCC
AACGACCGAAAACCTTTACTTCCAGGGCCATATGGCTAG –3′
```

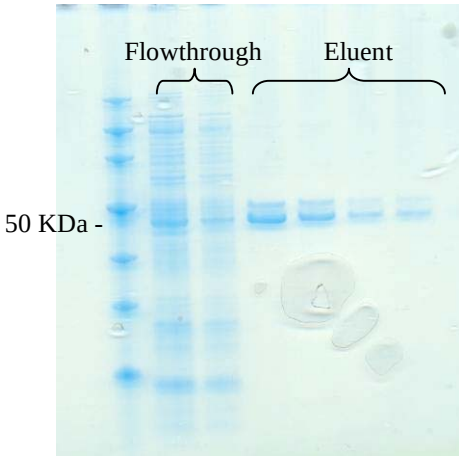
Protein Expression

Plasmid pMBP.E13· was co-transformed with plasmid pEVOL (pIPhe) into Novagen *E. coli* “Tuner (DE3)” cells, using the protocol supplied with the cells. The cells were grown overnight on agar plates containing chloramphenicol and kanamycin antibiotics. Cultures were then picked and grown in 10 mL LB broth overnight at 37 ºC, and then into 1 L of LB containing the appropriate antibiotics up to an $O.D_{600}$ = 0.6. At this point, protein expression was induced by the addition of 1 mM IPTG, 0.02 % *L*-Arabinose, and 2 mM *p*-Iodophenylalanine. The cells were then incubated at 30 ºC for 12 hours. Cells were then pelleted by centrifugation (8000 rpm, 8 min) and re-suspended in 15 mL of binding buffer (20 mM Tris-hydrochloride, 15 mm imidazole, 0.3 M NaCl, pH 7.8). Lysozyme (5 mg) and DNase (5 mg) were then added, and the mixtures shaken at 4ºC for 1 hour. The cells were then lysed by sonication (3 × 30 sec), and the cell debris collected by centrifugation (20000 rpm, 40 min). The supernatant was then purified by NiNTA affinity column washing first

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011

with binding buffer (~30 column volumes) and then eluting with elution buffer (20 mM Tris-hydrochloride, 500 mM imidazole, 0.3 M NaCl, pH 7.8). Eluent fractions were analysed by SDS-PAGE gel, and pure fractions combined and the protein precipitated overnight by the addition of ammonium sulphate to a final concentration of 85%. The protein was then collected by centrifugation (8000 rpm, 40 min) and re-suspended in 2.5 mL of phosphate buffer (20 mM NaH_2PO_4, 50 mM NaCl, pH 8.0). The protein was then purified by PD-10 column to give 3.5 mL of protein solution. Protein concentration was then determined by BCA assay (4.4 mg mL^{-1}). The protein was then analysed by LC-MS (Calc: 43708 Actual: 43709).



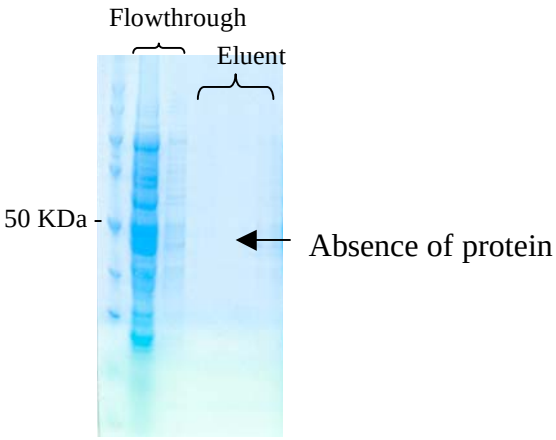


Protein Sequence: (X = pIPhe)

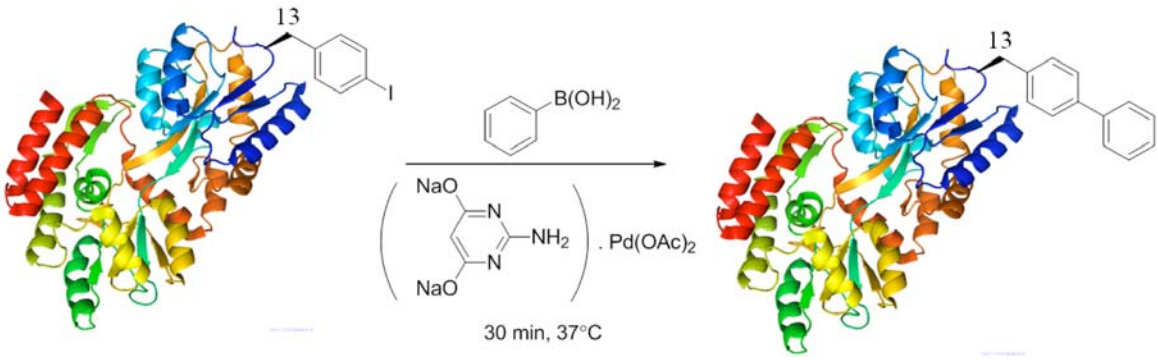
Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011

G S S H H H H H H G T K T X E G K L V I W I N G D K G Y N G L A E V G K K F E K D T G I
K V T V E H P D K L E E K F P Q V A A T G D G P D I I F W A H D R F G G Y A Q S G L L A
E I T P D K A F Q D K L Y P F T W D A V R Y N G K L I A Y P I A V E A L S L I Y N K D L
L P N P P K T W E E I P A L D K E L K A K G K S A L M F N L Q E P Y F T W P L I A A D
G G Y A F K Y E N G K Y D I K D V G V D N A G A K A G L T F L V D L I K N K H M N A D
T D Y S I A E A A F N K G E T A M T I N G P W A W S N I D T S K V N Y G V T V L P T F
K G Q P S K P F V G V L S A G I N A A S P N K E L A K E F L E N Y L L T D E G L E A V N
K D K P L G A V A L K S Y E E E L A K D P R I A A T M E N A Q K G E I M P N I P Q M S
A F W Y A V R T A V I N A A S G R Q T V D E A L K D A Q T G T D Y D I P T T E N L Y F
Q G H M A

Cells were also grown in the absence of *p*IPhe, and protein growth induced by the addition of L-Arabinose and IPTG. No protein was observed after purification as indicated by SDS-PAGE gel. This indicates the high fidelity of unnatural amino acid inorporation:

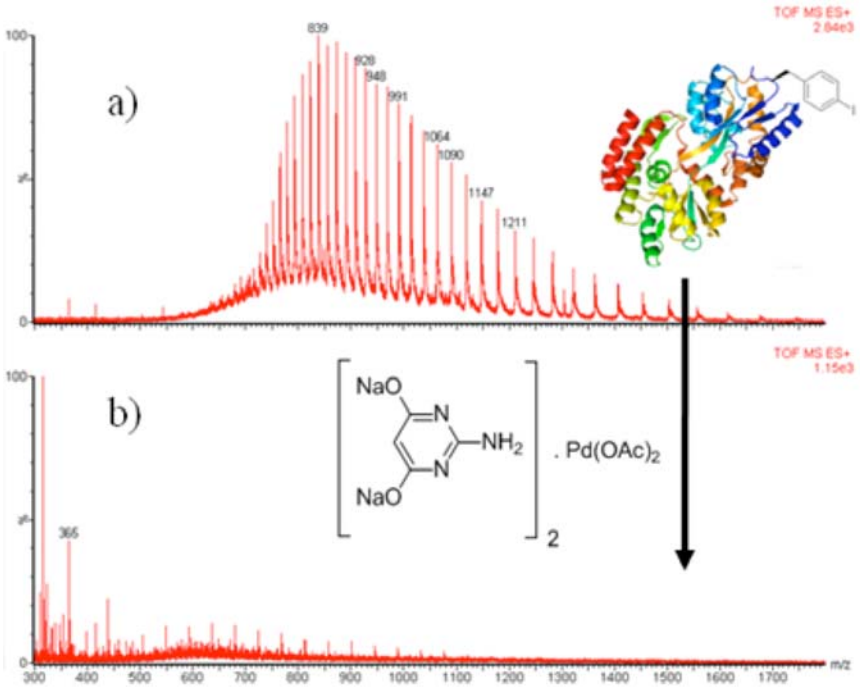


Initial Attempts on MBP



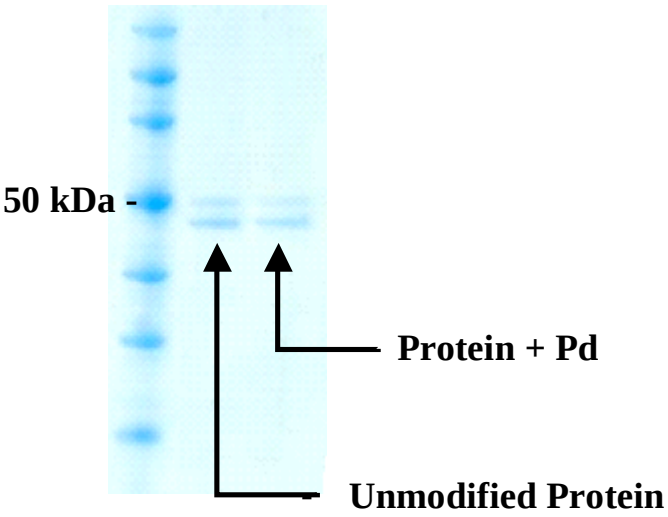
Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011

A 25 μL aliquot of MBP E13· (1.5 mg/mL in phosphate buffer, 0.4 nmol) was diluted with distilled water (25 μL). A stock solution of phenylboronic acid was prepared by dissolving 9 mg of phenylboronic acid (74 μmol) and 12 mg of Na_2HPO_4 (85 μmol) in 800 μL of water. An aliquot of this solution (7.5 μL, 0.7 μmol, 1800 equiv.) and an aliquot of Pd* solution (7.5 μL, 75 nmol, 188 equiv.) were added sequentially, and the reaction vortexed and then incubated at 37ºC for 45 min. The protein sample was then analysed by LC-MS, which showed the disappearance of the protein ion series:



The sample was then analysed by SDS-PAGE which indicated that the protein was still present within the reaction mixture undegraded:

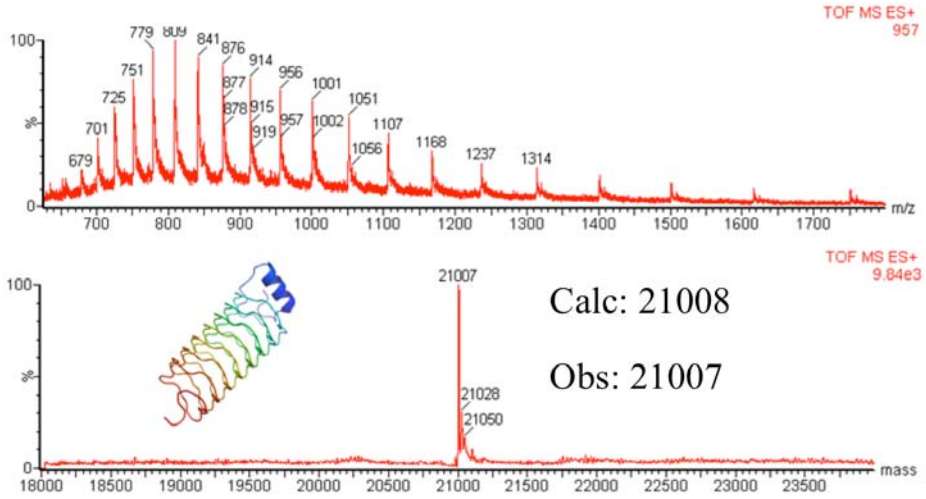
Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011



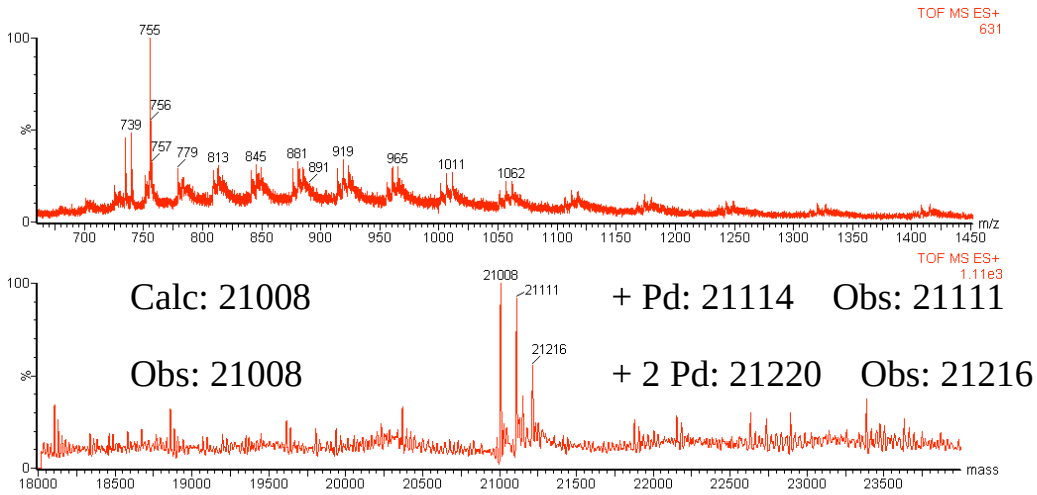
Adducts on Model Protein Np276

Np276 from *Nostoc punctiforme* (20 μL, 1 mg mL^{-1}) was incubated with Pd catalyst **3** stock solution (1.1 nmol, 5 equiv.) for 30 min at 37°C. The mixture was then analysed by LC-MS. This indicated the presence of a mixture of unmodified protein and also protein with one and two naked palladium atoms co-ordinated:

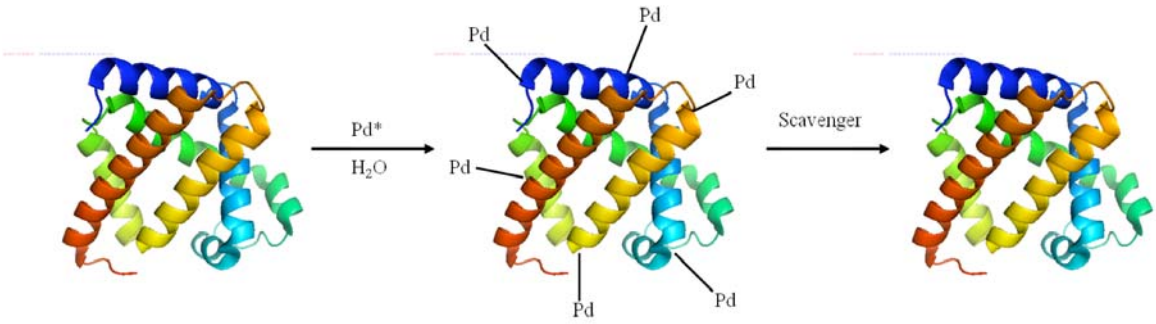
Unmodified:



Protein + 5 eqv. Pd:



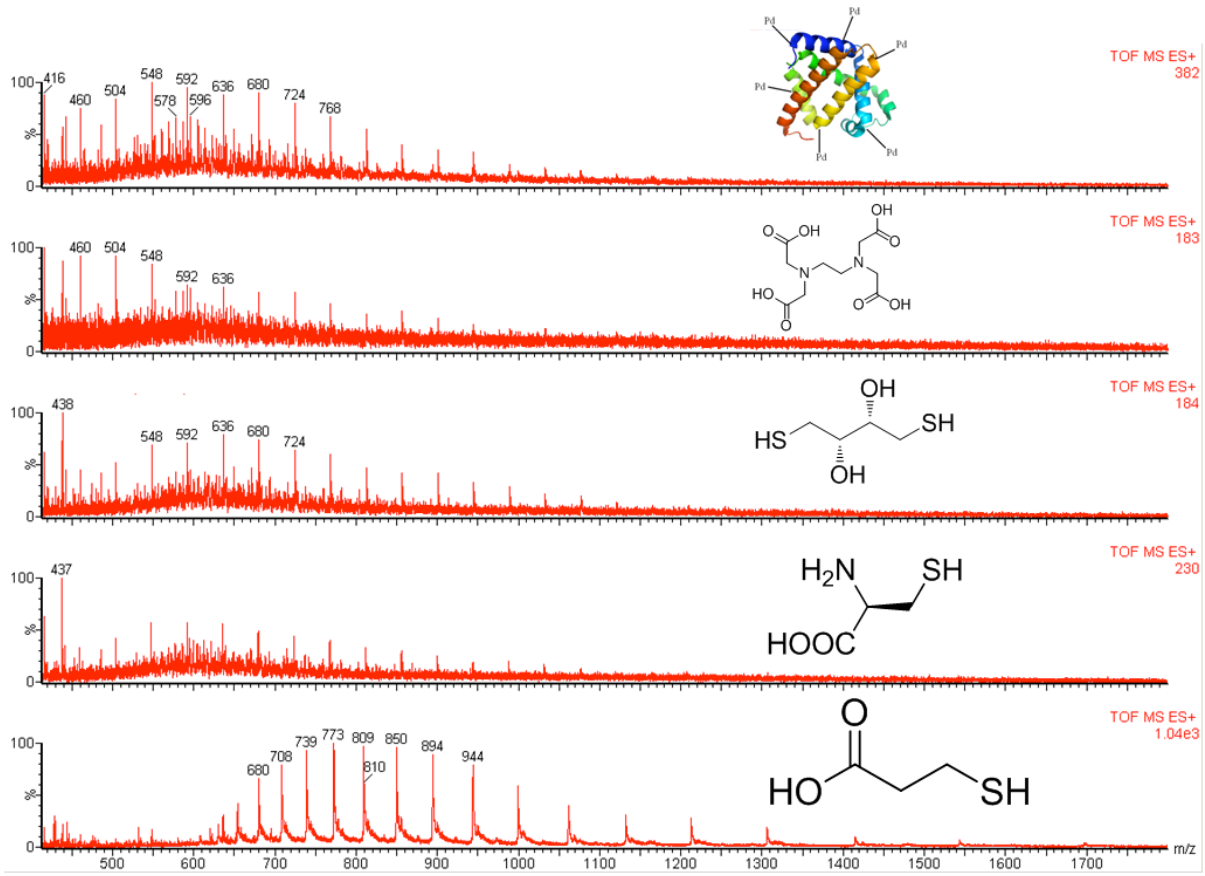
Scavenging Experiments on Myoglobin

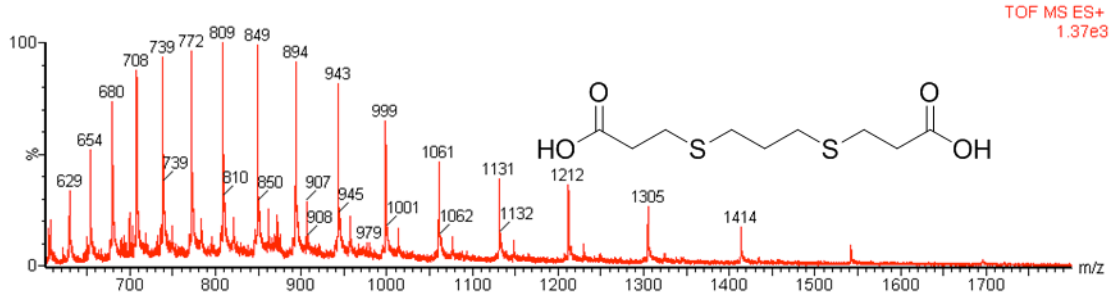


Equine heart myoglobin (50 µL, 1 mg mL-1) was incubated with Pd* stock catalyst (10 µL) for 1 hour at 37ºC. The ligands EDTA, DTT, cysteine, 3-mercaptopropionic acid and diad **2** were then added at varying concentrations as an aqueous solution, and incubated at 37 ºC for a further 30 min. The scavenging ability was then determined by LC-MS analysis. Only 3-mercaptopropionic acid and **2** had any effect on the protein trace, with an optimal concentration of 4.4 µmol/mL (3 equiv. w.r.t Pd) being found for 3-MPrAc, and 9 µmol/mL (6 equiv. w.r.t. Pd) for **2**. Above this concentration, a decrease in protein MS signal was observed

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011

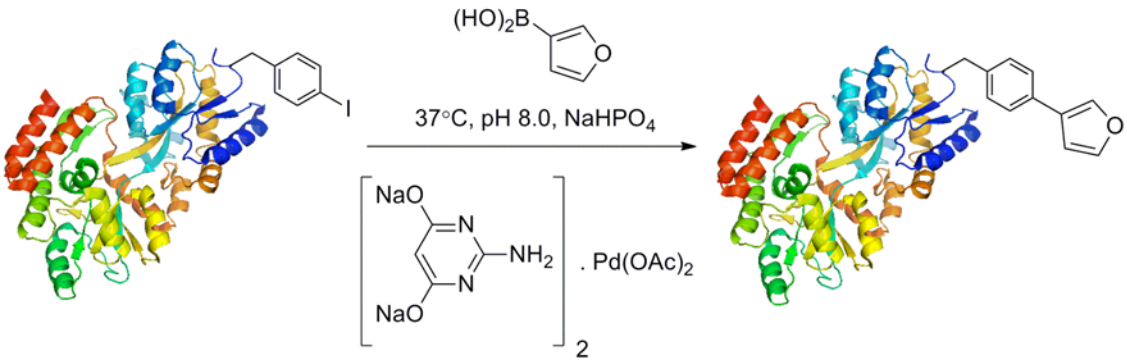
.



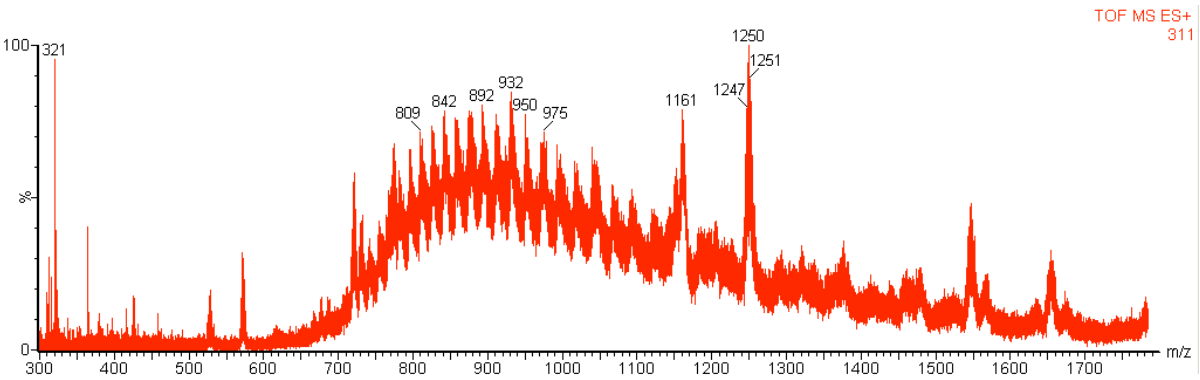


Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011

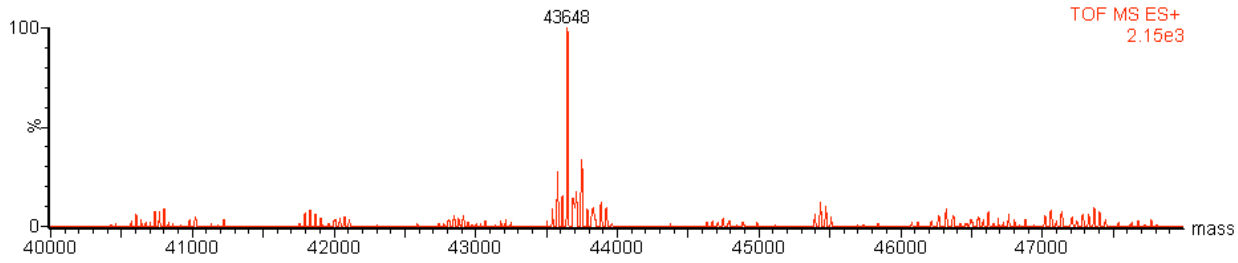
Cross-Coupling on MBP



An aliquot of maltose binding protein MBP.E13· (20 µL, 4.4 mg/mL, 2.0 nmol, 1 equiv) in phosphate buffer (20 mM NaH_2PO_4, 50 mM NaCl, pH 8.0) was added to a 500 µL eppendorf and warmed to 37°C. A stock solution of furan-3-boronic acid (11 mg/mL) and Na_2HPO_4 (16 mg/mL) was prepared in H_2O, and a 15 µL (1.3 µmol/ 680 equiv. boronic acid, 1.3 µmol Na_2HPO_4) aliquot added to the MBP. Stock palladium catalyst (10 µL, 100 nmol, 50 equiv) was then added, and the reaction vortexed to homogenise, then shaken at 37°C. After 2 hr, 20 µL of 3-MPrAc solution (5 µL/mL in water, 1.14 µmol, 570 equiv) was added, and the mixture shaken at 37°C for a further half an hour. The reaction was then analysed by LC-MS which indicated 100% conversion to the desired cross-coupled product. Calc: 43648 Obs: 43648.

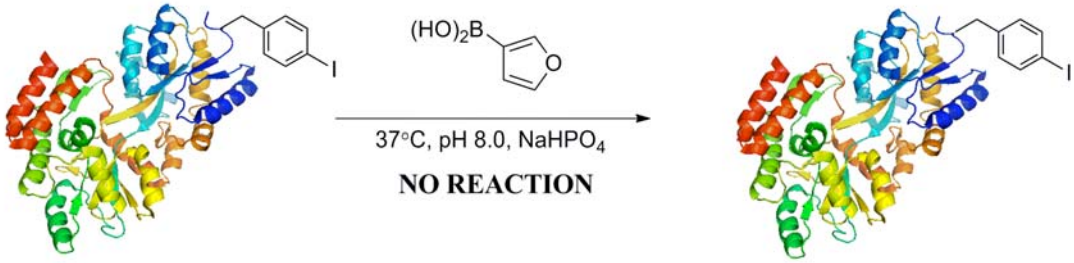


Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011



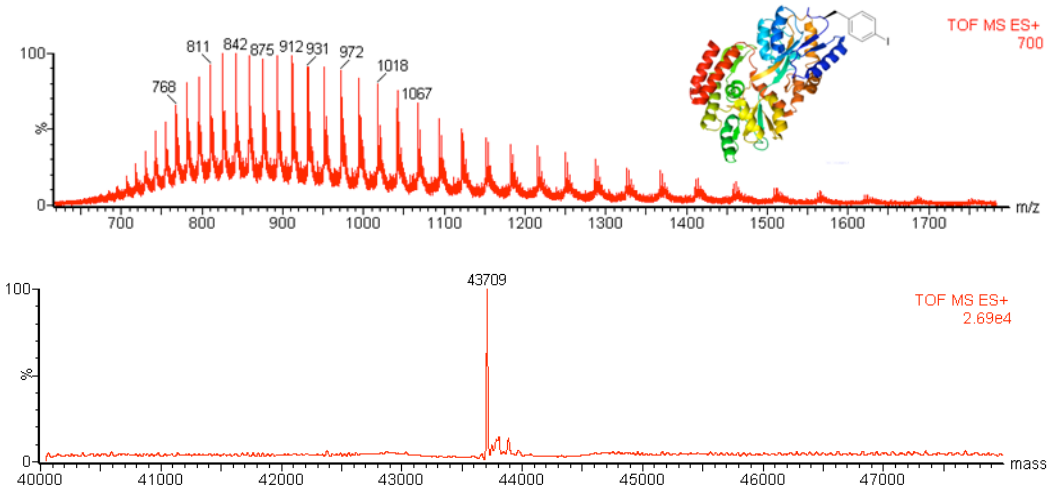
Control reactions

Boronic acid only

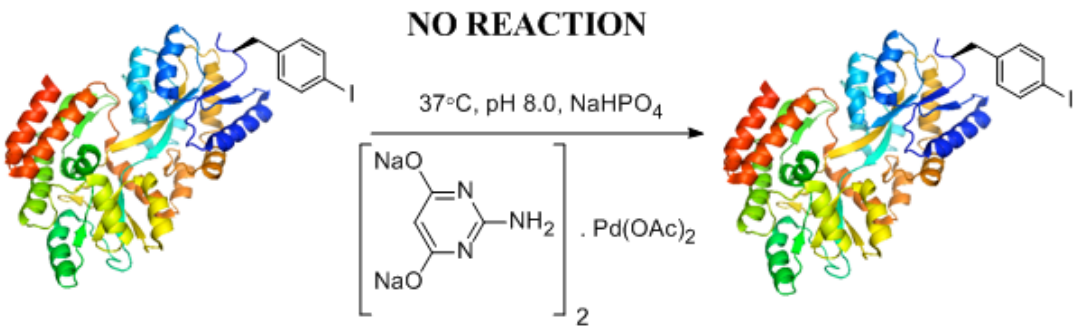


An aliquot of maltose binding protein (20 µL, 4.4 mg/mL, 2.0 nmol, 1 equiv) in phosphate buffer (20 mM NaH_2PO_4, 50 mM NaCl, pH 8.0) was added to a 500 µL eppendorf and warmed to 37°C. A stock solution of furan-3-boronic acid (11 mg/mL) and Na_2HPO_4 (16 mg/mL) was prepared in H_2O, and a 15 µL (1.3 µmol/ 680 equiv. boronic acid, 1.3 µmol Na_2HPO_4) aliquot added to the MBP. The reaction was vortexed to homogenise, then shaken at 37°C. After 2 hr, 20 µL of 3-MPrAc solution (5 µL/mL in water, 1.14 µmol, 570 equiv) was added, and the mixture shaken at 37°C for a further half an hour. The reaction was then analysed by LC-MS which indicated no reaction had occurred (Starting protein observed Calc: 43708 Obs: 43709).

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011

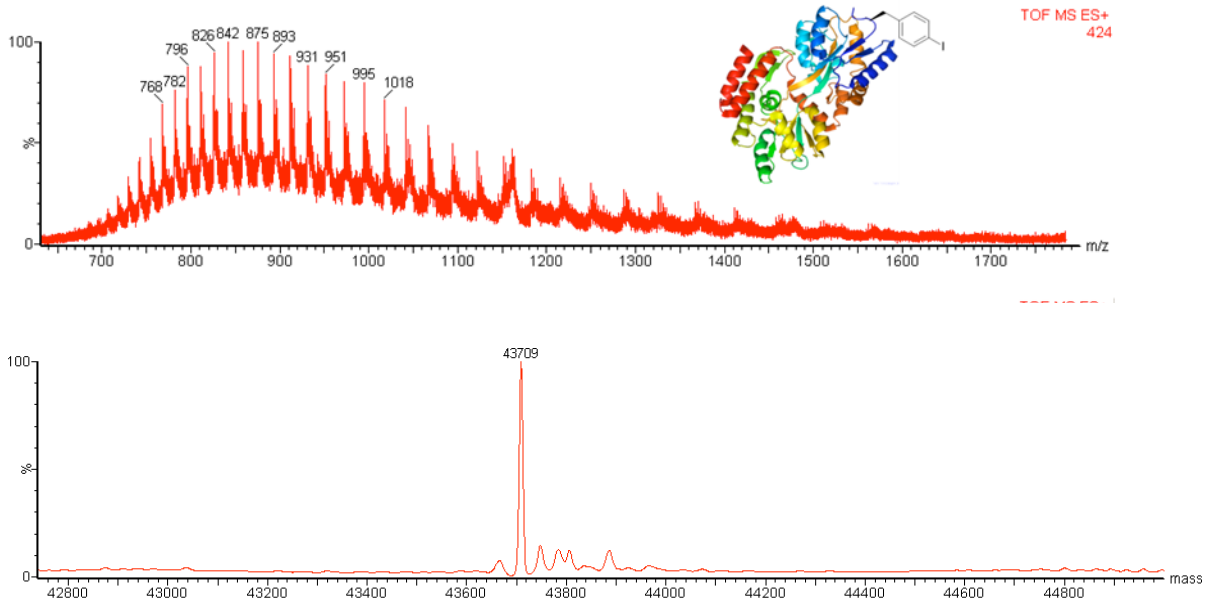


Palladium Only



An aliquot of maltose binding protein (20 µL, 4.4 mg/mL, 2.0 nmol, 1 equiv) in phosphate buffer (20 mM NaH_2PO_4, 50 mM NaCl, pH 8.0) was added to a 500 µL eppendorf and warmed to 37°C. A stock solution of Na_2HPO_4 (16 mg/mL) was prepared in H_2O, and a 15 µL (1.3 µmol Na_2HPO_4) aliquot added to the MBP. Stock palladium catalyst (10 µL, 100 nmol, 50 equiv) was then added, and the reaction vortexed to homogenise, then shaken at 37°C. After 2 hr, 20 µL of 3-MPrAc solution (5 µL/mL in water, 1.14 µmol, 570 equiv) was added, and the mixture shaken at 37°C for a further half an hour. The reaction was then analysed by LC-MS which indicated no reaction had occurred (Starting protein observed Calc: 43708 Obs: 43709).

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2011



Palladium and boronic acid, no scavenger

An aliquot of maltose binding protein (20 µL, 4.4 mg/mL, 2.0 nmol, 1 equiv) in phosphate buffer (20 mM NaH_2PO_4, 50 mM NaCl, pH 8.0) was added to a small eppendorf and warmed to 37°C. A stock solution of furan-3-boronic acid (11 mg/mL) and Na_2HPO_4 (16 mg/mL) was prepared, and a 15 µL (1.3 µmol/ 680 equiv. boronic acid, 1.3 µmol Na_2HPO_4) aliquot added to the MBP. Stock palladium catalyst (10 µL, 100 nmol, 50 equiv) was then added, and the reaction vortexed to homogenise, then shaken at 37°C. The reaction was then analysed by LC-MS. No protein peaks were observed due to co-ordination of the Pd to the protein, generating large number of isoforms.