Supporting Information

A rhodium complex-linked β-barrel protein as a hybrid biocatalyst for phenylacetylene polymerization

Akira Onoda,[†,‡] Kazuki Fukumoto,[†] Marcus Arlt,[§] Marco Bocola,[§] Ulrich Schwaneberg,[§] and Takashi Hayashi[*,†]

[†]*Department of Applied Chemistry, Graduate School of Engineering, Osaka University, Suita, Osaka 565-0871, Japan. Fax: + 81-6-6879-7930; Tel: + 81-6-6879-7928; E-mail: thayashi@chem.eng.osaka-u.ac.jp*

[‡]*Frontier Research Base for Global Young Researchers, Graduate School of Engineering, Osaka University, Suita, Osaka 565-0871, Japan.*

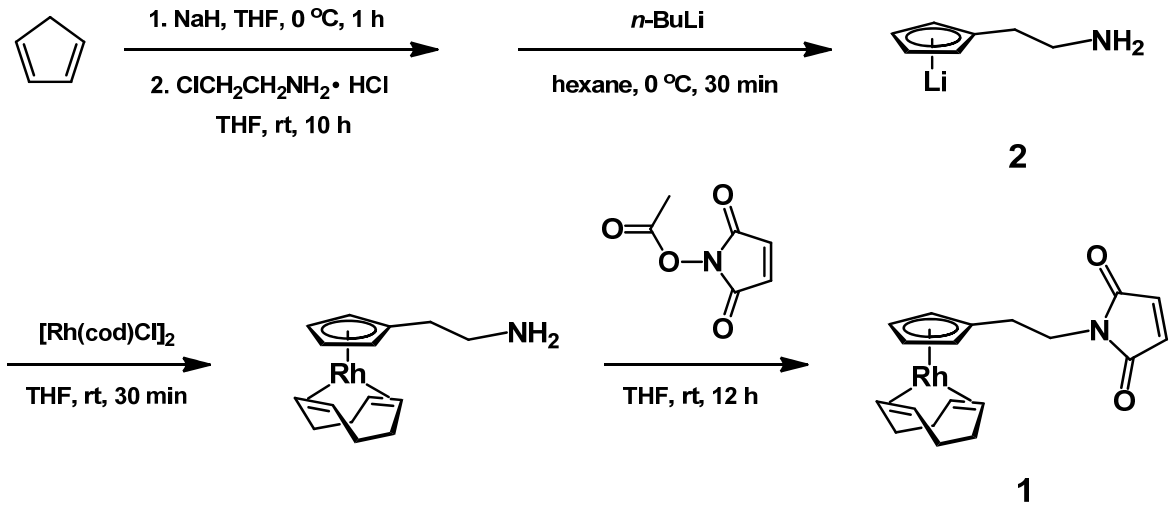
[§]*RWTH Aachen University, Worringerweg 1, D-52074 Aachen, Germany.*

Experimental Section

Materials. The pUC57 plasmid containing optimized nitrobindin (NB) (M75L, M148L) gene was purchased from GeneScript. Oligonucleotides were obtained from Invitrogen, Inc. (Japan). Restriction enzymes were obtained from Takara Bio Inc. (Japan). Nucleotide sequences were determined by Fasmac (Japan). All reagents of the highest guaranteed grade were purchased and used as received unless otherwise noted. A standard rhodium solution for inductively coupled plasma optical emission spectroscopy (ICP-OES) was purchased from Wako (Japan). Distilled water was demineralized by a Barnstead NANOpure DIamond^{TM} apparatus. $[Rh(cod)Cl]_2$ was synthesized according to the previous report. [S1]

Instruments. ^{1}H NMR spectra were recorded on a Bruker DPX400 NMR spectrometer. High-resolution magic angle spinning (HR-MAS) experiments were performed on a Varian Unity Inova 600 MHz NMR spectrometer equipped with a 4 mm gHX Nanoprobe. Chemical shifts were reported in ppm relative to the residual solvent resonances. ESI-TOF MS analyses were performed on an Applied Biosystems Mariner API-TOF Workstation or a Bruker micrOTOF focus III mass spectrometer and MALDI-TOF MS analyses were performed on a Bruker autoflex III mass spectrometer. UV-vis experiments were conducted using a Shimadzu UV-3150 double-beam spectrophotometer equipped with a thermostated cell holder with a 0.1 ºC deviation. Purification of the proteins was performed using a GE healthcare ÄKTA Purifier system at 4 ºC. Circular dichroism (CD) spectra were recorded by JASCO J720S spectrometer. ICP-OES was performed on a Shimadzu ICPS-8100 emission spectrometer. The pH values were monitored with a Horiba F-52 pH meter. Air-sensitive manipulations were performed in an MBraun glovebox. Gel permeation chromatography was performed on a TOSOH SC8020 apparatus with a refractive index detector with the TOSOH TSKgel G4000HHR column.

Scheme S1. Synthetic scheme of rhodium complex **1**.



Synthesis of $LiC_5H_4CH_2CH_2NH_2$ (2). To a solution of freshly prepared cyclopentadiene (20 mL, 240 mmol) in THF (300 mL) was added NaH (12.0 g, 290 mmol) at 0 ºC, and the mixture was stirred at 0 ºC for 30 min and at room temperature for additional 30 min. To the solution was added $ClCH_2CH_2NH_2 \cdot HCl$ (13.5 g, 120 mmol) and the mixture was stirred at room temperature for 10 h. The reaction was quenched by the addition of H_2O and the product was extracted with pentane. The organic phase was concentrated to ca. 30 mL and mixed with 2 M HCl_{aq} until the aqueous phase was adjusted to pH 1. The product was further extracted with an acidic solution and additional portion of 0.1 M HCl_{aq}, and the combined aqueous solution was then adjusted to pH 14 with 4 M $NaOH_{aq}$. The product was again extracted with pentane and the organic phase was dried over Na_2SO_4. The product in pentane solution was concentrated to ca. 10 mL and the solution was mixed with excess amount of *n*-BuLi (1.6 M in hexane, 15 mL). The product precipitated was collected by filtration and dried in vacuo to give air- and moisture-sensitive white powder, **2** (660 mg, 4.8%): ^{1}H NMR (400 MHz, THF-d_8): δ 5.61 (t, 2H, J = 2.4 Hz), 5.45 (t, 2H, J = 2.4 Hz), 2.69 (tt, 2H, J = 6.0 Hz, 6.0 Hz), 2.54 (t, 2H, J = 6.0 Hz), 1.08 (t, 2H, J = 6.0 Hz).

Synthesis of rhodium complex 1. $[Rh(cod)Cl]_2$ (79 mg, 0.16 mmol) and **2** (40 mg, 0.37 mmol) were dissolved in 5 mL of THF and the solution was stirred at room temperature for 30 min. To the solution was added *N*-methoxycarbonylmaleimide (55 mg, 0.37 mmol) in 10 mL of THF and the mixed solution was stirred at room temperature for 24 h. The product was purified by column chromatography (Al_2O_3, THF) and dried in vacuo. The residue was recrystallized from hexane to give air-stable yellow crystal **1** (22 mg 14.9 %): ^{1}H NMR (400 MHz, C_6D_6): δ 5.71 (s, 2H), 4.85 (m, 2H), 4.80 (m, 2H), 3.83 (m, 4H), 3.57 (m, 2H), 2.28 (m, 6H), 1.96 (m, 4H); ^{13}C NMR (100 MHz, C_6D_6): δ 169.9, 133.1, 102.0, 86.8, 85.8, 63.9, 38.7, 32.7, 26.7; ESI-TOF MS (positive mode) *m*/*z* calcd for $C_{19}H_{22}NO_2Rh$ [M + H^+] 400.078, Found 400.090.

Crystal structure determination. The crystal of **1** was mounted on a glass fiber, and the X-ray data was collected at 108 K on a Rigaku VariMax RAPID imaging plate area detector with graphite monochromated Mo-Kα radiation (0.71075 Å). The basic crystallographic parameters for **1** are listed in Table S1. The structure was solved by the direct method and expanded using SIR92.[S2] Non-hydrogen atoms were refined anisotropically and all H atoms were located at the calculated positions. All structures in the final stages of refinement showed no movement in the atom positions. The calculations were performed using Single-Crystal Structure Analysis Software, CrystalStructure (version 3.8.2) [S3] and SHELXL-97. [S4]

Expression and purification of the nitrobindin proteins. The pUC57 plasmid containing optimized NB (M75L, M148L) gene was used as a template for PCR with oligonucleotide primers (i) 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGATAGAACCATGAATCAA-CTGCAACAACTGC-3′ and (ii) 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATTA-CAGTTTGTCCAGGATGGC-3′. The PCR product was inserted into pDONR vector and then pDEST14 expression vector by the standard protocols for Gateway technology (Invitrogen). Successful mutagenesis, Q96C, was achieved with the QuickChange Mutagenesis kit (StrataGene). DNA sequencing was performed to verify the correct insertion of the gene sequence into the expression

vector. The resulting expression plasmid, pNB, was transformed into *E. coli* BL21(DE3). Each 1 L of a TB medium containing ampicillin (50 mg) and 1% (v/v) glucose was inoculated with 10 mL of the culture (OD = 0.5) of the relevant transformed cells. After the cells were grown aerobically with vigorous shaking at 37 ºC until the OD_{600} reached ~0.5, isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM to induce the protein expression. The incubation was continued at 37 ºC for approximately 6 h. The cells were harvested by centrifugation at 4000 ×g for 10 min. Tricine SDS-PAGE analysis showed that both the supernatant and pellet from the lysed cells contained an overexpressed protein (~18k Da).

The harvested cells from 4 L of culture were re-suspended in ca. 100 mL of a 20 mM Tris-HCl buffer (pH 8.0) containing 15 mM PMSF, 1mM EDTA, and 10 mM DTT and cell lysed by freeze–thaw cycles, followed by a reaction with 10 μL of benzonase® nuclease (Novagen) for 30 min at 4 ºC. The lysate was then centrifuged and the collected supernatant was dialyzed three times against 1 L of a 20 mM Tris-HCl buffer (pH 8.0). The dialyzed solution was loaded onto an anion-exchange column with HiTrap DEAE Fast Flow (GE healthcare) which was pre-equilibrated in a 5 mM Tris-HCl buffer (pH 9.0). The fraction of the target protein was collected by a 5 mM Tris-HCl buffer (pH 9.0) and 0.5 M NaCl, and concentrated using an Amicon stirred ultrafiltration cell with a 10-kDa molecular weight cut-off membrane (Millipore). The purified fractions were characterized by SDS-PAGE and MALDI-TOF MS. MALDI-TOF MS (positive mode) Q96C-NB: *m*/*z* calcd for $C_{830}H_{1304}N_{222}O_{248}S_2$ $[M + H]^+$ 18424.94, found 18425.83;

Conjugation of rhodium complex 1 with NB mutant. The NB mutant (10 μM) in 5 mL of a 10 mM Tris-HCl buffer (pH 8.0) was added DTT (2 mg/mL) to reduce a thiol group of the protein. The mixed solution was incubated at 4 ºC for 30 min and concentrated to 0.5 mL using an Amicon stirred ultrafiltration cell with a 10-kDa molecular weight cut-off membrane (Millipore). The reduced protein was purified using a HiTrap desalting column (GE healthcare) equilibrated with a 10 mM Tris-HCl buffer (pH 8.0). The collected protein solution (1.5 mL) was slowly added to **1** in DMSO (0.1 mg/20

μL) and the mixed solution was incubated at room temperature for 30 min. The precipitate of the rhodium complex was removed by filtration with membrane filters (0.44 μm) and the solution was again concentrated to 0.5 mL. The Q96C-NB–1 hybrid was purified by the HiTrap desalting column eluting with a 10 mM Tris-HCl buffer (pH 8.0). The purified fractions were characterized by MALDI-TOF MS. MALDI-TOF MS (positive mode) Q96C-NB–1: *m*/*z* calcd for $C_{830}H_{1304}N_{222}O_{248}S_2$ [M – cod]$^+$ 18424.94, found 18425.83.

Conjugation of rhodium complex with A125C-Mb. The conjugation of **1** with A125C-Mb mutant was performed by the method similar to Q96C-NB–1. The purified fractions were characterized by MALDI-TOF MS. MALDI-TOF MS (positive mode) A125C-Mb–**1**: *m*/*z* calcd for $C_{830}H_{1304}N_{222}O_{248}S_2$ [M – cod]$^+$ 17653.10, found 17653.81.

Polymerization of phenylacetylene. All emulsion polymerizations were carried out by the hybrid catalyst (10 μM) and phenylacetylene (1 M) in 1.5 mL of a buffer at 25 ºC for 24 h. After polymerization, the organic phase was separated by microsyringe and evaporated to dryness. The residue was washed with hexane and methanol to remove the unreacted monomer and dried in vacuo to give a poly(phenylacetylene) (PPA) product.

Gel permeation chromatography (GPC). Chloroform was used as an eluent at a flow rate of 1.0 mL/min. The molecular weight calibration curve was obtained with standard polystyrenes.

Determination of stereostructure of PPAs. The PPA product dissolved in $CDCl_3$ was transferred to a 4 mm magic angle spinning (MAS) rotor and the NMR spectrum was recorded by high-resolution magic angle spinning (HR-MAS) method. The ratio of *cis* and *trans* PPAs was determined based on the peak intensities of each signal.

Molecular modeling of hybrid catalyst. Modeling of the hybrid catalyst was performed using YASARA[S5] Structure Vers. 11.6.16 employing force field AMBER03[S6] for protein residues and GAFF[S7] using AM1/BCC[S8] partial charges for the catalyst covalently bound to Cys96. The X-ray structure of **1** was used as the starting structure. The metal was replaced by cobalt, since no parameters

are available for rhodium. To maintain the correct coordination geometry, the distances from the metal to all five carbon atoms of the Cp ligand were constrained to 2.2 Å and the η^5 coordination is represented by five force field arrows. To maintain η^2 coordination of the cod ligand, four force field arrows were defined and the distances were constrained to 2.0 Å, according to the X-ray structure of **1**. The total charge of the Rh complex **1** was set to zero and the charge of the metal was set to +1 for Rh^+. The bond orders for the Cp ligand were defined as 1.67 Å to maintain an aromatic system. As basis for the hybrid catalyst, the X-ray structure of NB with bound heme ligand (pdbID 3EMM) was used. We removed the heme ligand and water molecules and mutated Gln96 to cysteine. The catalyst was placed manually in the cavity with its maleimide moiety adjacent to Cys96 and defined a bond from Cys Cγ atom to the C1 or C2 atom of the maleimide group. The constructed hybrid catalyst was solvated in a box of TIP3P water molecules using periodic boundaries at pH 7 and a density of 0.997 g/mL. Several starting structures were analyzed and two favorable orientations were identified for covalent attachment to the reactive maleimide atoms by steepest descent minimization and simulated annealing. The minimized structures were relaxed using molecular dynamics calculations at 298 K for 500 ps and snapshots were taken every 25 ps to analyze the binding modes. In the case of the hybrid catalyst linked via C1 of maleimide, the binding mode was not stable after 500 ps and an additional 2000 ps of MD-simulation was performed. The most favorable two binding modes are shown in Fig. 2c and Fig. S5.

Table S1: Crystallographic Data for **1**.

parameters	**1**
empirical formula	$C_{19}H_{22}NO_2Rh_1$
formula weight	399.29
crystal system	Monoclinic
a, Å	6.3936(2)
b, Å	11.9067(4)
c, Å	21.0945(6)
α, deg	90.00
β, deg	91.9750(10)
γ, deg	90.00
V, Å^3	90.00
β, deg	91.9750(10)
γ, deg	90.00
V, Å^3	1604.90(9)
space group	$P2_1/n$
Z	4
ρ_{calc}, g cm^{-1}	0.064 (0.300)
μ (MoKα), cm^{-1}	96.5 (89.7)
temp	108
data	15370
unique data	3679
R_1	2.96
wR_2	5.80
GOF	1.054

[a] $R1 = \Sigma(|Fo| - |Fc|)/ \Sigma |Fo|$. [b] $wR2 = \{\Sigma(w(Fo^2 - Fc^2)^2)/\Sigma w(Fo^2)^2\}^{1/2}$.

Optimized codon sequence of NB (M75L,M148L):

AATCAACTGCAACAACTGCAAAATCCGGGC**G**AGAGTCCGCCGGTTCATCCGTTCGTGGC
ACCGCTGTCCTATCTGCTGGGTACCTGGCGCGGCCAGGGTGAAGGCGAGTATCCGACCATT
CCGAGCTTTCGCTATGGCGAAGAGATCCGTTTCAGCCATTCGGGTAAACCGGTGATTGCCT
ATACCCAAAAAACGTGGAAACTGGAATCGGGTGCACCGCTGCACGCAGAGAGTGGTTATT
TTCGCCCGCGTCCGGATGGTTCTATTGAAGTGGTTATCGCACAGTCGACCGGTCTGGTGGA
AGTTCAAAAAGGCACGTATAATGTGGATGAGCAGAGTATTAAACTGAAATCTGACCTGGT
GGGCAACGCGTCCAAAGTTAAAGAAATCAGCCGCGAATTCGAGCTGGTTGACGGTAAACT
GAGTTATGTGGTTCGTCTGAGCACGACCACGAATCCGCTGCAACCGCACCTGAAAGCCAT
CCTGGACAAACTGTAA

Amino acid sequence of NB (M75L,M148L):

MNQLQQLQNPGESPPVHPFVAPLSYLLGTWRGQGEGEYPTIPSFRYGEEIRFSHSGKPVIAYTQ
KTWKLESGAPLHAESGYFRPRPDGSIEVVIAQSTGLVEVQKGTYNVDEQSIKLKSDLVGNASK
VKEISREFELVDGKLSYVVRLSTTTNPLQPHLKAILDKL

Forward primer for Gateway cloning: 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCG-AAGGAGATAGAACCATGAATCAACTGCAACAACTGC-3′. Reverse primer for Gateway cloning: 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATTACAGTTTGTCCAGGAT-GGC-3′.

Forward primers for Q96C-NB: 5′-GTTCTATTGAAGTGGTTATCGCATGCTCGAC-CGGTCTGGTGGAAG-3′. Reverse primers for Q96C-NB: 5′-CTTCCACCAGACCGGTCGAGC-ATGCGATAACCACTTCAATAGAAC-3′.

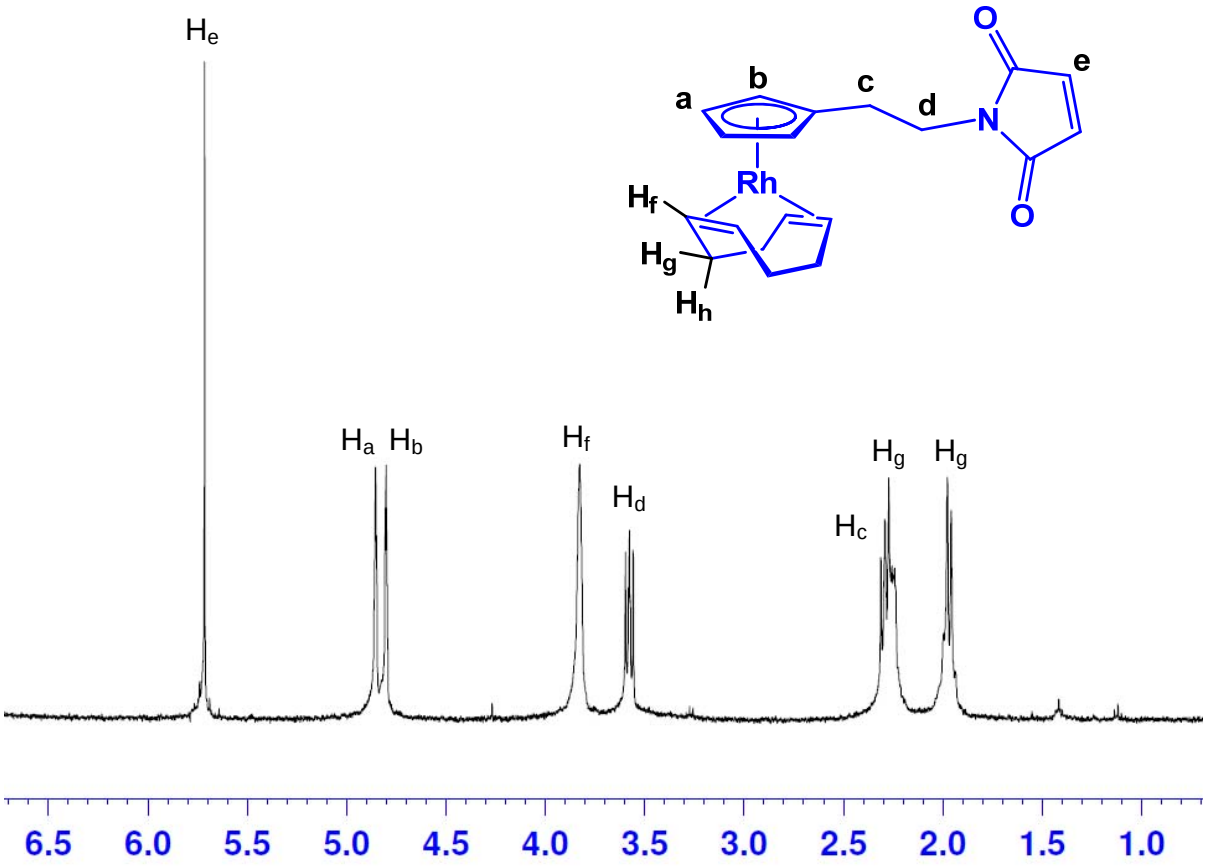


Fig. S1 400 MHz ^{1}H NMR spectrum of **1** in $CDCl_3$ at 296 K.

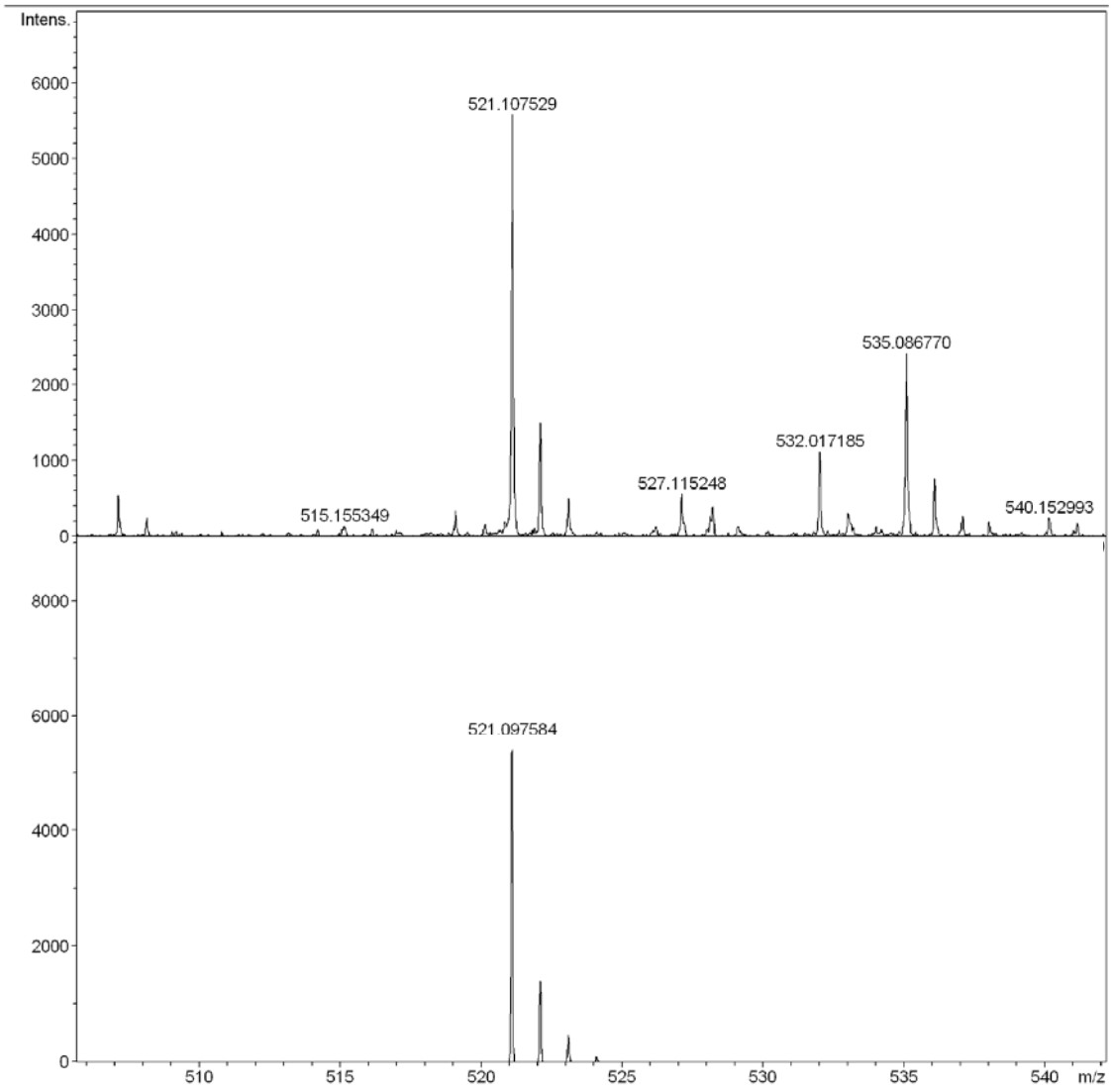


Fig. S2 ESI-TOF MS spectrum of **1** coupled with cysteine (upper). Simulated spectrum is shown at the bottom.

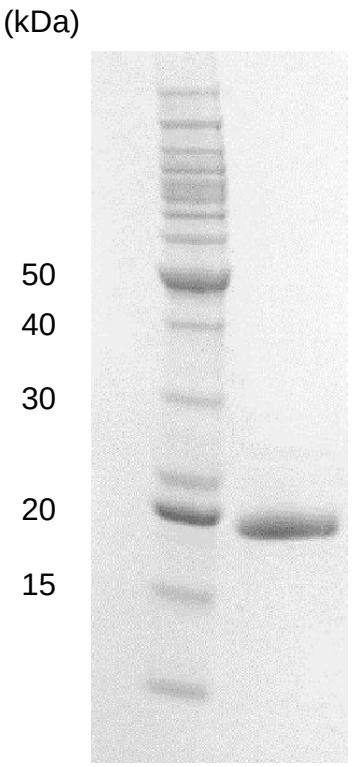


Fig. S3 SDS-PAGE analysis of Q96C-NB–**1** protein (right lane) with molecular weight markers (left lane).

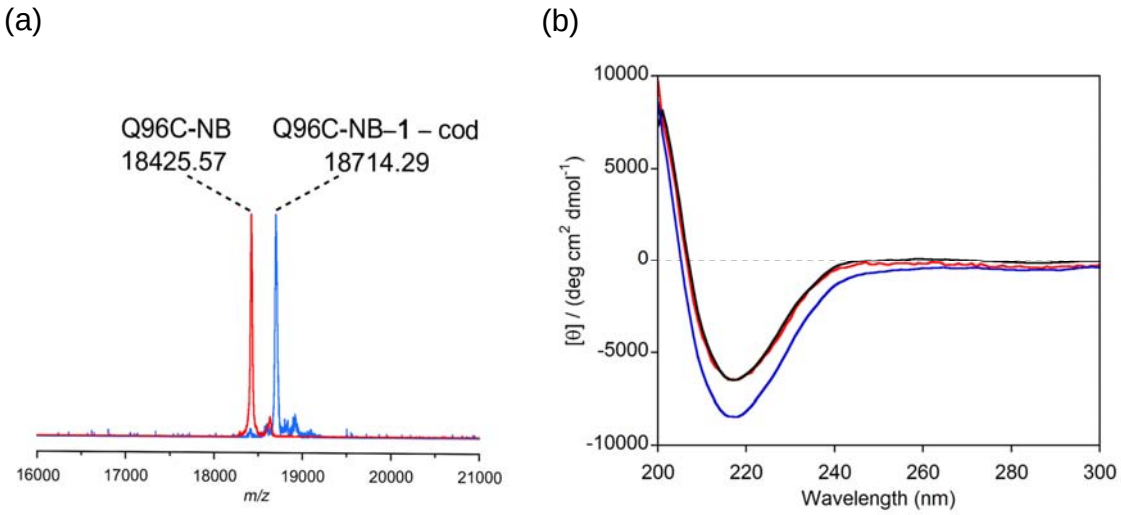


Fig. S4 (a) MALDI-TOF MS spectra of Q96C-NB (in red) and Q96C-NB–1 (in blue). (b) CD spectra of holo NB (in black), Q96C-NB (in red), and Q96C-NB–**1** (in blue).

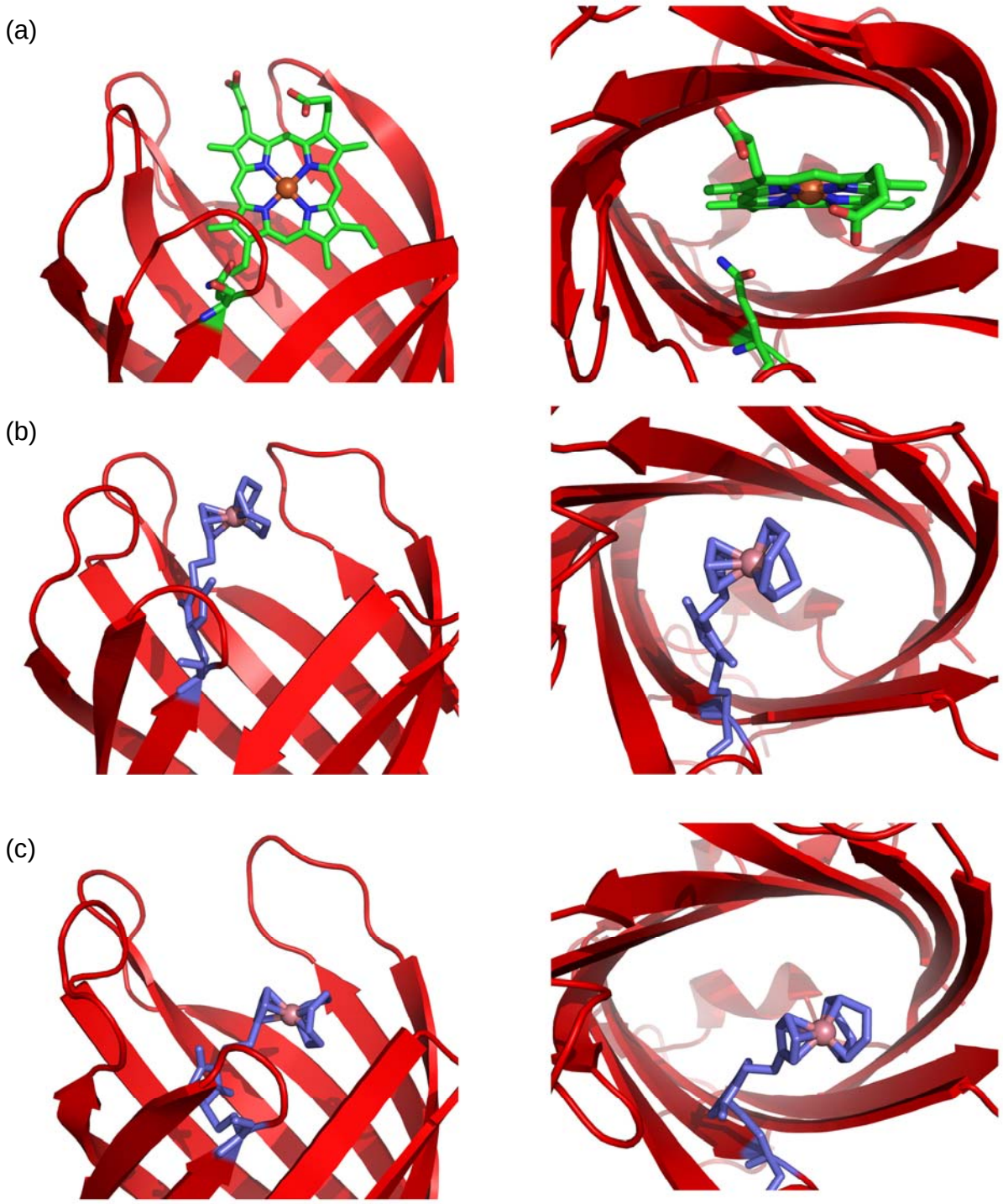


Fig. S5 Molecular modeling of the hybrid catalyst. (a) native nitrobindin (PDBID 3EMM), (b) model 1 of Q96C-NB–**1**, and (c) model 2 of Q96C-NB–**1**.

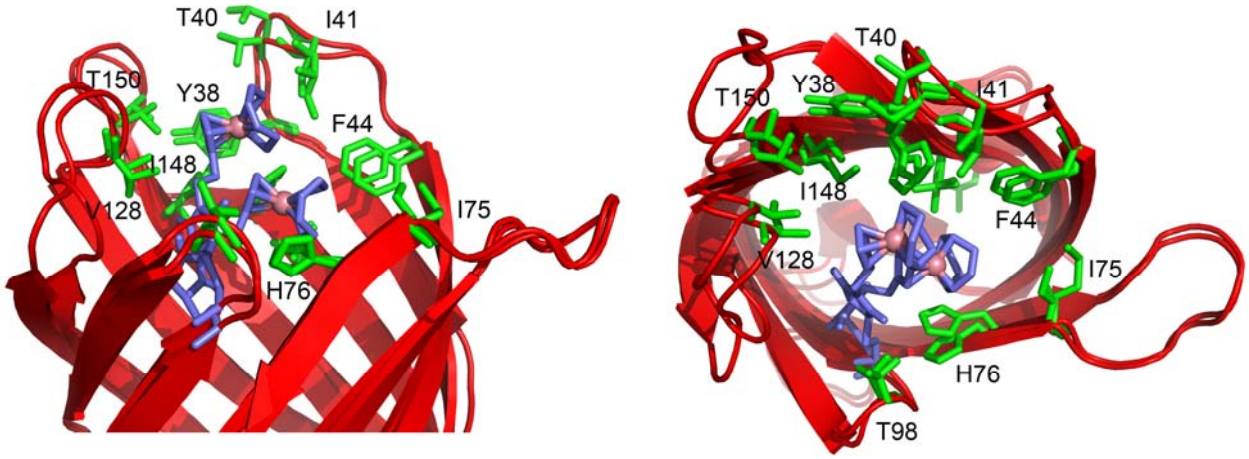


Fig. S6 Close-lying amino acid residues (in green) around the Rh complex in the calculated structures. The hydrophobic Rh complex positions closely with the residues interacting with the hydrophobic heme molecule in the native nitrobindin.

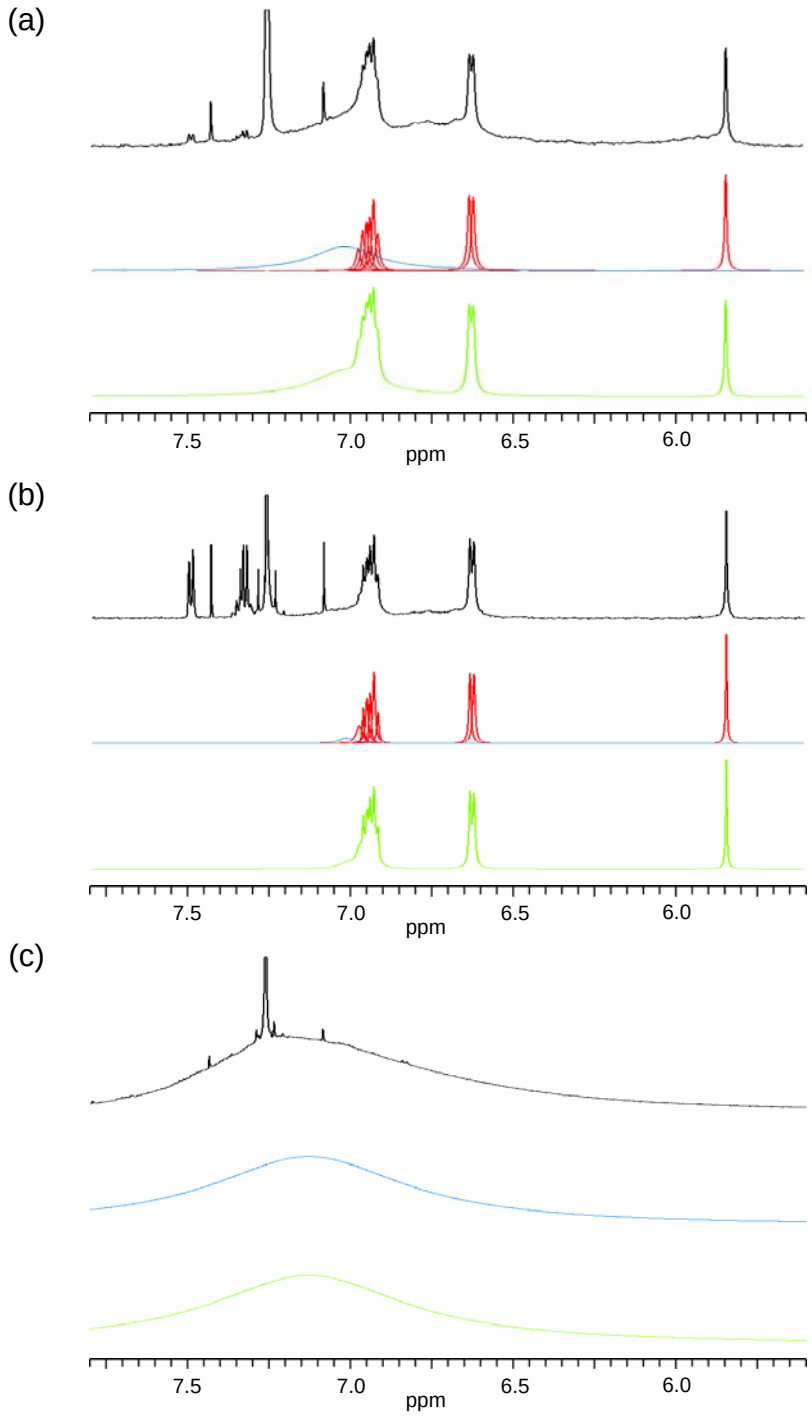


Fig. S7 Determination of stereostructure of PPAs by ^{1}H NMR (600 MHz). The spectrum was recorded by the high-resolution magic angle spinning (HR-MAS) method. (a) PPA prepared by **1,** (b) PPA prepared by Q96C-NB–**1**, and (c) reference sample for *trans*-PPA prepared by a WCl_6 catalyst. The observed spectra (in black), simulated spectra of *cis* (in red) and *trans* (in blue), and the sum of the simulated spectra (in green).

References

S1. J. W. Kang, K. Moseley and P. M. Maitlis, *J. Am. Chem. Soc.*, 1969, **91**, 5970–5977.

S2. A. Altomare, G. Cascarano, C. Giacovazzo and A. Guagliardi, *J. Appl. Cryst.*, 1993, **26**, 343–350.

S3. *Crystal Structure Analysis Package (Version 3.8.2).* , 2007, Rigaku Americas, 9009 TX, USA.

S4. G. M. Scheldrick, *SHELEX-97, Program for the Refinement of Crystal, University of Gottingen, Germany*, 1997.

S5. E. Krieger, T. Darden, S. Nabuurs, A. Finkelstein and G. Vriend, *Proteins: Struct. Funct. Bioinf.*, 2004, **57**, 678–683.

S6. Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang and P. Kollman, *J. Comput. Chem.*, 2003, **24**, 1999–2012.

S7. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

S8. A. Jakalian, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2002, **23**, 1623–1641.