Table of Contents Page

Materials

Unless otherwise noted, all reagents were obtained from commercial suppliers and used without further purification. Benzene, dimethylformamide (DMF), acetonitrile (ACN), tetrahydrofuran (THF) and methylene chloride (CH_2Cl_2), were obtained from a PureSolv MD solvent purification system by Innovative Technology (solvent deoxygenated by N_2 sparge and dried over alumina). "Extra Dry" grade methanol purchased from Acros was utilized. Deuterated solvents were obtained from Cambridge Isotope labs. Two fluorescent probes (Alexa Fluor® 594 C_5 maleimide and Alexa Fluor® 488 cadaverine, sodium salt) were purchased from Life Technologies™. Silicycle silica gel plates (250 mm, 60 F254) were used for analytical TLC, and preparative chromatography was performed using SiliCycle SiliaFlash silica gel (230-400 mesh). Known compounds including α,α,α',α'-tetramethyl-1,3-benzenedipropionitrile[1], cis-$Rh_2(OAc)_2(OCOCF_3)_2$[2], carbonate **1**[3], 2,6-Bis(2'-pyridyl)-4-pyridone[4], were prepared as previously reported.

Plasmid pET11c-tHisF was provided by the Sterner group of University of Regensburg, Germany.[5] Plasmid pEVOL-pAzF was provided by the Schultz group of the Scripps Research Institute, CA.[6] *E. coli* DH5α and BL21 (DE3) cells were purchased from Invitrogen (Carlsbad, CA). Bacillus amyloliquefaciens strain was purchased from American Type Culture Collection (Mannasas, VA). Nde I, Xho I restriction enzyme, T4 DNA ligase, Taq DNA polymerase and Phusion HF polymerase (Cat# 530S) were purchased from New England Biolabs (Ipswitch, MA). Luria broth (LB) and rich medium (2YT) and Agar media were purchased from Research Products International (Mt. Prospect, IL). Qiagen DNA extraction kit (Cat# 28706) and plasmid isolation kit (Cat# 27106) were purchased from QIAGEN Inc. (Valencia, CA) and used according to the

manufacturer's instructions. DNA purification kit (Zymo, Cat# D4004) was purchased from Zymo research (Irvine, CA) and used as recommended. All genes were confirmed by sequencing at the University of Chicago Comprehensive Cancer Center DNA Sequencing & Genotyping Facility (900 E. 57th Street, Room 1230H, Chicago, IL 60637). Electroporation was carried out on a Bio-Rad MicroPulser using method Ec2. Ni-nitrilotriacetic acid (Ni-NTA) resin and Pierce® BCA Protein Assay Kits (Cat# 23225) were purchased from Fisher Scientific International, Inc. (Hampton, NH), and the manufacturer's instructions were followed when using both products (for Ni-NTA resin, 8 mL resin were used, with buffers delivered by a peristaltic pump at a rate of 1 mL/min, in a 4 °C cold cabinet). Amicon® 10 kD spin filters for centrifugal concentration were purchased from EMD Millipore (Billerica, MA) and used at 4,000 g at 4 °C. PD-10 desalting columns (Cat# 17-0851-01) and Hitrap desalting columns (Cat# 11-0003-29) were purchased from GE Healthcare (Pittsburg, PA).

General procedures

Unless otherwise specified, all reactions were prepared in flame or oven-dried glassware under an inert N_2 atmosphere using either syringe or cannula techniques. TLC plates were visualized using 254 nm ultraviolet light. Flash column chromatography was carried out using Silicycle 230-400 mesh silica gel. ^{1}H and ^{13}C NMR spectra were recorded at 500 MHz and 125 MHz, respectively, on a Bruker DMX-500 or DRX-500 spectrometer, and chemical shifts are reported relative to residual solvent peaks. Chemical shifts are reported in ppm and coupling constants are reported in Hz. Yields determined by HPLC were calculated from internal standards (anisole for cyclopropanation and 1,2,4-

trimethoxybenzene for silane insertion) and reported as the average of two trials set up in parallel. High resolution ESI mass spectra were obtained from the University of Chicago mass spectrometry facility using an Agilent Technologies 6224 TOF LC/MS. MALDI-MS spectra were recorded on AB SCIEX Voyager-DE PRO MALDI-Tof system. Amicon® 15 mL 10 kD cutoff centrifugal filter was used to concentrate or wash protein solutions. Protein concentrations were measured using the Pierce® BCA Protein Assay Kit and protein stocks were then stored at -20 °C until use.

Cloning, expression and protein purification

Standard cloning procedures and site directed mutagenesis:

A gene encoding the cyclase subunit (tHisF) of the imidazole glycerol phosphate synthase enzyme complex from Thermotoga maritima was amplified from pET11c-tHisF20 by PCR using gene specific primers containing NdeI (forward) and XhoI (reverse) restriction sites. The gene was cloned into the NdeI and XhoI sites of pET22b so that scaffolds would be expressed with a C-terminal hexa-histidine tag for Ni-NTA affinity chromatography. Amber mutations were introduced into the tHisF gene at positions L50, G176 and I199 by site directed overlap extension (SOE) PCR[7]. For each mutation, two separate polymerase chain reactions were performed, each using a perfectly complementary flanking primer at the 5' and 3' end of the sequence and a mutagenic primer. The PCR conditions were as follows: Phusion HF buffer 1x, 0.2 mM dNTPs each, 0.5 μM forward primer, 0.5 μM reverse primer, 0.02 U/μL Phusion polymerase and 1 ng/mL template plasmid. The resulting two overlapping fragments that

contained the base pair substitution were then assembled in a second PCR using the flanking primers resulting in the full-length mutated gene. For phytase gene cloning, genomic DNA was isolated from *Bacillus amyloliquefaciens* (ATCC#23350) and genomic PCR was done using gene specific primers having above restriction sites on the flanking region. Genomic PCR was performed in one step using the same conditions as above, except template (genomic DNA) concentration was increased to 500 ng/mL. An amber stop codon (Y104Az) was introduced for ArM formation, and three alanine mutations (N99A, N100A, D101A) were introduced to improve access to the site of Az mutation within the scaffold. Nucleotide sequences for the all the primers are summarized in Table S1.

#	Primer name	Primer sequence
1	T7 for	5'-GCG AAA TTA ATA CGA CTC ACT ATA-3'
2	T7 rev	5'-TTA TGC TAG TTA TTG CTC AGC GG-3'
3	L50Az for	5'-GAA CTC GTT TTT TAG GAT ATC ACC GCG-3'
4	L50Az rev	5'-CGC GGT GAT ATC CTA AAA AAC GAG TTC-3'
5	G176Az for	5'-AGT ATC GAC AGA TAG GGC ACA AAA TCG-3'
6	G176Az rev	5'-CGA TTT TGT GCC CTA TCT GTC GAT ACT-3'
7	I199Az for	5'-ACA CTT CCC ATC TAG GCT TCC GGT GGT-3'
8	I199Az rev	5'-ACC ACC GGA AGC CTA GAT GGG AAG TGT-3'
9	C9ala for	5'-AGA ATA ATC GCG GCG CTC GAT GTG AAA-3'
10	C9ala rev	5'-TTT CAC ATC GAG CGC CGC GAT TAT TCT-3'
11	K243C for	5'-GAG TAC CTC AAA TGC CAC GGA GTG AAC-3'
12	K243C rev	5'-GTT CAC TCC GTG GCA TTT GAG GTA CTC-3'
13	D174C for	5'-C TCACC AGT ATC TGC AGA GAC GGC-3'

14	D174C rev	5’-GCC GTC TCT GCA GAT ACT GGTGA G-3’
15	phyA104 for	5’-CCTGCGATTTAGCTGGACCCCAAG-3’
16	phyA104 rev	5’-CTTGGGGTCCAGCTAAATCGCAGG-3’
17	Phy for	5’-GCAACATATGTCTGATCCTTATCATTTTACCG-3’
18	Phy rev	5’-AGCACTCGAGTTATTTTCCGCTTCTGTCAGTCA-3’

Table S1. Nucleotide sequences for the primers

PCR amplified fragments and plasmid vector pET22b were restriction digest with Nde I and Xho I enzymes in recommended buffer at 37 °C for 2 hours. Digested DNA was cleaned by agarose gel extraction using commercial kit before ligation. Ligation was set-up with a molar ratio of 1:3 (plasmid: insert) in 10 µL reaction mix. Typically a ligase reaction mix had 1 ng/L digested plasmid vector, 9 ng/mL of the insert, 1 µL 10X ligase buffer and 1 U/mL ligase. Reaction mix was incubated at 16 °C overnight, cleaned using DNA purification kits and transformed into *E. coli* DH5 cells. Cells were spread on LB ampicillin plates (6.25 g LB powder mix, 4 g agar, 250 mL DDI water, 0.1 mg/mL ampicillin) before recovering in SOC medium for 1 hour at 37 °C. Plates were incubated at 37 °C overnight; individual colonies that appeared next day were tested for gene fragments by colony PCR. Clones that showed amplification for desired fragments were inoculated on LB broth having 0.1 mg/mL ampicillin and grown overnight at 37 °C, 250 rpm. Recombinant plasmid from these overnight grown cultures were isolated using kit and given for sequencing. Plasmid sequencing was done by the U Chicago sequencing facility staff and T7 for and T7 rev primers were used for sequencing reactions. For FRET experiments, double mutants were created by replacing L50 with Az (L50Az) and C9 with A (C9A), and this construct was used as template to introduce cysteine residues at different positions (D174C and K243C).

Expression and purification protocol:

pET22b-tHisF (or mutants thereof) and pEVOL-pAzF were co-transformed into electrocompetent *E. coli* BL21 (DE3)[8], these cells were allowed to recover in SOC medium (37 °C, 50 min), then plated onto LB amp+Cm agar plates (6.25 g LB powder mix, 4 g agar, 250 mL DDI water, 0.1 mg/mL ampicillin, 0.05 mg/mL chloramphenicol), and incubated at 37 °C for 16 h. Several colonies appeared on overnight-incubated plates; a single colony from this plate was inoculated in 5 mL 2YT medium having antibiotics with the same concentrations as above. The culture was incubated overnight at 30 °C with constant shaking at 250 rpm. On the following day, 3 mL of the overnight cultures was used to inoculate 300 mL of fresh 2YT media having the same antibiotics, in 1 L Erlenmeyer flask. The culture was incubated at 30 °C, 250 rpm, and protein expression was induced by adding 1mM IPTG, 2mM 4-Azido-phenyl alanine and 1% (w/v) L-arabinose when OD_{600} reached 1. The induced culture was allowed to grow for 12 hours, and then the cells were harvested by centrifugation at 4 °C, 3000 x g for 20 minutes. Cell pellets were re-suspended in 30 mL PBS (pH 7.5) and sonicated (40 amplitude, 30 second burst, 10 minute total process). Lysed culture was then clarified at 16000 x g, 4 °C for 30 minutes and supernatant thus obtained was purified by Ni-NTA resin using manufacturer's instructions. Purified protein was buffer exchanged to 10 mM Tris (pH 7.5) and measured by Pierce® BCA Protein Assay Kit as recommended.

Cofactor and Probe Synthesis

Compound **2**: To a 50 mL glass bomb were added α, α, α', α'-tetramethyl-1,3-benzenedipropionitrile[1] (3.365 g, 14 mmol), $[Ir(COD)(OMe)]_2$ (30.5 mg, 0.046 mmol,

0.003 equiv), 4,4'-di-tert-butylbipyridine (24.7 mg, 0.092 mmol, 0.006 equiv), and B_2pin_2 (2.489 g, 9.8 mmol, 0.7 equiv)[9]. The bomb was evacuated and refilled with N_2 three times. Under a positive flow of N_2, THF (22.0 mL) was added. The bomb was then sealed and heated in an 80 °C oil bath for 48 h. The solvent was removed under reduced pressure, and the resulting residue was purified by flash chromatography (silica gel, 85:15 hexanes/EtOAc) to yield a crude product. The crude product was dissolved in 100 mL MeOH, and hydrogen peroxide solution (30 % (w/w) in H_2O, 19 mL, 167.6 mmol, 12.0 equiv) was added. The resulting solution was stirred at room temperature for 2 h, and the solvent was evaporated. Purification by flash chromatography (silica gel, 3:1 hexanes/EtOAc) afforded phenol 2a as a white solid (2.401 g, 68%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.72 (d, J = 1.0 Hz, 2H), 6.71 (d, J = 1.0 Hz, 1H), 2.74 (s, 4H), 1.35 (s, 12H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 156.0, 137.5, 124.9, 124.5, 116.3, 46.5, 33.6, 26.7. HRMS-ESI (m/z): calcd for $C_{16}H_{19}N_2O$ $[M-H]^-$: 255.1492, found: 255.1504.

Phenol 2a (2.401 g, 9.37 mmol) and KOH (3.152 g, 56.2 mmol, 6.0 equiv) were dissolved in 16 mL ethylene glycol[1]. The resulting solution was heated at 180 °C for 6 h. After cooling the reaction to room temperature, the contents were partitioned between 24 mL of $CHCl_3$ and 24 mL of H_2O. The aqueous layer was collected, acidified with 6 M aqueous HCl to pH 1, and extracted with EtOAc (80 mL x 3). The organic layer was washed successively with 50 mL of H_2O and 50 mL of brine, dried over Na_2SO_4, and concentrated under reduced pressure. Purification by flash chromatography (silica gel, 19:1 CH_2Cl_2/MeOH) afforded 2b as white solid (2.185 g, 79%). ^{1}H NMR (500 MHz, CD_3OD) δ 6.47 (m, 3H), 2.74 (s, 4H), 1.13 (s, 12H). ^{13}C NMR (125 MHz, CD_3OD) δ

181.7, 157.8, 140.4, 125.0, 116.4, 47.2, 44.5, 25.8. HRMS-ESI (m/z): calcd for $C_{16}H_{22}O_5Na$ $[M+Na]^+$: 317.1365, found: 317.1357.

Phenol diacid 2b (48 mg, 0.16 mmol), cis-$Rh_2(OAc)_2(OCOCF_3)_2$[2] (100 mg, 0.16 mmol, 1.0 equiv) and potassium carbonate (47 mg, 0.32 mmol, 2.0 equiv) were added into a 25 mL round-bottom flask. 8 mL of THF was added, and the resulting suspension was heated at 50 °C for 3 h. The solvent was evaporated, and purification by flash chromatography (silica gel, 4:1 benzene/acetonitrile) yielded 2 as purple solid (77 mg, 80%, bis-acetonitrile adduct)[10]. ^{1}H NMR (500 MHz, CD_3CN) δ 6.37 (s, 1H), 6.33 (s, 2H), 2.52 (s, 4H), 2.00 (s, 6H), 1.81 (s, 6H), 0.96 (s, 12H). ^{13}C NMR (125 MHz, CD_3CN) δ 197.6, 191.7, 156.4, 140.5, 123.9, 115.6, 47.5, 46.4, 26.0, 23.5. HRMS-ESI (m/z): calcd for $C_{20}H_{26}ClO_9Rh_2$ $[M+Cl]^-$: 650.9376, found: 650.9395.

Cofactor **3**: Complex 2 (54 mg, 0.088 mmol) and sodium hydride (60 % dispersion in mineral oil, 3.2 mg, 0.08 mmol, 0.9 equiv) were added into a 10 mL round-bottom flask. 3 mL THF was added and, the resulting suspension was heated at 50 °C for 1 h and cooled to room temperature. A solution of carbonate **1**[3] (31 mg, 0.098 mmol, 1.1 equiv) in 1 mL THF was added. The reaction was stirred at room temperature for 1 h, and the solvent was evaporated. Purification by flash chromatography (silica gel, 4:1 benzene/acetonitrile) yielded cofactor 3 as a dark green solid (50mg, 72%). ^{1}H NMR (500 MHz, CD_3CN) δ 6.73 (d, J = 1.5 Hz, 1H), 6.68 (d, J = 1.5 Hz, 2H), 4.14 (d, J = 6.9 Hz, 2H), 2.58 (s, 4H), 2.44 – 2.22 (m, 4H), 2.20 (m, 2H), 1.49 (m, 2H), 0.92 (s, 6H), 0.84 (m, 3H). ^{13}C NMR (125 MHz, CD_3CN) δ 197.6, 192.0, 154.7, 150.9, 140.7, 129.6, 121.4, 99.0, 73.6, 47.1, 46.6, 33.5, 25.9, 23.9, 23.7, 23.5, 21.7. HRMS-ESI (m/z): calcd for $C_{31}H_{42}NO_{11}Rh_2$ $[M+NH_4]^+$: 810.0868, found: 810.0831.

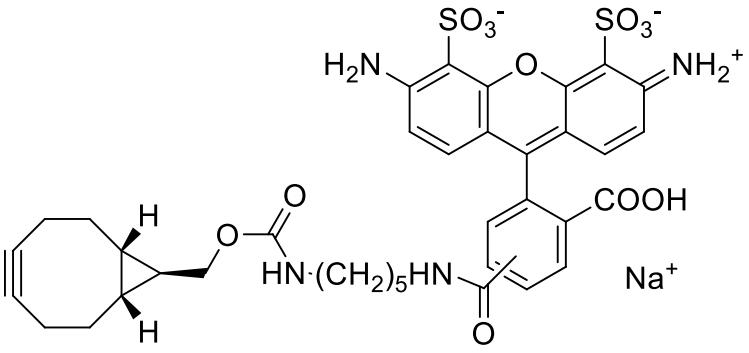
Compound **5**: 2,6-bis(2'-pyridyl)-4-pyridone[4] (39 mg, 0.16 mmol) and 1 (48 mg, 0.15 mmol, 0.95 equiv) were dissolved in 12 mL THF. Cesium carbonate (73 mg, 0.22 mmol, 1.5 equiv) was added, and the suspension was stirred at room temperature overnight. The mixture was filtered through celite and the filtrate was concentrated down to yellow oil. The crude was purified by flash chromatography (silica gel, 49:1 CH_2Cl_2/MeOH)) to afford compound 5 (32 mg, 48 %) as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.80 (d, *J* = 4.3 Hz, 2H), 8.64 (d, *J* = 8.0 Hz, 2H), 8.34 (s, 2H), 7.95 (t, *J* = 7.7 Hz, 2H), 7.47 – 7.40 (m, 2H), 4.24 (d, *J* = 6.5 Hz, 2H), 2.47 (m, 2H), 2.32 (m, 2H), 2.18 (m, 2H), 1.42 (m, 2H), 0.87 (m, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 160.2, 157.9, 155.4, 152.7, 149.4, 137.0, 124.3, 121.4, 113.6, 98.9, 73.9, 33.3, 31.1, 23.4, 23.2, 21.4. HRMS-ESI (m/z): calcd for $C_{26}H_{24}N_3O_3$ $[M+H]^+$: 426.1818, found: 426.1812.

Cofactor **6**: 1 (29 mg, 0.068 mmol) was dissolved in 1 mL dichloromethane. A solution of $CuCl_2$ (9.2 mg, 0.068 mmol, 1.0 equiv) in 1 mL of DCM/methanol 1:1 was added dropwise. The resulting solution was stirred at room temperature for 4 h, and the solvent was evaporated. The residue was washed with cold dichloromethane and cold methanol and dried to yield 6 (32 mg, 84%) as a paramagnetic, blue solid[11]. HRMS-ESI (m/z): calcd for $C_{26}H_{23}ClCuN_3O_3$ $[M-Cl]^+$: 523.0724, found: 523.0722.

Cofactor **7**: 5 (21.2 mg, 0.05 mmol) was dissolved in 1 mL THF. A solution of $MnCl_2$•$4H_2O$ (49.5 mg, 0.25 mmol, 5.0 equiv) in 5 mL THF was added in one portion. A yellow precipitate formed immediately after addition. The resulting suspension was stirred for 15 minutes and filtered. The solid was washed with copious THF and dried under vacuum to afford 7 (10 mg, 36%) as a paramagnetic, light yellow solid[12]. HRMS-ESI (m/z): calcd for $C_{26}H_{23}ClMnN_3O_3$ $[M-Cl]^+$: 515.0809, found: 515.0792.

Probe **8**: In a 1.5 mL microcentrifuge tube, Alexa Fluor® 488 cadaverine, sodium salt (0.5 mg, 0.78 µmol), 1 (0.50 mg, 1.56 µmol, 2 equiv), and N-ethyldiisopropylamine (0.7 µL, 4.02 µmol, 5.2 equiv) were dissolved in 1 mL DMF. The tube was sealed with aluminum foil and shaken at room temperature for 48 h, and the solvent was evaporated. The residue was dissolved in 9:1 water/acetonitrile, purified by reversed-phase HPLC using water and acetonitrile, and lyophilized to a fine powder. See supporting information for HPLC trace. HRMS-ESI (m/z): calcd for $C_{37}H_{37}N_4O_{12}S_2$ $[M]^+$: 793.1850, found: 793.1890.

HPLC trace of probe ***8***



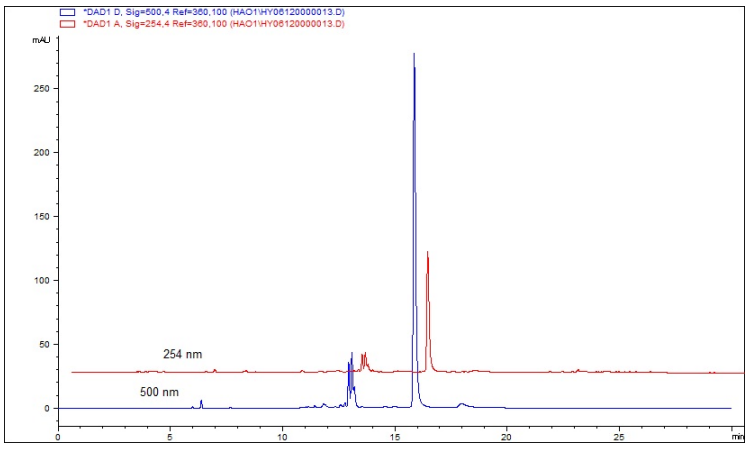
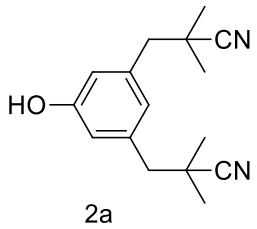


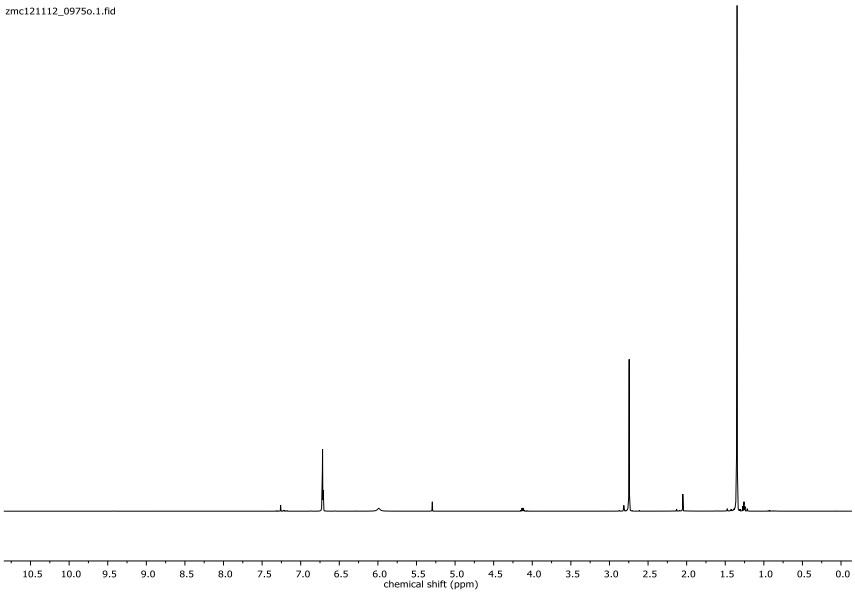
Figure S1**.** Analytical HPLC trace of purified **9**. Method: 0% to 50% B from 0-20 min, 50% to 100% B from 20-25 min. Solvent A: water + 0.1% TFA; solvent B: acetonitrile.

NMR spectra

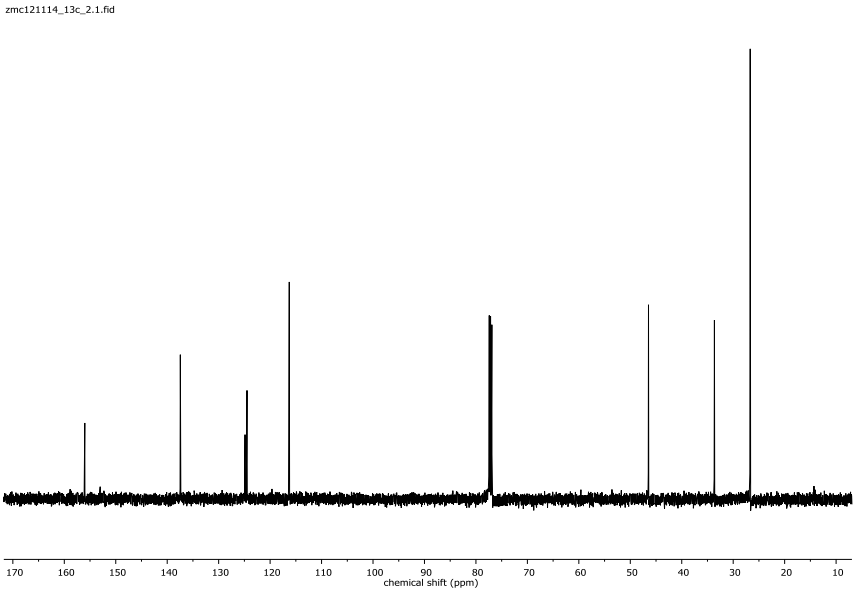


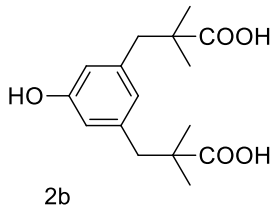
2a

^{1}HNMR of 2a



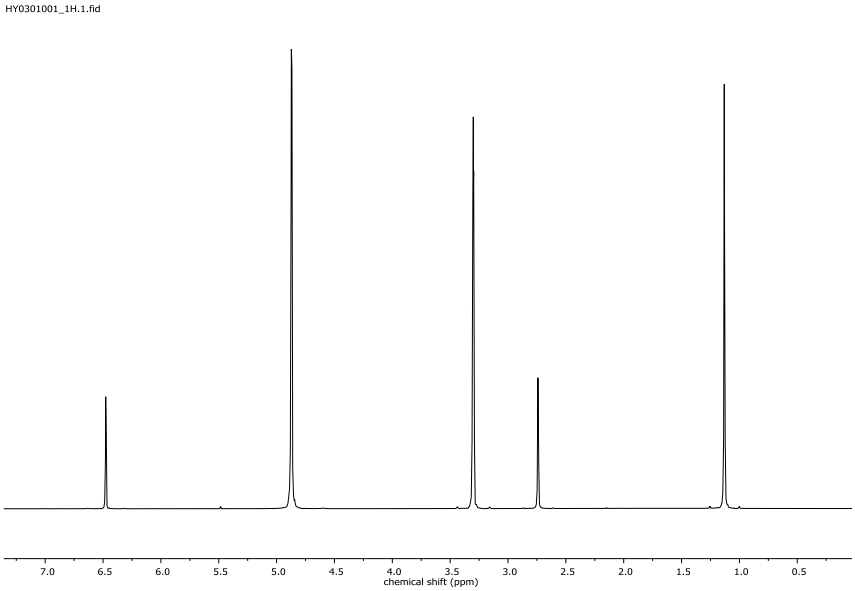
^{13}CNMR of 2a



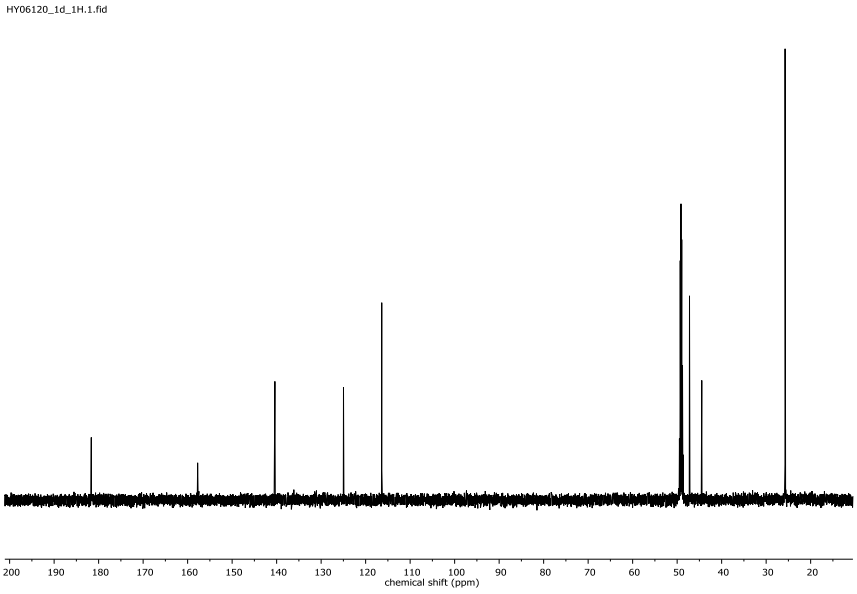


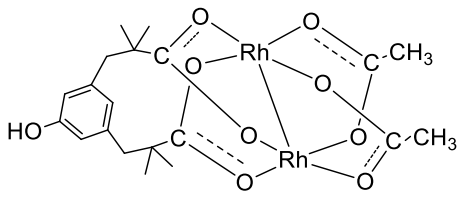
2b

^{1}HNMR of 2b



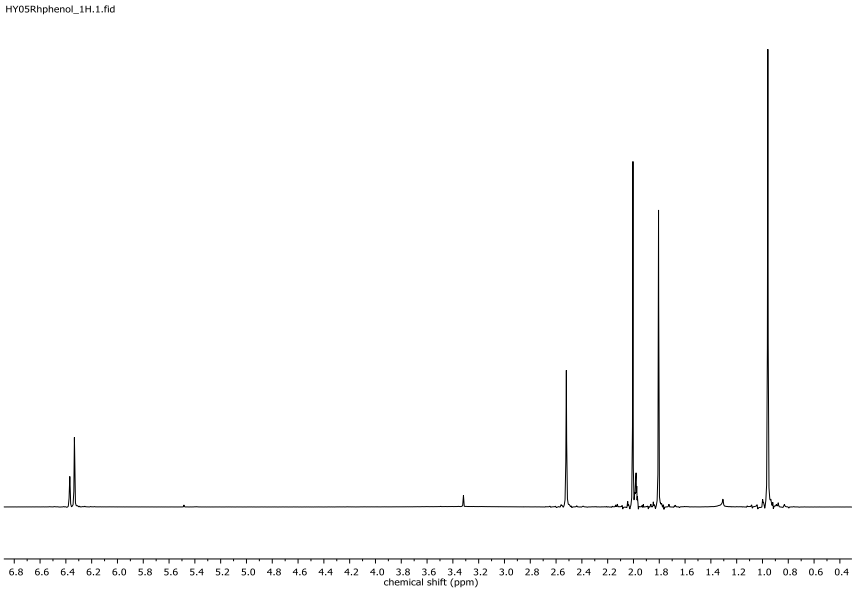
^{13}CNMR of 2b



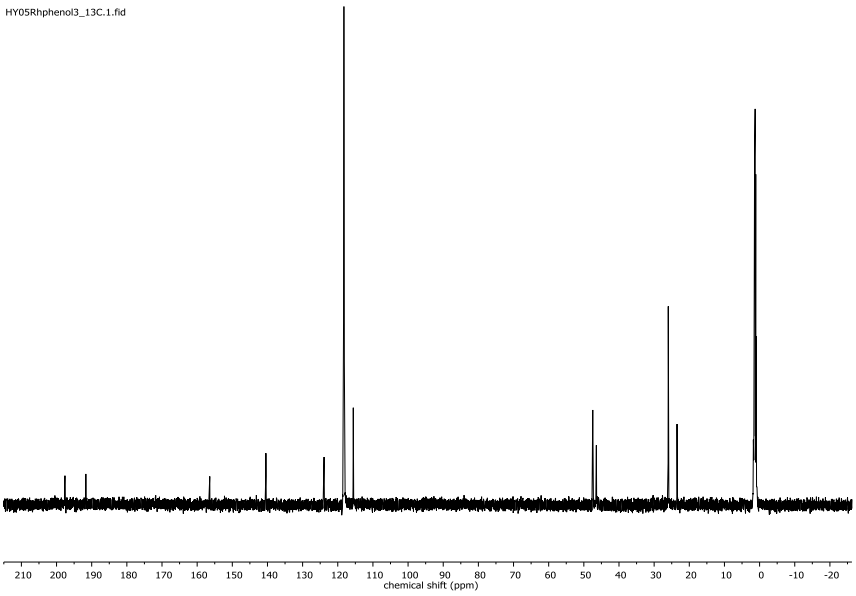


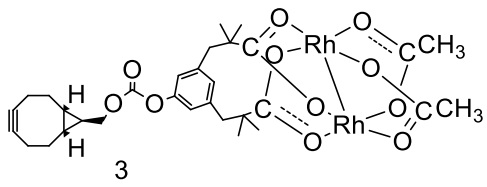
2

^{1}HNMR of 2



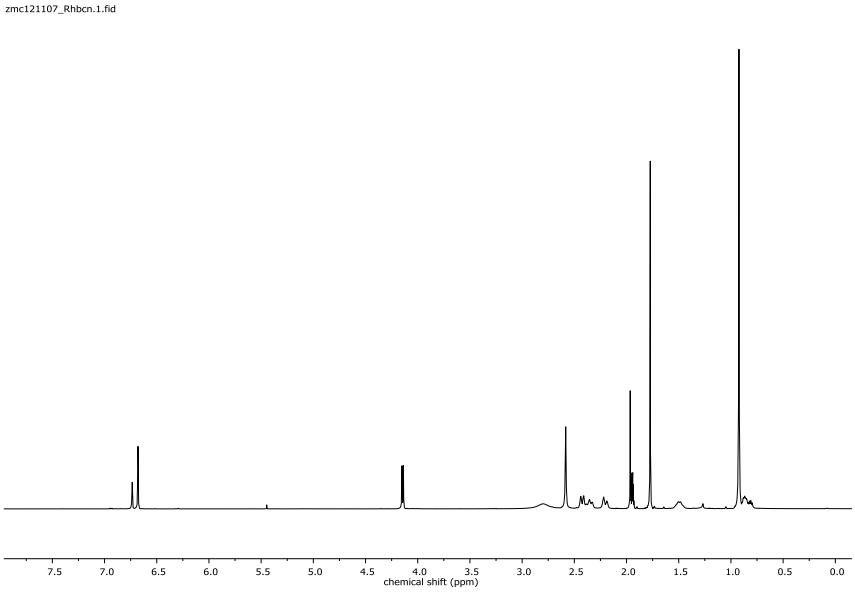
^{13}CNMR of 2



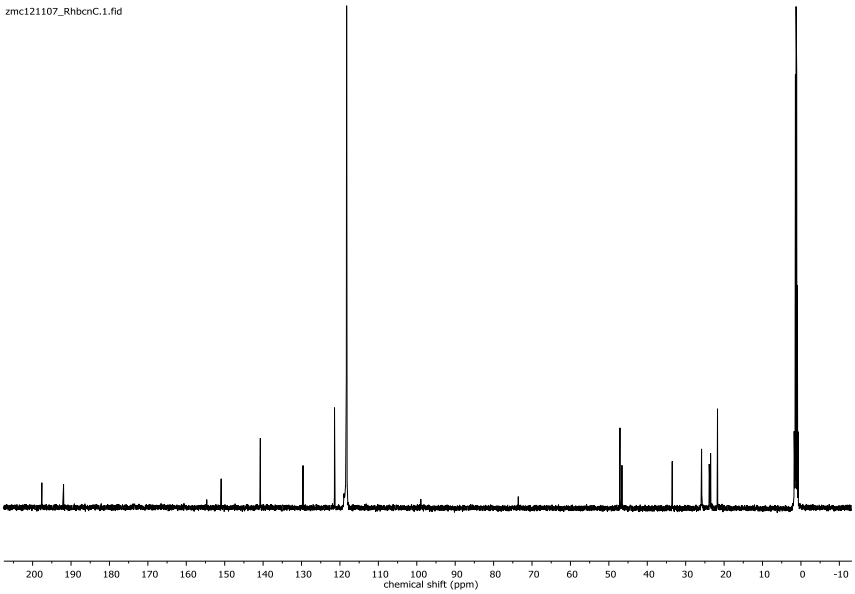


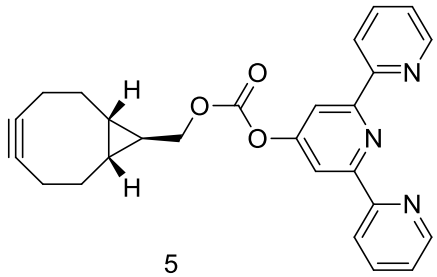
3

^{1}HNMR of 3



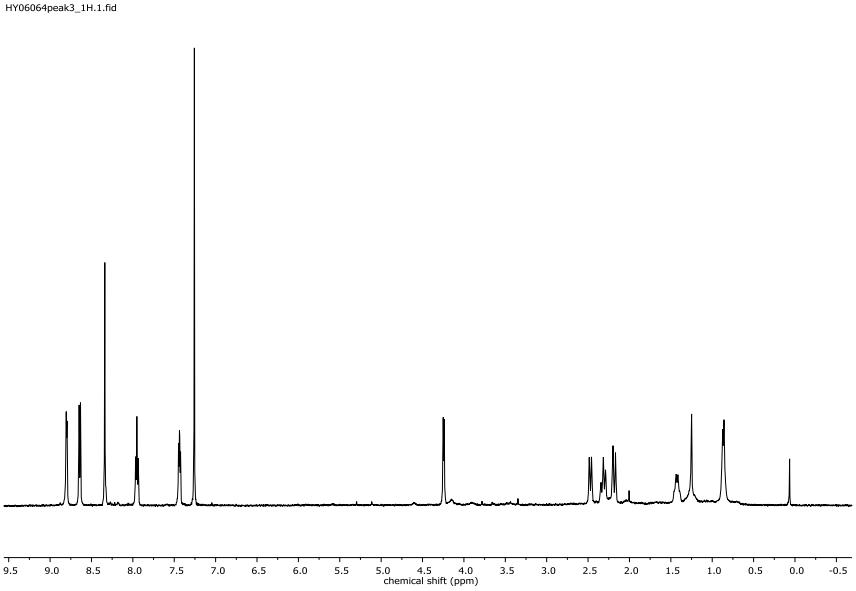
^{13}CNMR of 3



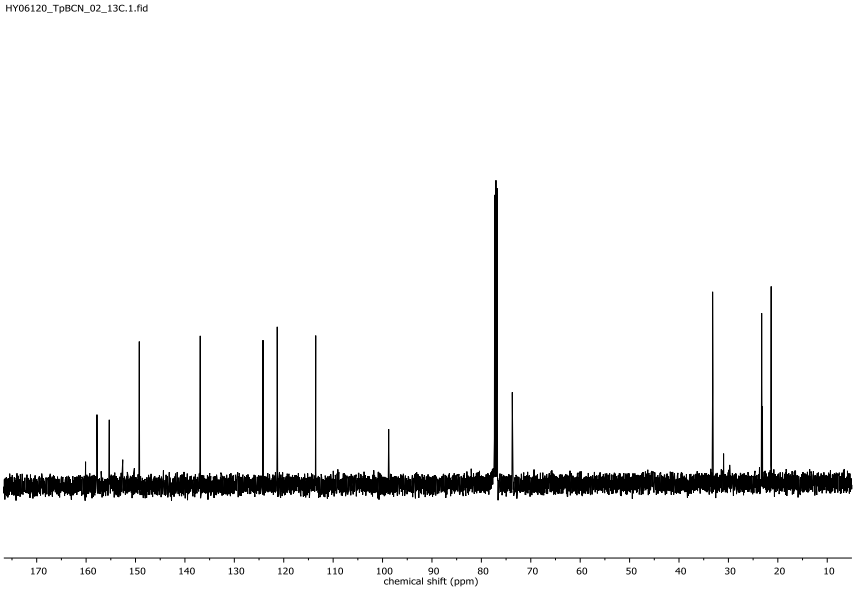


5

^{1}HNMR of 5



^{13}CNMR of 5



ArM formation

*Bioconjugation of **3**:*

A solution of tHisF (480 µL, 75 µM tHisF in 10 mM Tris, pH 7.5) and a solution of cofactor **3** (120 µL, 1.188 mg/mL in acetonitrile, 5.0 equiv) were added to an microcentrifuge tube and shaken in the dark at 4 °C for 48 h. The final concentrations were: 60 µM tHisF-Az50, 300 µM **3**, 20 vol% acetonitrile. Due to the high hydrophobicity of cofactor **3**, coelution of **3** with tHisF in size exclusion chromatography or anion exchange chromatography was observed (other cofactors **6**, **7** were removed by simple gel filtration due to their relatively good hydrophilicity). Hence, the reaction was purified by reversed-phase preparative HPLC and exchanged to 10 mM Tris (pH 7.5) with centrifugal filters. The removal of excess cofactor **3** was confirmed by analytical HPLC analysis, the concentration of product was determined with Pierce® BCA Protein Assay Kit, and the conversion was estimated by MALDI-MS analysis. Because apparent in-situ bioconjugation under MALDI-MS conditions was observed, a modified sinapinic acid matrix was used to quench any unreacted azide (10 mg/mL of sinapinic acid and 3 mg/mL of cyclooctaalkyne alcohol in 50:50 water/acetonitrile, 0.1 % TFA final conc.).

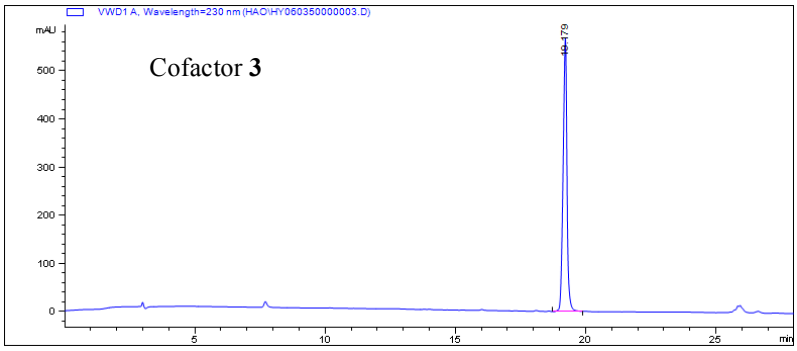
*Bioconjugation of **6** or **7**:*

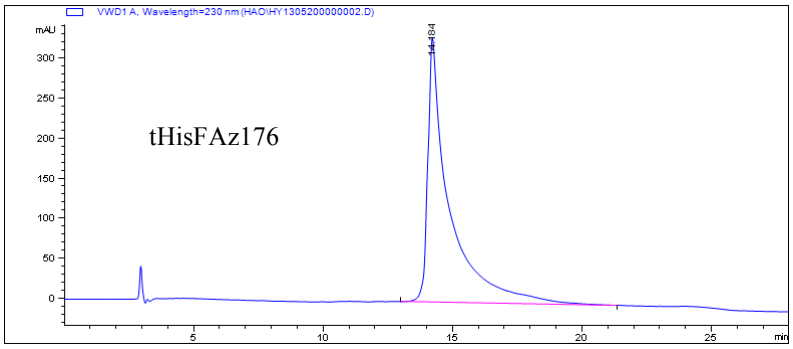
General procedure for the conjugation of cofactor **6** or **7**: A tHisF solution in 10 mM Tris (pH 7.5) and a solution of cofactor in methanol were mixed and incubated in the dark at 4 °C for 24 h. The final concentrations were: 60 µM tHisF, 300 µM, 15 vol% methanol. The reaction was desalted by gel-filtration with PD-10 desalting columns to remove

excess cofactors. The purity of metalloenzyme was determined by analytic HPLC analysis. Protein concentrations were measured using the Pierce® BCA Protein Assay Kit. The conversion was estimated by MALDI-MS analysis with a modified sinapinic acid matrix described above.

HPLC monitoring and analysis of RhBCN bioconjugation:

Analytic HPLC runs were performed on an Agilent 1100 Series HPLC system using Vydac 218TP54 column (C18, 300 Å, 5 µm, 4.6 mm i.d. x 250 mm), with a flow rate of 1.0 mL/min and detection wavelength set at 230 nm. The following gradient was used: 20 % to 64 % B from 0-15 min, 64 % from 15-20 min, 64 % to 80 % from 20-22 min, 80 % from 22-25 min, 80 % to 20 % from 25-28 min (solvent A: water containing 0.1% TFA; solvent B: CH_3CN).





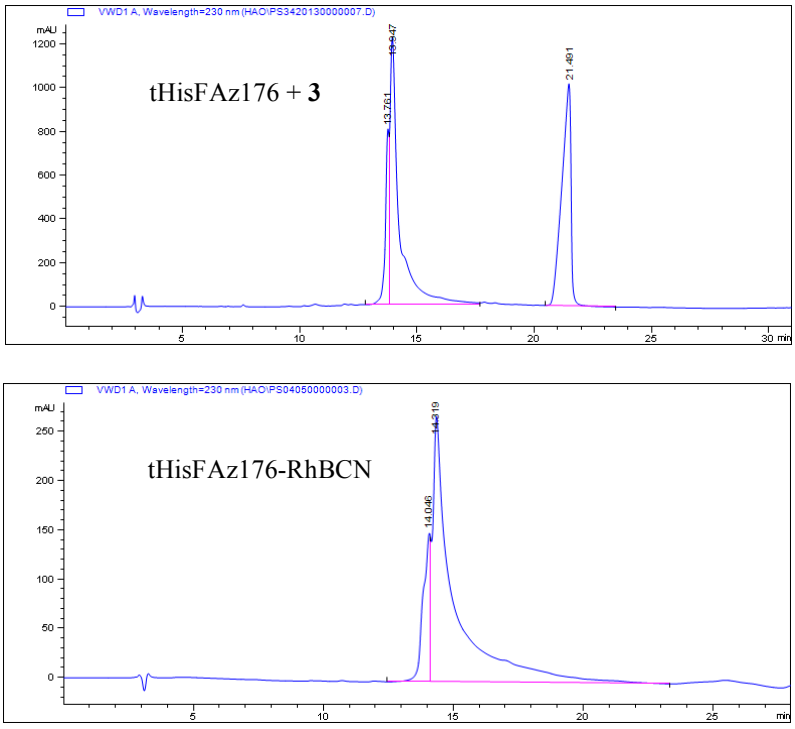
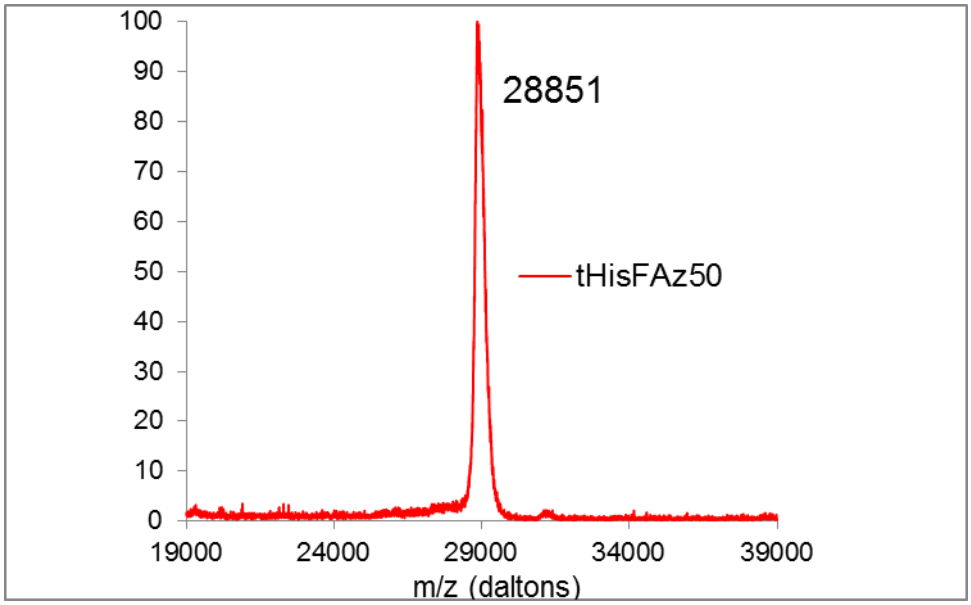
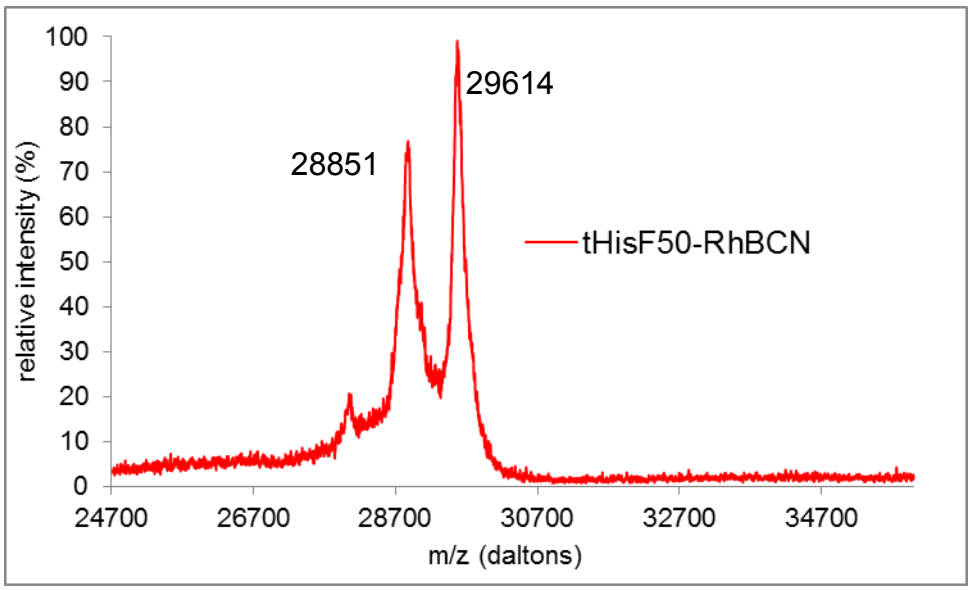
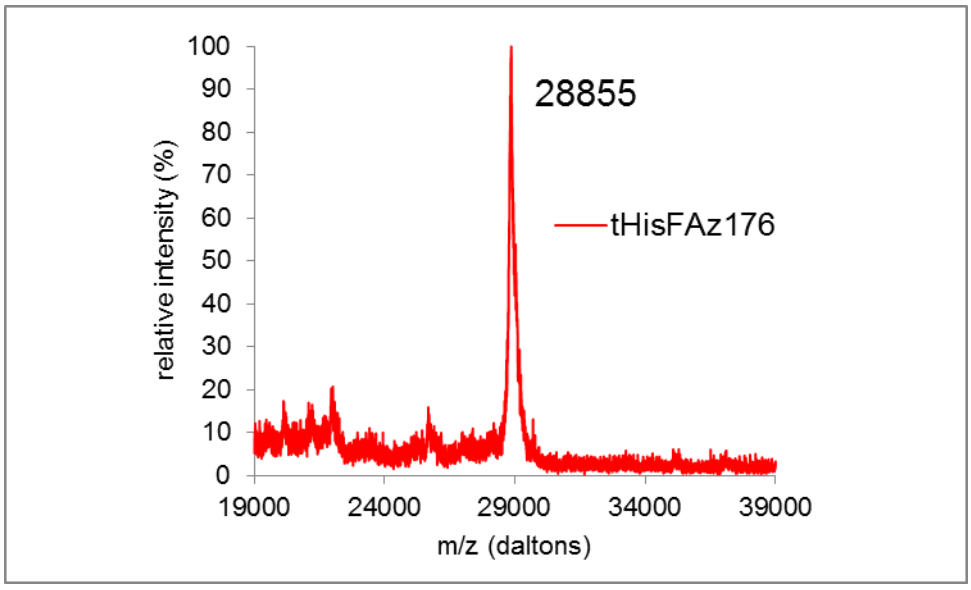


Figure S2. Comparison of analytical HPLC runs of cofactor **3**, tHisF protein, crude bioconjugation reaction, and purified tHisF_RhBCN. The residual amount of **3** in tHisF_RhBCN is not detectable by HPLC (Slight variations in retention time of protein or cofactors were observed in HPLC).

Representative MALDI-MS spectra for ArM:




100
90
80
70
60
50
40
30
20
10
0
relative intensity (%)
29614
28851
tHisF50-RhBCN
24700
26700
28700
30700
32700
34700
m/z (daltons)


100
90
80
70
60
50
40
30
20
10
0
relative intensity (%)
28855
tHisFAz176
19000
24000
29000
34000
39000
m/z (daltons)

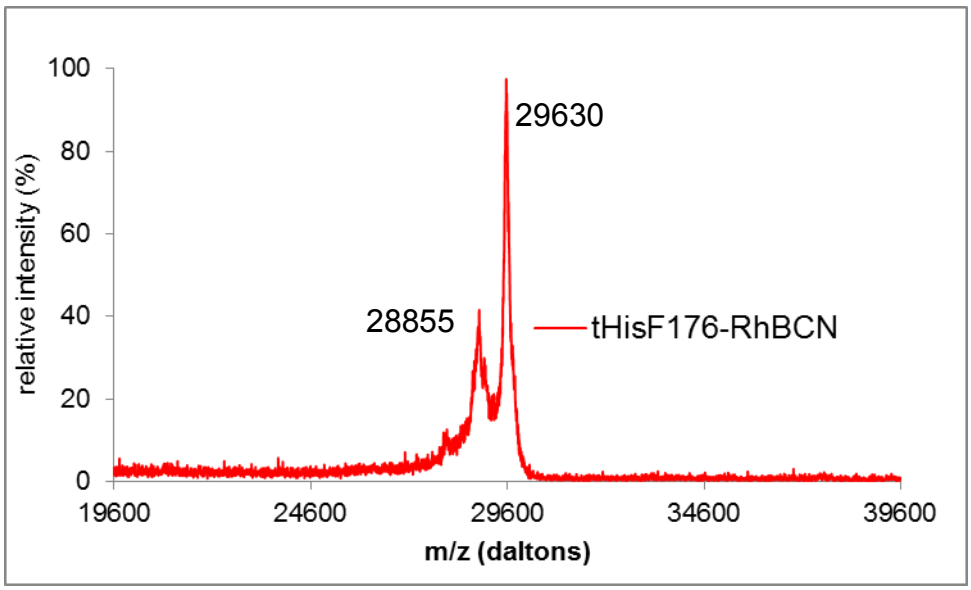
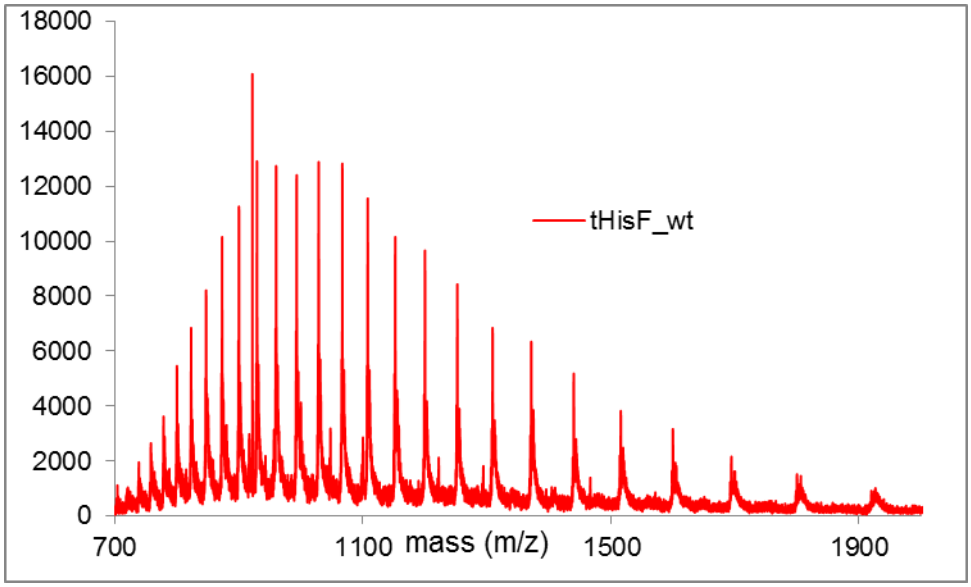


Figure S3. MALDI-MS spectra of representative bioconjugations

Representative ESI-MS spectra for ArM (or scaffold):

In ESI-TOF MS analysis, a sample of protein was desalted with centrifugal filters to a mixture of water: acetonitrile: glacial acetic acid (49.5: 49.5: 1, v/v). The final protein concentration was 50 μM.



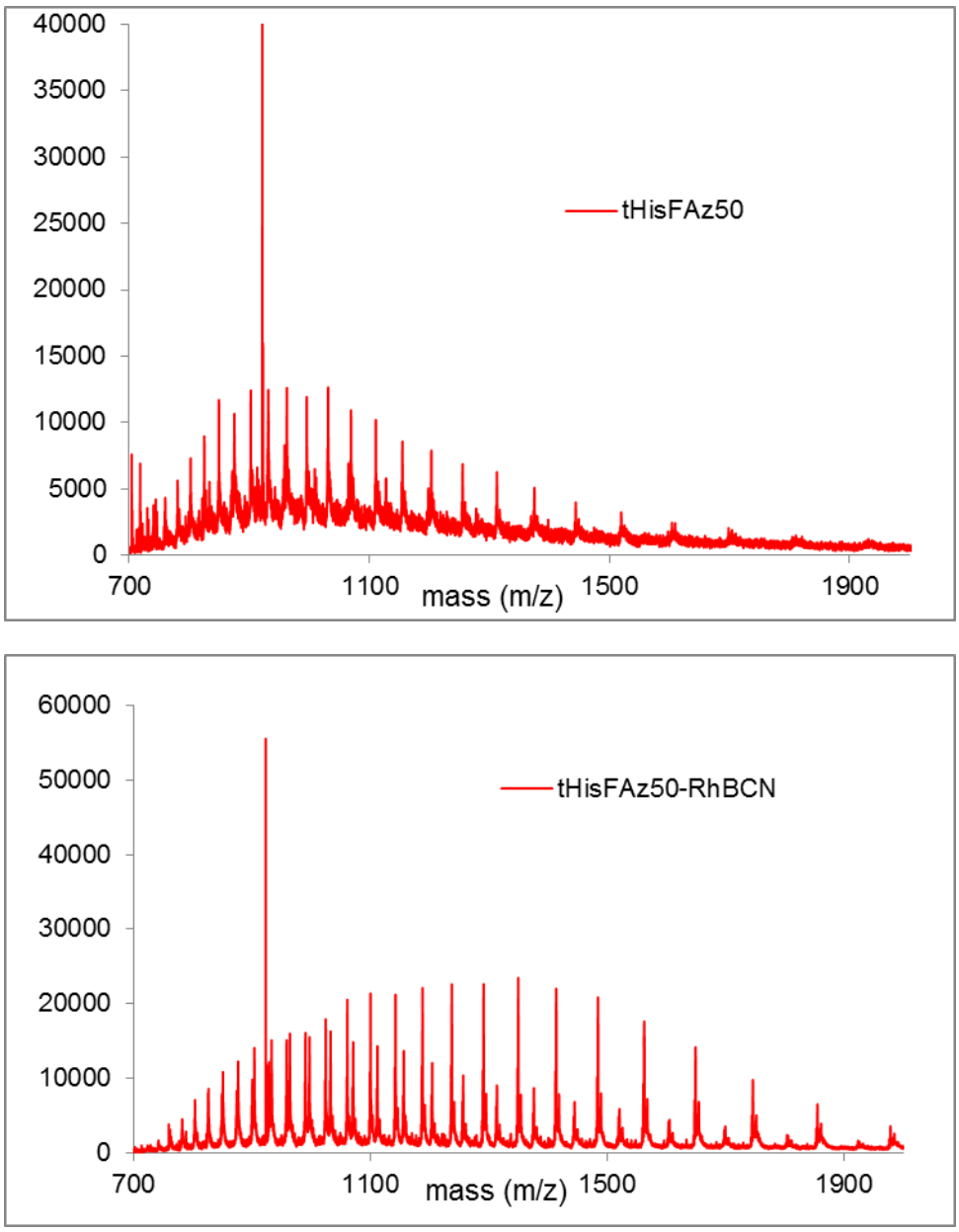


Figure S4. ESI-MS spectra for tHisF-wt, tHisFAz50 and tHisF50-RhBCN (the corresponding deconvoluted spectra are shown in the article)

Circular Dichroism (CD) and Fluorescence Analysis

CD analysis of scaffold proteins and ArMs was conducted by loading the protein solutions into a 0.1 mm quartz cuvette. CD spectra were obtained on AVIV-202 CD spectrophotometer (AVIV Biomedical, Inc.) between 280 and 200 nm with a 1 nm

increment at room temperature. Fluorescence measurements of scaffold proteins and ArMs were acquired at 290 nm using Fluorolog-3 spectrofluorometer (Horiba Jobin Yvon, Inc.) in phosphate buffer (pH 8.0). Proteins were treated with 60% acetonitrile and 6 M guanidine chloride (GdmCl) separately and fluorescence was measured at 290nm along with non-treated protein to check for any structural perturbation.

Circular dichroism of tHisF and tHisFAz50:

Purified wild type tHisF and tHisFAz50 mutant were exchanged into 100 mM sodium phosphate buffer (pH 7.5, 150 mM NaCl), and concentrated to 10 mg/mL to provide 10X stocks for CD experiments. The protein was diluted to a final concentration of 2 mg/mL into buffer and 200 µL of each protein solution was loaded into a 2 mm quartz cuvette (Hellma QC) and CD spectra were obtained on an AVIV 202SF CD spectropolarimeter at 24 °C. For each protein, CD spectra were taken in triplicate between 260 nm and 195 nm in 1 nm increments with a 2 second integration time. Data were averaged and background spectra from free buffer solutions were subtracted.

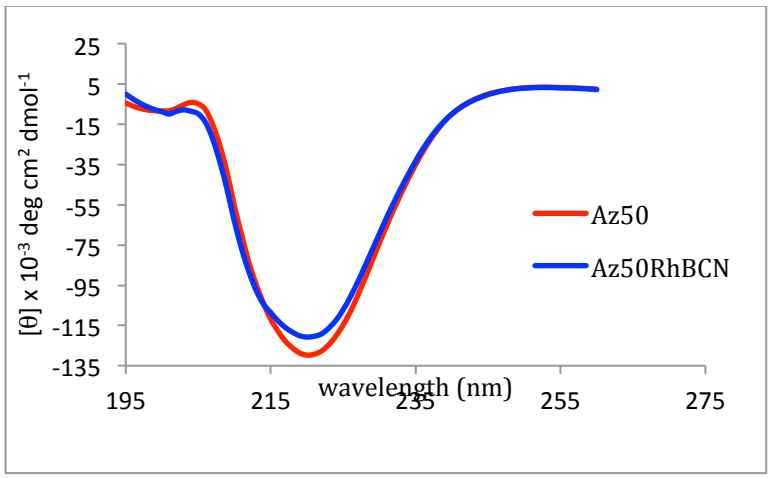


Figure S5. Comparison of CD spectra of tHisFAz50 and tHisFAz50-RhBCN

Fluorescence Analysis:

For organic solvent tolerance measurements, tHisFAz50 was first concentrated down to 100 µM in 10mM tris buffer (pH 7.5). 10 µL of this concentrated protein was then added to 90 µL Acetonitrile, 90 µL DMF and 90 µL methanol in separate 1.5 mL microcentrifuge tubes. Thus final protein concentration becomes 10 µM and buffer to organic solvent ratio becomes 10:90 (v/v) for each sample. A control sample was also prepared where 10 µL of concentrated protein was added to 90 µL buffer (10 mM Tris, pH 7.5). All the samples were incubated at room temperature for an hour and fluorescence emission spectra was measured in a Tecan infinite M200pro plate reader. At 60% acetonitrile concentration no change was observed as described in the article. Typically, 20% acetonitrile concentration was used in bioconjugation reactions.

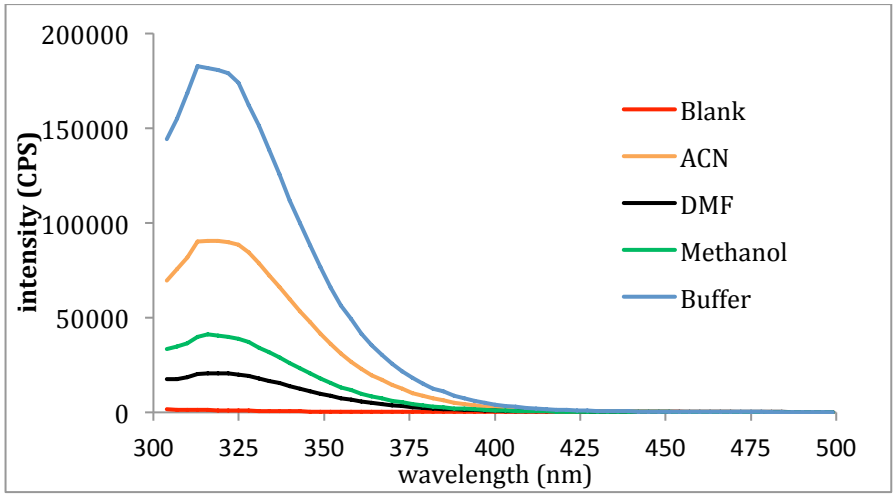


Figure S6. Fluorescence emission spectra of tHisFAz50 in different organic solvents

Dual-Labelling FRET Measurements[13]

Double labeling of Azide(1)/Cys mutants:

Purified tHisFAz50/Cys double mutants were exchanged into labeling buffer (10 mM Tris pH 7.5), by washing 3X 100 fold dilutions using a 10 kDa cutoff 15 mL Amicon® Ultra Centrifugal Filtration Device. 10 equivalents of Alexa594-maleimide were then added to label the free cysteine thiol of 100 μM protein. Reaction was done at room temperature for 16 hours in the dark with constant shaking. Excess unreacted dye was removed by desalting columns followed by buffer exchange 3 times and concentrated to 80 μL. The dye to protein ratio (Alexa594-maleimide:tHisF) was estimated by measuring the absorbance at 280 and 595 nm for tHisF and Alexa594-maleimide, respectively. 10 equivalents of Alexa488-BCN (probe **8**) were then added and the reaction was allowed to proceed at room temperature for 16 h. Excess dye was removed by exhaustive dialysis against water. Protein was aliquoted and stored in 10 mM Tris buffer (pH 7.5) and 150 mM NaCl. The dye to protein ratio (Alexa488:tHisF) was determined by measuring the absorbance at 280 and 495 nm for tHisF and Alexa488, respectively. Extinction coefficients of tHisF (11460 $M^{-1}cm^{-1}$ at 280 nm)[14], Alexa488-BCN (70,000 $M^{-1}cm^{-1}$ at 495 nm, Provided by Invitrogen), Alexa594-maleimide (80000 $M^{-1}cm^{-1}$ at 595 nm)[15] were used to calculate molar concentration of the protein and the dyes. Reactions performed at 100 μM protein concentration typically gave > 80 % labeling after 16 h reaction at 24 °C. Labeling reactions were tried at different conditions and low concentrations of proteins (less than 100 μM) showed slow reaction and a significant decrease in yield. Also negligible labeling was observed when 50mM phosphate buffer (pH 7) was used. Fluorescence Scanning was done by running the labeled proteins (both

singly and doubly) on 10% SDS PAGE. The protein gel was scanned for Alexa488-BCN and Alexa594-maleimide using a fluorescence scanner (Bio-Rad FX pro plus, see Fig. S7).

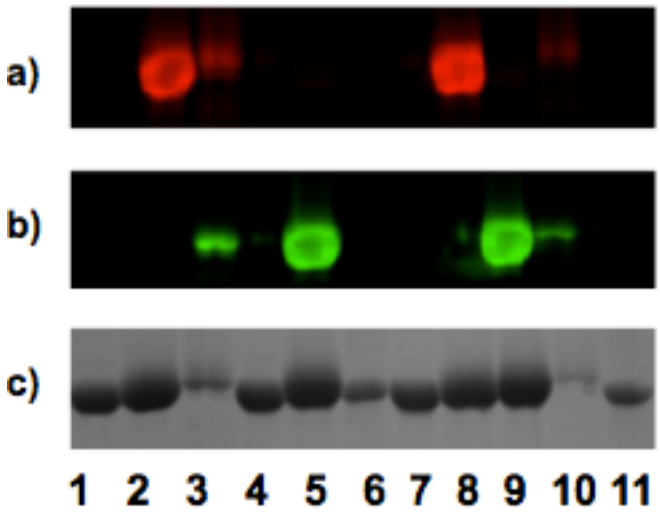


Figure S7. Fluorescence scanned SDS-PAGE picture of labeled proteins (Lane 1, 4 and 7: control protein (tHisF-C9A); lane 2, 8: tHisF-C9A-L50Az-D174C-Alexa488 and tHisF-C9A-L50Az-K243C-Alexa488 respectively; lane 3 and 10: doubly labeled, tHisF-C9A-L50Az-D174C-Alexa488-Alexa594 and tHisF-C9A-L50Az-K243C-Alexa488-Alex594 respectively; lane 5 and 9: tHisF-C9A-L50Az-D174C-Alexa594 and tHisF-C9A-L50Az-K243C-Alex594 respectively; lane 6 and 11: protein markers.)

Steady state FRET measurement:

Steady-state fluorescence measurements were performed at 4 °C on a Fluoromax-3 spectrofluorometer (Horiba Jobin Yvon Inc.). Fluorescence was measured for tHisFAlexa488-BCN (donor) and doubly labeled mutant where tHisFAlexa594-maleimide was the acceptor. The reduction of donor fluorescence emission by the acceptor was recorded between 500 and 700 nm and corrected for the buffer blank. The protein concentrations of tHisFAlexa488-BCN and tHisFAlexa594-maleimide were kept at 2 μM. The experiments were conducted in 10 mM Tris-HCl buffer (pH 7.5). The

distance between residues of tHisFAlexa488-BCN and tHisFAlexa594-maleimide was estimated spectroscopically by FRET. The distance is given by: $R = R_0(E^{-1}-1)^{1/6}$,[16] where R is calculated in Å, R_0 is the Foster critical distance, and E is the FRET efficiency given by $E=(1-F_{da}/F_d) \times 1/F_a$, where $F_{da,}$ F_d and F_a are fluorescence intensities in both, donor, and acceptor channels. R_0 is given by: $R_0 = 9.79 \times 10^3 (\kappa^2 J \Phi_D \eta^{-4})^{1/6}$, where κ2 is the orientation factor, η is the refractive index of the buffer, Φ_D is the quantum yield of the donor, and J is the overlap integral in cm^3 /M given by, $J=\int F_D(\lambda)\varepsilon(\lambda)\lambda_4 d\lambda / F_D(\lambda) d\lambda$, where λ is the wavelength in cm, $F_D(\lambda)$ is the corrected fluorescence of the donor, and $\varepsilon(\lambda)$ is the acceptor molar absorption coefficient in M^{-1} cm^{-1}. J was obtained by numerical integration of normalized spectra. InstruView v-0.5 software (Columbia University) was used to calculate the R_0 values using the defined parameters.[15]

Time-domain lifetimes were measured on a ChronosBH lifetime fluorometer (ISS, Inc.) using Time-Correlated Single Photon Counting (TCSPC) methods. The fluorometer contained Becker-Hickl SPC-130 detection electronics and a HPM-100-40 Hybrid PMT detector. Tunable picosecond pulsed excitation was provided by a Fianium SC400 supercontiuum laser source and integrated AOTF. Emission wavelengths were selected with bandpass filters (Semrock and Chroma). The Instrument Response Function (IRF) was measured to be approximately 120 ps FWHM in a 1% scattering solution of Ludox LS colloidal silica. Lifetimes were fit via a forward convolution method in the Vinci control and analysis software.[15]

Serial	Double mutant	Energy transfer	Distance (R) in A^0

		$E=(1-F_{da}/F_d) \times 1/F_a$		$E=1/1+(R/R_0)^6$	
		Steady state	Lifetime	Steady state	Lifetime
1	Az50, D174C (top)	0.54	0.63	58	55
2	Az50, K243C (bottom)	0.11	0.15	88	80

Table S2. R_0 values calculated from steady state FRET measurement

(F_{da}: Fluorescence of doubly labeled protein

F_d: Fluorescence of donor protein

F_a: Fluorescence of acceptor protein)

ArM catalysis

Preparation of standard products for cyclopropanation

In a 25 mL round-bottom flask, 4-methoxystyrene (482 μL, 3.6 mmol) and rhodium acetate dimer (10.6 mg, 0.024 mmol, 0.0067 equiv) were dissolved in 7.5 mL ether. A solution of ethyl diazoacetate (351 μL, 3.0 mmol) in 5 mL ether was added dropwisely over 30 minutes. The reaction was stirred at room temperature for 8 h and the solvent was evaporated. Purification by flash chromatography (silica gel, 19:1 hexanes/EtOAc) afforded both *cis*-isomer (60 mg, 9 %) and *trans*-isomer (146 mg, 22 %). The characterization of products is consistent with literature reports[17, 18].

Preparation of standard products for silane insertion

In a 25 mL round-bottom flask, methyldiphenylsilane (67 μL, 0.44 mmol, 1.1 equiv) and rhodium acetate dimer (1.8 mg, 0.004 mmol, 0.01 equiv) were dissolved in 4 mL hexane. A solution of methyl phenyldiazoacetate (59 μL, 0.4 mmol) in 4 mL hexane was added dropwisely over 1 h. The reaction was stirred at room temperature overnight and the solvent was evaporated. Purification by flash chromatography (silica gel, 19:1

hexanes/EtOAc) afforded the product (54 mg, 39 %), which is consistent with literature report[19]. The O-H insertion product methyl DL-mandelate was prepared as previously reported[20].

Catalytic cyclopropanation

In a 1.5 mL microcentrifuge tube, tHisFA176-RhBCN solution (80 µL, 100 µM), 90 µl phosphate buffer (0.1 M, pH 7.5), and 26 µL THF were added. A mixed solution of styrene and diazoacetate in THF (4 µL, styrene 600 mM, diazoacetate 200 mM) was added. The resulting solution was left shaking at room temperature overnight. The final concentrations of the reagents were: 12 mM styrene, 4 mM diazoacetate, 40 µM tHisF-Az176-RhBCN solution. The reaction was quenched by adding 800 µL chloroform to the closed vials and immediately vortexing the mixture. The vial was then opened and 20 µL internal standard (18.4 mM anisole in acetonitrile) was added. The mixture was vortexted and centrifuged (16,000xg, 1 min). The bottom organic layer was evaporated and re-dissolved in 200 µL acetonitrile and analyzed by HPLC.

Catalytic silane Insertion

In a 1.5 mL microcentrifuge tube, tHisF-Az176-RhBCN solution (75 µL, 130 µM), 95 µL phosphate buffer (0.1 M, pH 7.5) and 22 µL THF were added. A mixed solution of silane and diazoacetate in THF (8 µL, styrene 25 mM, diazoacetate 125 mM) was added. The resulting solution was left stirring at room temperature overnight. The final concentrations of the reagents were: 1 mM silane, 5 mM diazoacetate, 40 µM tHisFA176-RhBCN solution. The reaction was quenched by adding 800 µL chloroform

to the closed vials and immediately vortexing the mixture. The vial was then opened and 8 µL internal standard (26.8 mM 1,2,4-trimethoxybenzene in acetonitrile) was added. The mixture was vortexted and centrifuged (16,000 x g, 1 min). The bottom organic layer was evaporated and re-dissolved in 200 µL hexane and analyzed by HPLC.

HPLC analysis of cyclopropanation:

The analytic HPLC run for cyclopropanation was performed on an Agilent 1100 Series HPLC system using an Agilent Eclipse Plus C18 column (95 Å, 3.5 µM, 4.6 mm i.d. x 150 mm), with with a flow rate of 1.0 mL/min and detection wavelength set at 230 nm. The following gradient was used: 20 % to 50 % B from 0-5 min, 50 % from 5-10 min, 50 % to 80 % from 10-15 min, 80 % from 15-18 min, 80 % to 20 % from 18-20 min (solvent A: water containing 0.1% TFA; solvent B: CH_3CN).

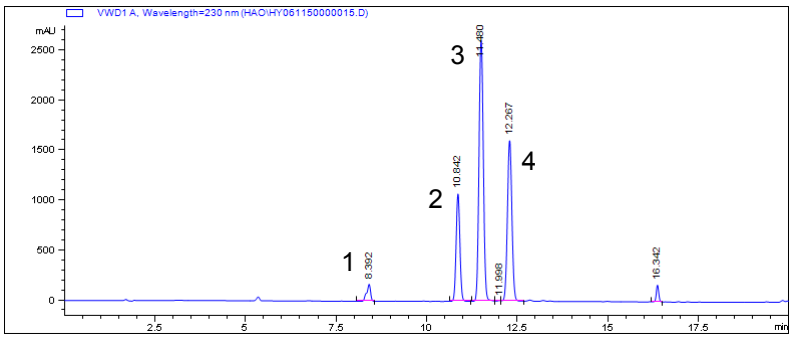


Figure S8. HPLC traces of cyclopropanation catalyzed by tHisF176-RhBCN (1: anisole internal standard; 2: cis-product isomer; 3: 4-methoxystyrene starting material; 4: trans-product isomer)

HPLC analysis of silane insertion:

The analytic HPLC run for cyclopropanation was performed on an Agilent 1200 UHPLC system using a Lux® 3u Cellulose-1 column (1000 Å, 3.0 µM, 4.6 mm i.d. x 250 mm),

with with a flow rate of 1.0 mL/min and detection wavelength set at 230 nm. The following gradient was used: 97 % B from 0-15 min, 97-90 % from 15-17 min, 90 % from 17-26 min, 90-97 % from 26-28 min, 97 % from 28-30 min (solvent A: isopropanol; solvent B: hexane).

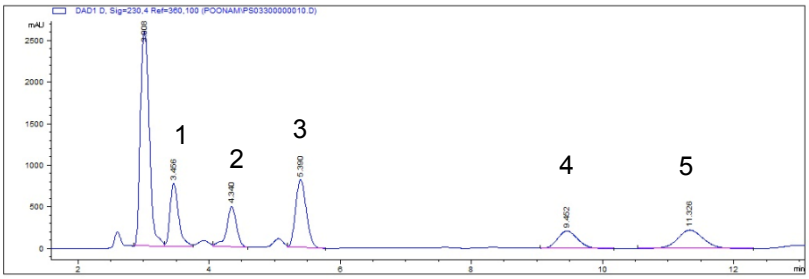


Figure S9. HPLC trace of silane insertion catalyzed by tHisF176-RhBCN (1: phenyldimethysilane; 2: one product enantiomer; 3: methyl phenyldiazoacetate; 4: the other enantiomer; 5: 1,2,4-trimethoxybenzene internal standard)

Table S3. Summary of yields for silane insertion and cyclopropanation catalyzed by ArM and small molecule dirhodium catalysts.

Catalyst	Silane insertion (relative to silane)		Cyclopropanation (%)[a]
	Si-H insertion (%)[a]	Diazo hydrolysis (%)[a]	
tHisFAz176-RhBCN (top)	28	322	81 (cis: trans =1: 1.8)
tHisFAz50-RhBCN (middle)	5	327	69 (cis: trans = 1: 1.8)
tHisFA199-RhBCN (bottom)	6	98	60 (cis: trans =1: 1.8)
2-OAc	80	418	99 (cis: trans =1: 1.8)

[a]All the yields were calculated relative to limiting reagent (silane for Si-H insertion and olefin for cyclopropanation) in the reactions by analysis of HPLC traces for crude reaction mixtures.

DFT model of ArM

The phenyl azide adduct of cofactor **3** (Figure S10a) was optimized using density functional theory calculation (DFT, B3LYP, LANL2DZ) using Gaussian09 (Figure S10b). The mutation wizard in Pymol was used to convert residue Ile199 of tHisF

(PDB#1THF) to Ala. The Ala methyl and terminal phenyl of the cofactor **3** adduct were fused using the "fuse" command in Pymol. Bond angles were manually adjusted to provide a rough model of the ArM (Figure S10c). While crude, this model provides some idea of the relative scale of scaffold and cofactor.

Figure S10a. Structure of phenyl azide-**3** adduct

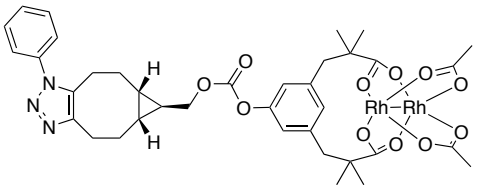


Figure S10b. DFT-optimized structure of adduct

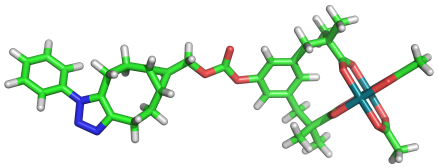
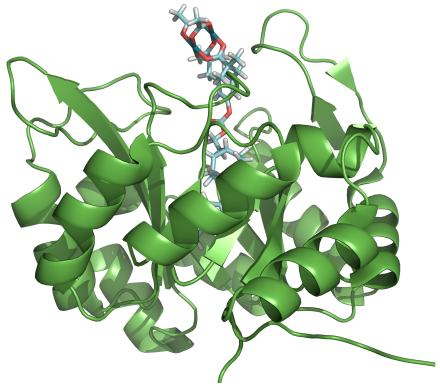


Figure S10c. tHisF-Az199-RhBCN (note that higher mutants, 50 and 176, should project cofactor further into solution)



References

1. C. G. Espino, K. W. Fiori, M. Kim, J. Du Bois, *J. Am. Chem. Soc.* **2004**, *126*, 15378-15379.
2. Y. Lou, T. P. Remarchuk, E. J. Corey, *J. Am. Chem. Soc.* **2005**, *127*, 14223-14230.

3. J. Dommerholt, S. Schmidt, R. Temming, L. J. A. Hendriks, F. P. J. T. Rutjes, J. C. M. van Hest, D. J. Lefeber, P. Friedl, F. L. van Delft, *Angew. Chem. Int. Ed.* **2010**, *49*, 9422-9425.
4. E. C. Constable, M. D. Ward, *J. Chem. Soc., Dalton Trans.* **1990**, 1405-1409.
5. M. T. Reetz, M. Rentzsch, A. Pletsch, A. Taglieber, F. Hollmann, R. J. G. Mondiere, N. Dickmann, B. Hocker, S. Cerrone, M. C. Haeger, R. Sterner, *ChemBioChem* **2008**, *9*, 552-564.
6. T. S. Young, I. Ahmad, J. A. Yin, P. G. Schultz, *J. Mol. Biol.* **2010**, *395*, 361–374.
7. K. L. Heckman, L. R. Pease, *Nat. Protoc.* **2007**, *2*, 924-932.
8. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular cloning: A laboratory manual*, Cold Spring Harbor Laboratory Press: 4th edition, **2012**, Cold Spring Harbor, N.Y.
9. J. M. Murphy, X. Liao, J. F. Hartwig, *J. Am. Chem. Soc.* **2007**, *129*, 15434-15435.
10. R. Sambasivan, Z. T. Ball, *J. Am. Chem. Soc.* **2010**, *132*, 9289-9291.
11. K. Suntharalingam, A. J. P. White, R. Vilar, *Inorg. Chem.* **2010**, *49*, 8371.
12. H. Chen, R. Tagore, S. Das, C. Incarvito, J. W. Faller, R. H. Crabtree, G. W. Brudvig, *Inorg. Chem.* **2005**, *44*, 7661-7670.
13. R. Roy, S. Hohng, T. Ha, *Nat. Methods* **2008**, *5*, 507-516.
14. The extinction coefficient of tHisF was determined according to the Lambert-Beer law by using the concentration obtained from a Pierce® BCA Protein Assay Kit.
15. Q. Guo, J. E. Jureller, J. T. Warren, E. Solomaha, J. Florian, W. J. Tang, *J. Biol. Chem.* **2008**, *283*, 23836-23845.
16. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, 3rd ed., Springer, New York, **2006**, p. 443-472.
17. M. L. Rosenberg, A. Krivokapic, M. Tilset, *Org. Lett.* **2009**, *11*, 547-550.
18. Y. Chen, X. P. Zhang, *J. Org. Chem.* **2007**, *72*, 5931-5934.
19. L. A. Akin, P. C. Ong, J. S. Panek, R. J. Staples, P. Stavropoulos, *Organometallics* **2000**, *19*, 2896-2908.
20. J. Jr. Jones, A. J. Kresge, *J. Org. Chem.* **1993**, *58*, 2658-2662.