Supporting Information

New reactions and products resulting from alternative interactions between the P450 enzyme and redox partners

Wei Zhang[1], Yi Liu[1,2], Jinyong Yan[1], Shaona Cao[1], Fali Bai[1], Ying Yang[1], Shaohua Huang[1], Lishan Yao[1], Yojiro Anzai[3], Fumio Kato[3], Larissa M. Podust[4], David H. Sherman[5]*, Shengying Li[1]*

[1]CAS Key Laboratory of Biofuels, and Shandong Provincial Key Laboratory of Energy Genetics, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, Shandong 266101, China

[2]University of Chinese Academy of Sciences, Beijing 110039, China

[3]Faculty of Pharmaceutical Sciences, Toho University, Funabashi, Chiba 274-8510, Japan

[4]Center for Discovery and Innovation in Parasitic Diseases, and Department of Pathology, University of California, San Francisco, CA 94158

[5]Life Sciences Institute, Departments of Medicinal Chemistry, Chemistry, and Microbiology and Immunology, University of Michigan, 210 Washtenaw Avenue, Ann Arbor, MI 48109, USA

*Corresponding author

davidhs@umich.edu, lishengying@qibebt.ac.cn

Methods

General experimental procedures. Liquid chromatography mass spectrometry (LC-MS) analysis was carried out on an Agilent 1290-6430 spectrometer using a Waters symmetry column (4.6 mm × 150 mm, RP18) with the water (+ 0.1% formic acid) and acetonitrile (+ 0.1% formic acid) biphasic solvent system. High resolution mass spectra (HRMS) were recorded by a Dionex Ultimate 3000 coupled to a Bruker Maxis Q-TOF. Nuclear magnetic resonance (NMR) spectra were acquired on a Bruker 600 MHz spectrometer. NMR data were processed using MestReNova software. Reverse phase high performance liquid chromatography (RP-HPLC) was employed to purify mycinamicin derivatives using a Waters XBridge 5 μm C18 column with a solvent system of acetonitrile and water supplemented with 0.1% trifluoroacetic acid (TFA). The UV-visible spectra were taken using a Spectrophotometer DU 800 (Beckman Coulter). The codon-optimized gene encoding spinach ferredoxin was synthesized by Genscipt. Primer synthesis and DNA sequencing was performed by Sangon Biotech at Shanghai, China.

Preparation of RhFRED, RhFRED-Fdx, MycG-RhFRED, and MycG-RhFRED-Fdx. The gene encoding RhFRED domain was PCR-amplified using the previously constructed pET28b-*pikC-RhFRED* (*J. Am. Chem. Soc.* **2007**, *129*, 12940-12941) as template and the primer pair of RhFRED_*Nde*I/RhFRED_*Hind*III (Table S1). The gel-cleaned PCR fragment was double digested by *Nde*I and *Hind*III and subsequently inserted into the *Nde*I/*Hind*III pre-treated pET28b to afford pET28b-*RhFRED*. The hybrid gene of *RhFRED-fdx* was prepared by overlap extension PCR (*Nat. Prot.* **2007**, 2, 924-932). Specifically, the gene fragment encoding the FMN domain of RhFRED was amplified from pET28b-*RhFRED* using a pair of primers including RhFRED_*Nde*I and FMN_Fdx. The codon optimized *fdx* gene was rescued from pUC57-*fdx* using the primers Fdx_FMN and Fdx_*Hind*III. The two PCR products with designed overlap sequence were mixed, annealed, extended, and amplified with the RhFRED_*Nde*I/Fdx_*Hind*III primer pair, resulting in the fused gene of *RhFRED-fdx*. This hybrid gene was ligated into pET28b to generate pET28b-*RhFRED-fdx*. To get the P450-reductase fusion genes, the *mycG* gene with stop codon removed was first cloned into pET28b using the *Nde*I/*Eco*RI sites, giving rise to the intermediate plasmid pET28b-*mycG*. Next, the re-prepared *RhFRED* (primers: RhFRED_*Eco*RI and RhFRED_*Hind*III) and *RhFRED-fdx* (primers: RhFRED_*Eco*RI and Fdx_*Hind*III) fragments bearing 5' *Eco*RI and 3' *Hind*III restriction sites were ligated into the *Eco*RI/*Hind*III pre-digested pET28b-*mycG*, resulting in pET28b-*mycG-RhFRED* and pET28b-*mycG-RhFRED-fdx*, respectively. Protein expression and purification of RhFRED, RhFRED-Fdx, MycG-RhFRED, and MycG-RhFRED-Fdx followed

the previously established protocol (*J. Am. Chem. Soc.* **2007**, *129*, 12940-12941). The functional concentration of P450 enzymes was calculated from the CO-bound reduced difference spectrum (*J. Biol. Chem.* **1964**, 239, 2379-2385) using the extinction coefficient ($\epsilon_{450\text{-}490}$) of 91,000 M^{-1} cm^{-1}. The concentration of reductases was determined by Bradford assay (*Anal. Biochem.* **1976**, 72, 248-254) using bovine serum albumin (BSA) as standard.

MycG enzymatic assay. The standard assay containing 1 μM MycG, 100 μM (or 10 and 1 μM for investigation of stoichiometric effect) RhFRED or RhFRED-Fdx, 0.5 mM substrate, and 2.5 mM NADPH in 100 μl of reaction buffer (50 mM NaH_2PO_4, pH 7.3, 10% glycerol) was carried out at 28 ^oC for 2 h and quenched twice with equal volume of chloroform. The organic extracts were combined and dried by N_2 flow and re-dissolved in 100 μl of methanol for LC-MS analysis. For the self-sufficient MycG enzymes, 1 μM MycG-RhFRED or MycG-RhFRED-Fdx was used for replacement of 1 μM MycG and 100 μM reductase in the above standard reaction. Measurements of NADPH coupling efficiency followed Ba *et al.* (*Biotechnol. Bioeng.* **2013**, 110, 2815-2825) with slight modifications.

Detection of formaldehyde using Purpald reagent. The reaction containing 1 μM MycG, 100 μM RhFRED or RhFRED-Fdx, 0.5 mM substrate, and 2.5 mM NADPH in 100 μL reaction buffer (50 mM NaH_2PO_4, pH 7.3, 10% glycerol) was carried out at 28 ^oC. After 2h incubation, 50 μL of 2 M NaOH solution containing 150 mM Purpald reagent (Sigma Aldrich) was mixed with 50 μL reaction solution. The $A_{550\ nm}$ of mixture was measured using the plate reader of Synergy HT (BioTek). However, this assay is not suitable for quantification due to the reactive nature of formaldehyde.

One phase exponential decay kinetics for MycG reactions using M-IV as substrate. The standard assay contained 1 μM MycG, 100 μM RhFRED or RhFRED-Fdx, 0.5 mM M-IV, and 2.5 mM NADPH in 100 μl of reaction buffer (50 mM NaH_2PO_4, pH 7.3, 10% glycerol). For the self-sufficient MycG catalysed reaction, 1 μM MycG-RhFRED or MycG-RhFRED-Fdx was used to replace of 1 μM MycG and 100 μM reductase in the above described standard reaction. The reactions were carried out at 28 ^oC, and quenched by adding equal volume of chloroform at 0min, 5min, 10min, 20min, 60min and 120min. The reaction extracts were subject to HPLC analysis. The areas under curve (AUCs) of the substrate M-IV at different time points were used for fitting the one phase exponential decay curve to deduce the rate constant (*k*). All experiments were performed

in triplicate.

Determination of NADPH coupling efficiencies. By following the previously established procedure (*Biotechnol. Bioeng.* **2013**, 110, 2815-2825), the NADPH coupling efficiencies of different MycG reactions were determined as follow: NADPH consumption was monitored at 340 nm using the plate reader of Synergy HT (BioTek), and calculated with the extinction coefficient of 6.22 mM^{-1} cm^{-1}. The substrate consumption was measured by HPLC when NADPH was depleted. All measurements were performed in duplicate. Coupling efficiency was calculated as the percentage of NADPH used for product formation (i.e. substrate consumption) over the total NADPH consumption.

Measurement of H_2O_2 concentration in MycG reactions. From different MycG reactions, 10μL aliquot of reaction mixture was taken into 50μl deinoized water and thoroughly mixed with 40μL chloroform to quench the reaction at 0min, 5min, 10min, 20min, 60min and 120min, respectively. The mixture was centrifuged and the supernatant was used for the Amplex® Red Hydrogen Peroxide/Peroxidase Assay (Life Technologies). Following the manufacturer's protocol, appropriate amounts of 20 mM H_2O_2 working solution were diluted into 1x Reaction Buffer to generate H_2O_2 serial concentrations ranging from 0 to 10 μM in a 96-well microplate. The absorbance at 560 nm of each sample well was recorded by the Synergy HT plate reader (BioTek), and the standard curve was established. Next, 10μl MycG reaction supernatant was added into 40μl 1x Reaction Buffer, and another 50 μL of 100 μM Amplex® Red reagent and 0.2 U/mL horseradish peroxigenase (HRP) was added into each microplate well to initiate the reaction. The reactions were incubated at room temperature for 30 minutes in darkness, after which $A_{560\ nm}$ values were taken and used to calculate the concentration of H_2O_2 by fitting into the standard curve. All experiments were performed in duplicate.

The effects of catalase, superoxide dismutase (SOD) and ascorbate on MycG reactions. To the standard MycG reaction containing 1 μM MycG, 100 μM RhFRED, and 0.5 mM substrate, 20 U bovine liver catalase (Sigma Aldrich), 2 U SOD from bovine erythrocytes (Sigma Aldrich), 20 mM ascorbate (Sigma Aldrich), and the combination of these three ROS scavengers were added, respectively, prior to the addition of 2.5 mM (final concentration) NADPH in 100 μl of reaction mixture. After 2 h incubation at 28 ^{o}C, the reaction extracts were analyzed by HPLC to calculate the overall conversion percentage and the ratio between demethylated product and oxidized products. To evaluate the dose responses, 0, 2, 4, 20, and 200 U catalase; 0, 0.2, 1, 2, and

20 U SOD; and 0, 1, 5, 10, 100 mM ascorbate were added into the parallel MycG + RhFRED reactions, and the dose response curves were drawn with the overall conversion percentages and the ratios between demethylated product and oxidized products under the different concentrations of ROS scavengers. All experiments were performed in duplicate.

HRMS data for demethylated products. dMe-M-I: $[M + H]^+$, 698.4114 (*calc.* 698.4110 for $C_{36}H_{60}NO_{12}$); d2Me-M-I: $[M + H]^+$, 684.3946 (*calc.* 684.3953 for $C_{35}H_{58}NO_{12}$); dMe-M-II: $[M + H]^+$, 714.4074 (*calc.* 714.4059 for $C_{36}H_{60}NO_{13}$); dMe-M-III: $[M + H]^+$, 668.4019 (*calc.* 668.4004 for $C_{35}H_{58}NO_{11}$); d2Me-M-III: $[M + H]^+$, 654.3859 (*calc.* 654.3848 for $C_{34}H_{56}NO_{11}$); dMe-M-IV: $[M + H]^+$, 682.4161 (*calc.* 682.4160 for $C_{36}H_{60}NO_{11}$); dMe-M-V: $[M + H]^+$, 698.4116 (*calc.* 698.4110 for $C_{36}H_{60}NO_{12}$).

^{1}H NMR data for dMe-M-IV in MeOD. 7.11 (1H, dd, *J* = 11.0,14.9 Hz, H-11); 6.64 (1H, dd, *J* = 9.2, 15.3 Hz, H-3); 6.47 (1H, d, *J* = 14.8 Hz, H-10); 6.29 (1H, dd, *J* = 11.0, 14.8 Hz, H-12); 6.12 (1H, dd, *J* = 9.2, 15.3 Hz, H-13); 5.94 (1H, d, *J* = 15.8 Hz, H-2); 4.93 (1H, dt, *J* = 2.9, 9.9 Hz, H-15); 4.61 (1H, d, *J* = 7.9 Hz, H-1’’); 4.36 (1H, d, *J* = 6.9 Hz, H-1’); 4.02 (1H, ABs, *J* = 4.9, 9.4 Hz, H-21); 3.78 (1H, m, H-3’’); 3.68(1H, m, H-21); 3.62(1H, m, H-4’); 3.59 (3H, s, OMe-3’’); 3.55 (1H, m, H-5’’); 3.53 (3H, s, OMe-2’’); 3.52 (1H, m, H-5’); 3.35(1H, m, H-2’); 3.31(1H, m, H-5); 3.22 (1H, m, H-4’’); 3.08 (1H, dd, *J* = 2.9, 8.1 Hz, H-2’’); 2.84(1H, m, H-4); 2.72 (3H, s, NHMe); 2.57 (1H, m, H-14); 2.54 (1H, m, H-3’); 2.52 (1H, m, H-8); 1.96 (1H, m, H-16); 1.75(1H, m, H-4’); 1.64 (2H, m, H-7); 1.60 (1H, m, H-6); 1.54(1H, m, H-16); 1.31(3H, d, *J* = 6.4 Hz, H-20); 1.30 (3H, d, *J* = 6.4 Hz, H-6’’); 1.25 (3H, d, *J* = 7.0 Hz, H-6’); 1.18(3H, d, *J* = 7.0 Hz, H-18); 1.01 (3H, d, *J* = 7.0 Hz, H-19); 0.97 (3H, t, *J* = 7.6, 7.6 Hz, H-17).

The DNA sequences encoding Fdx, RhFRED, and RhFRED-Fdx

1. The spinach ferredoxin gene - *fdx*, codon optimized for *E. coli*

ATGGCGGCCTATAAAGTGACCCTGGTTACGCCGACCGGCAACGTCGAATTTCAGTGCCCGGATGA
CGTGTACATTCTGGATGCAGCTGAAGAAGAAGGTATCGACCTGCCGTATAGCTGTCGTGCGGGCTC
TTGCAGTTCCTGTGCCGGTAAACTGAAAACGGGCTCACTGAATCAAGATGACCAGTCGTTCCTGG
ATGACGATCAAATTGACGAAGGTTGGGTTCTGACCTGCGCAGCTTACCCGGTCTCTGATGTGACG
ATCGAAACCCATAAAGAAGAAGAACTGACGGCATAA

2. The gene encoding RhFRED (cyan: the natural linker between the P450 domain and the reductase domain; yellow: FMN domain; green: [2Fe-2S] domain)

gtgctgcaccggcatcaaccggtcaccatcggagaacccgccgcccgggcggtgtcccgcaccgtcaccgtcgagcgcctggaccggatcgccgacgac
gtgctgcgcctcgtcctgcgcgacgccggcggaaagacattacccacgtggactcccggcgcccatatcgacctcgacctcggcgcgctgtcgcgccagta
ctccctgtgcggcgcgcccgatgcgccgagctacgagattgccgtgcacctggatcccgagagccgcggcggttcgcgctacatccacgaacagctcgagg
tgggaagcccgctccggatgcgcggccctcggaaccatttcgcgctcgaccccggcgccgagcactacgtgttcgtcgccggcggcatcggcatcacccca
gtcctggccatggccgaccacgcccgcgcccgggggtggagctacgaactgcactactgcggccgaaaccgttccggcatggcctatctcgagcgtgtcgc
cgggcacggtgaccgggccgccctgcacgtgtccgaggaaggcacccggatcgacctcgccgccctcctcgccgagcccgcccccggcgtccagatctac
gcgtgcgggcccgggcggctgctcgccggactcgaggacgcgagccggaactggcccgacggggcgctgcacgtcgagcacttcacctcgtccctcgcg
gcgctcgatccggacgtcgagcacgccttcgacctcgaactgcgtgactcggggctgaccgtgcgggtcgaacccacccagaccgtcctcgacgcgttgcg
cgccaacaacatcgacgtgcccagcgactgcgaggaaggcctctgcggctcgtgcgaggtcgccgtcctcgacggcgaggtcgaccatcgcgacacggtg
ctgaccaaggccgagcgggcggcgaaccggcagatgatgacctgctgctcgcgtgcctgtggcgaccggctggccctgcgactctga

3. The hybrid gene encoding RhFRED-Fdx

gtgctgcaccggcatcaaccggtcaccatcggagaacccgccgcccgggcggtgtcccgcaccgtcaccgtcgagcgcctggaccggatcgccgacgac
gtgctgcgcctcgtcctgcgcgacgccggcggaaagacattacccacgtggactcccggcgcccatatcgacctcgacctcggcgcgctgtcgcgccagta
ctccctgtgcggcgcgcccgatgcgccgagctacgagattgccgtgcacctggatcccgagagccgcggcggttcgcgctacatccacgaacagctcgagg
tgggaagcccgctccggatgcgcggccctcggaaccatttcgcgctcgaccccggcgccgagcactacgtgttcgtcgccggcggcatcggcatcacccca
gtcctggccatggccgaccacgcccgcgcccgggggtggagctacgaactgcactactgcggccgaaaccgttccggcatggcctatctcgagcgtgtcgc
cgggcacggtgaccgggccgccctgcacgtgtccgaggaaggcacccggatcgacctcgccgccctcctcgccgagcccgcccccggcgtccagatctac
gcgtgcgggcccgggcggctgctcgccggactcgaggacgcgagccggaactggcccgacggggcgctgcacgtcgagcacttcacctcgtccctcgcg
gcgctcgatccggacgtcgagcacGCGGCCTATAAAGTGACCCTGGTTACGCCGACCGGCAACGTCGAATTTC
AGTGCCCGGATGACGTGTACATTCTGGATGCAGCTGAAGAAGAAGGTATCGACCTGCCGTATAGC
TGTCGTGCGGGCTCTTGCAGTTCCTGTGCCGGTAAACTGAAAACGGGCTCACTGAATCAAGATGA
CCAGTCGTTCCTGGATGACGATCAAATTGACGAAGGTTGGGTTCTGACCTGCGCAGCTTACCCGG
TCTCTGATGTGACGATCGAAACCCATAAAGAAGAAGAACTGACGGCATAA

Table S1. Primers used in this study

Primer name	**Primer sequence** (5'→3')	**Function of underlined bases**
RhFRED_*Nde*I	CAGATTCATATGGTGCTGCACCGCCATCAA	*Nde*I site
RhFRED_*Hind*III	CAATGAAGCTTTCAGAGTCGCAGGGCCAG	*Hind*III site
FMN_Fdx	ACTTTATAGGCCGCGTGCTCGACGTCCG	*fdx* overhang
Fdx_FMN	CGGACGTCGAGCACGCGGCCTATAAAGT	FMN domain overhang
Fdx_*Hind*III	GCCTGAAGCTTTTATGCCGTCAGTTCTTC	*Hind*III site
MycG_*Nde*I	CGGTCATATGACTTCAGCTGAACCTAGGGCG	*Nde*I site
MycG_*Eco*RI	GGCAGAATTCCCACACGACCGGCAGCTCGAG	*Eco*RI site
RhFRED_*Eco*RI	CAGGGAATTCGTGCTGCACCGCCATCAACCG	*Eco*RI site

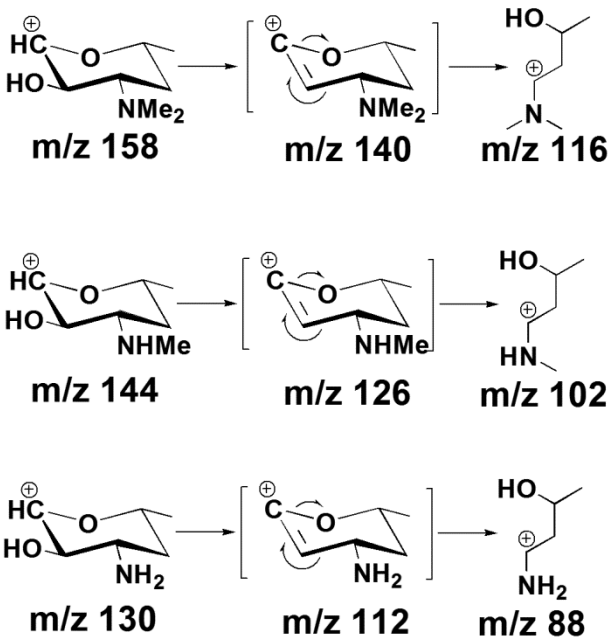

Figure S1. Explanation of the secondary mass spectra of diglycosylated mycinamicins (top), *N*-mono-demethylated mycinamicins (middle), and *N*-double-demethylated mycinamicins (bottom).

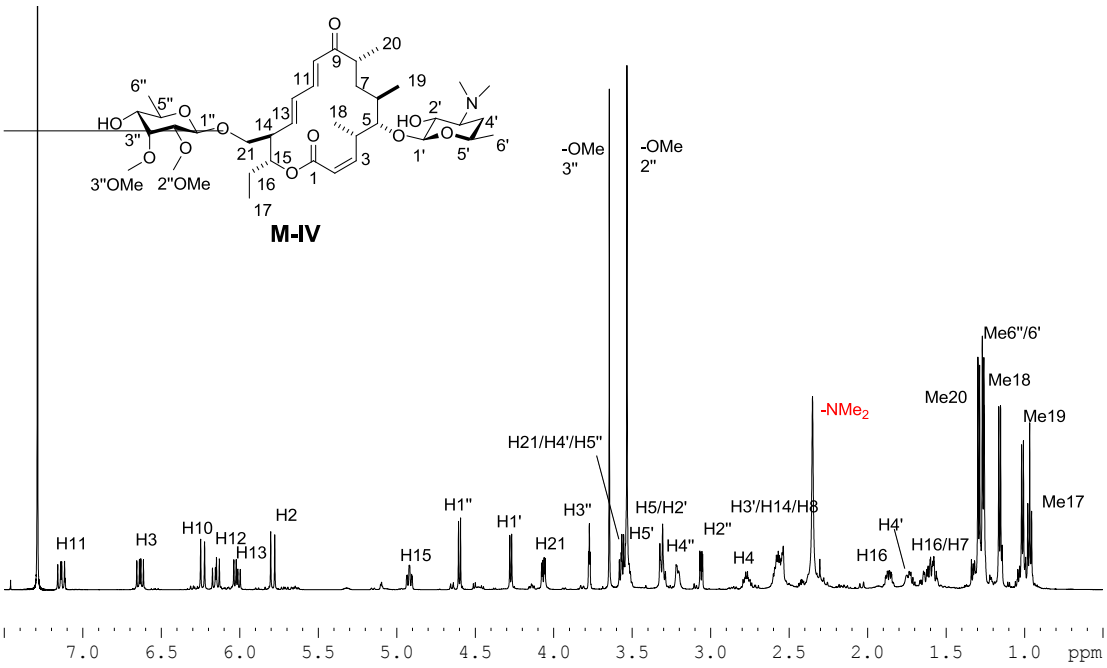

Figure S2. ^{1}H NMR spectrum of M-IV in $CDCl_3$.

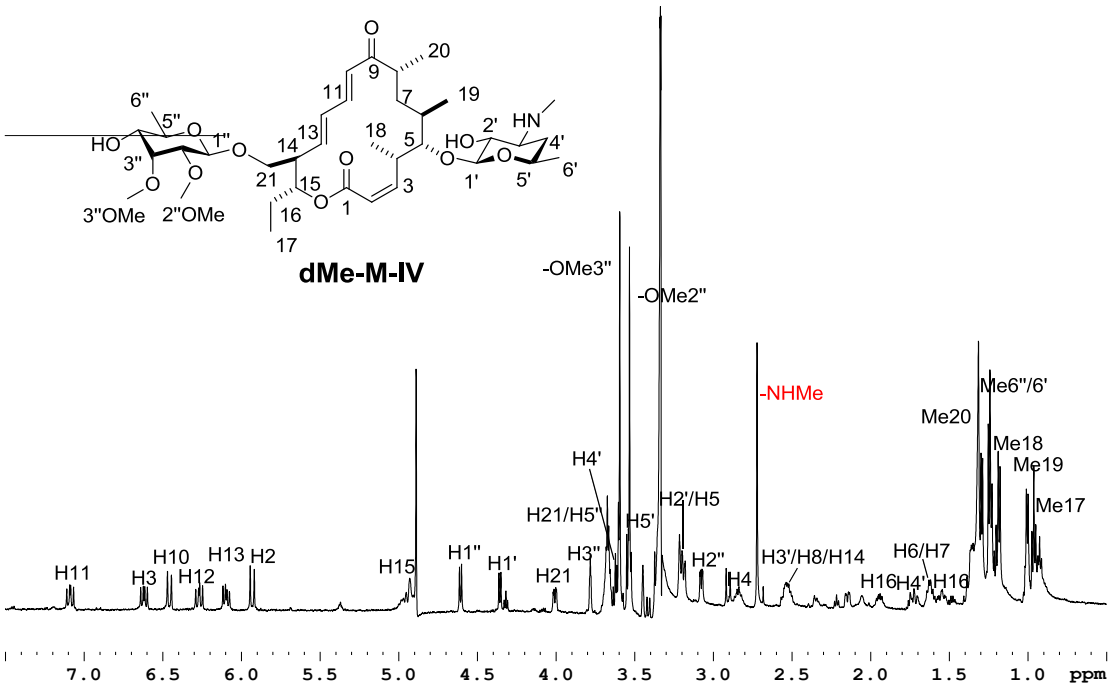

Figure S3. ^{1}H NMR spectrum of dMe-M-IV in MeOD.

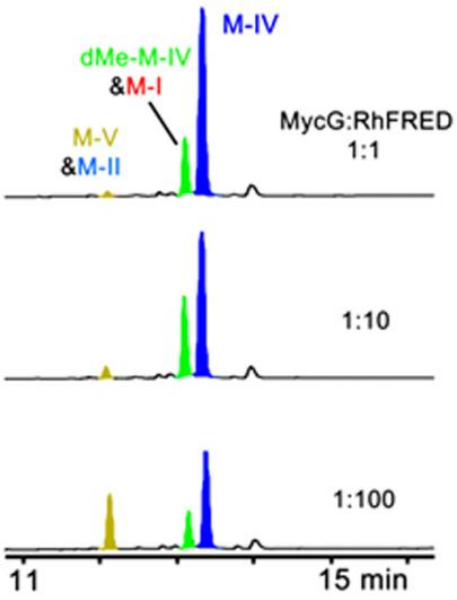

Figure S4. The stoichiometric effect of RhFRED on the activity of MycG against M-IV (2 h reaction)

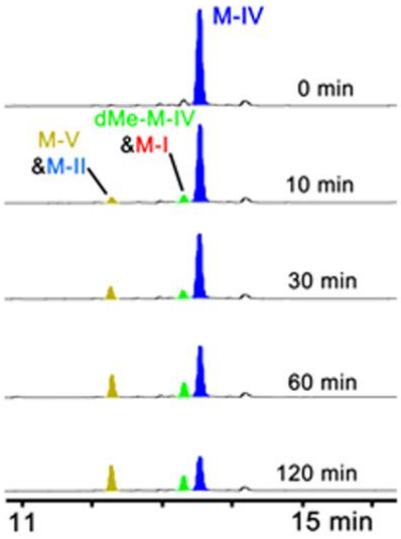

Figure S5. The time course of M-IV oxidation(s) and demethylation catalysed by MycG plus RhFRED (1:100).

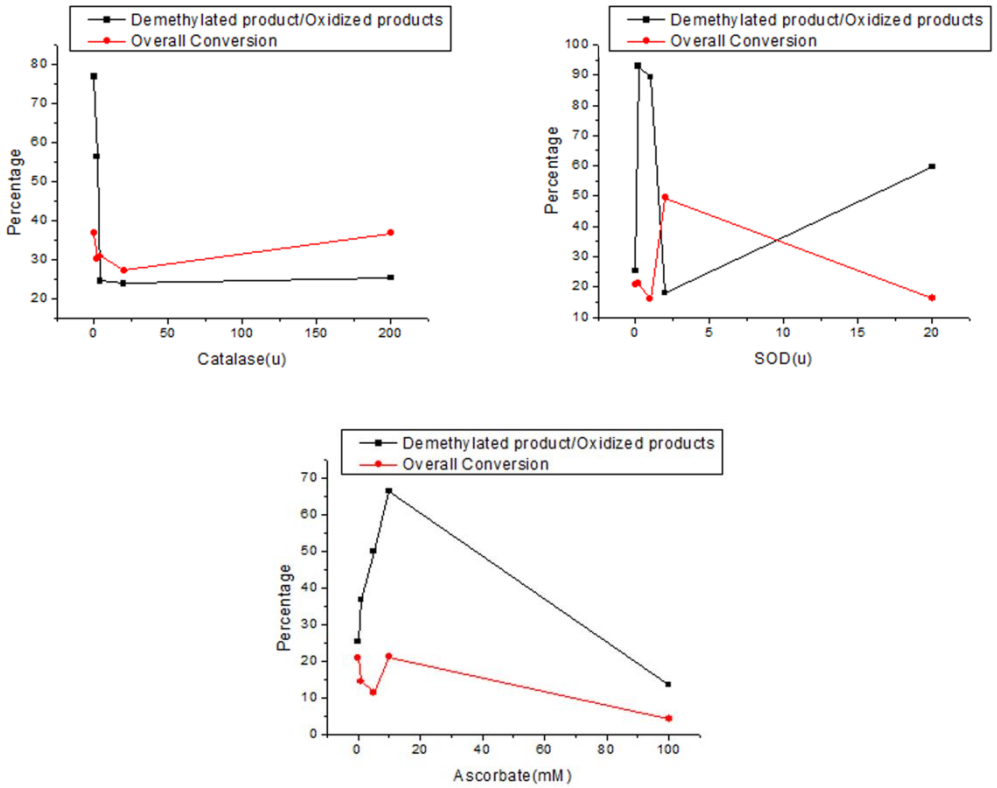

Figure S6. Dose response of the MycG + RhFRED reaction toward catalase, superoxide dismutase (SOD), and ascorbate

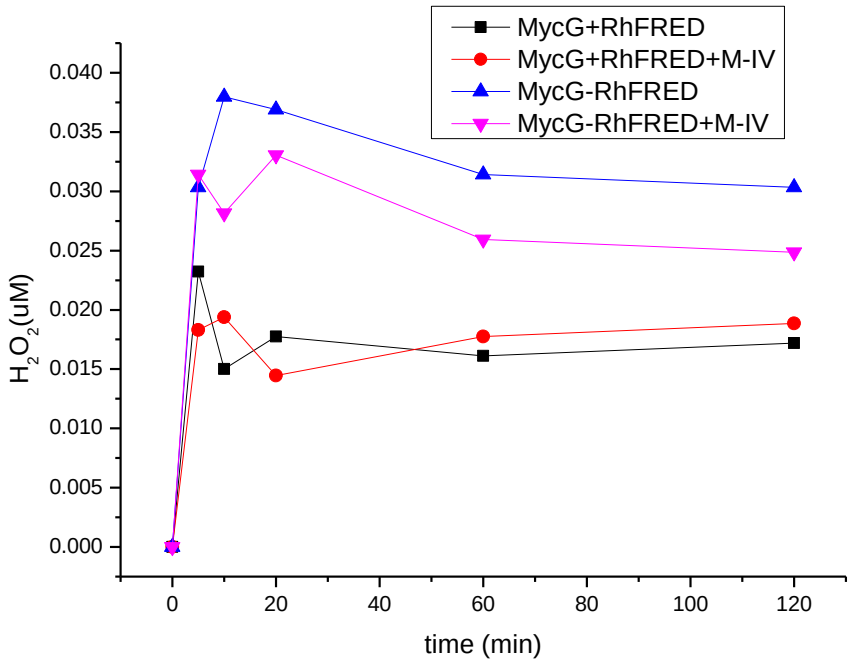

Figure S7. H_2O_2 production by different MycG systems in the presence and absence of the substrate M-IV

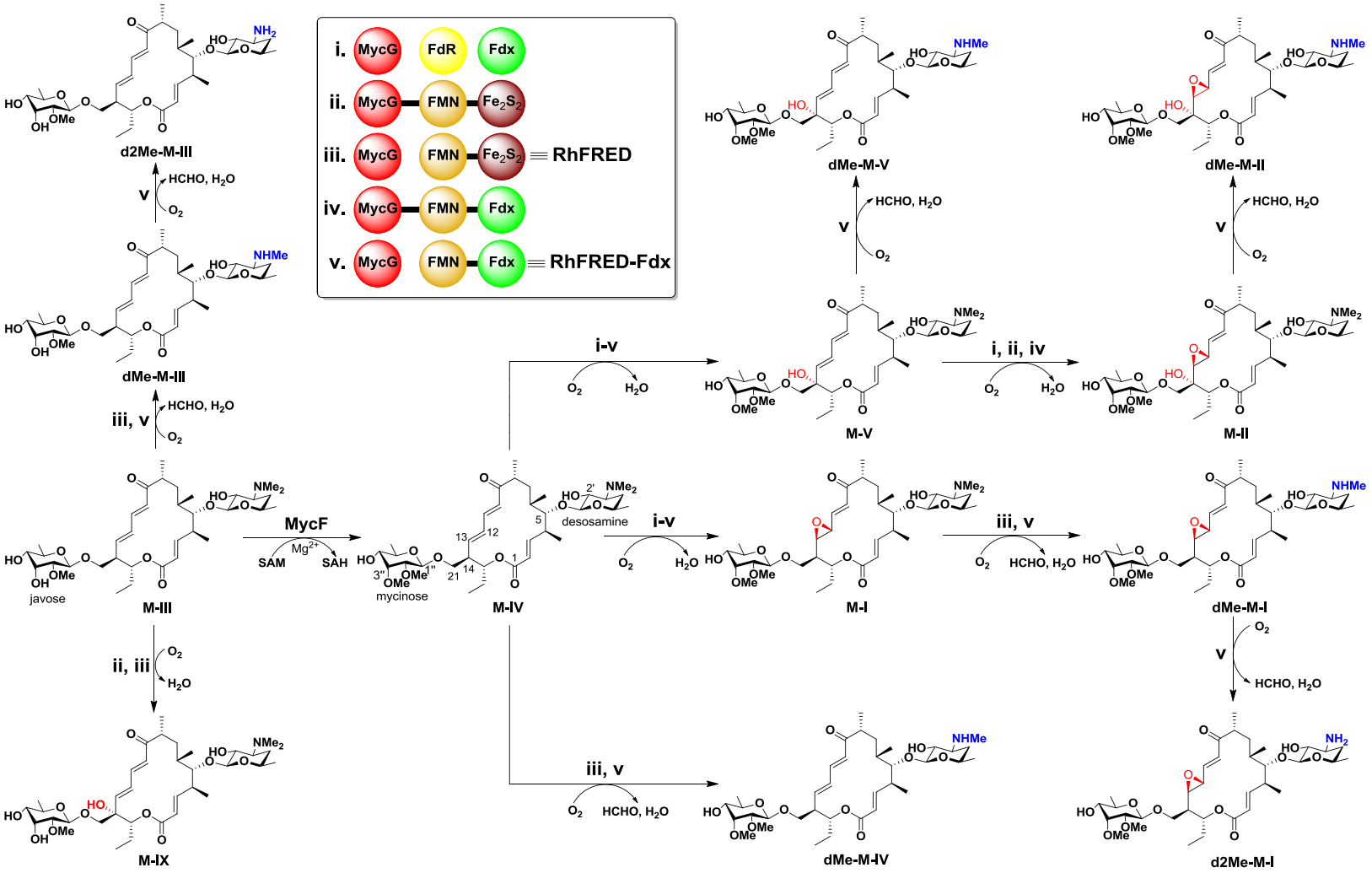

Figure S8. Oxidation and demethylation of mycinamicins catalysed by MycG using five different redox systems. (i) MycG + spinach FdR + spinach Fdx (*Chem. Biol.* 2008, 15, 950-959), (ii) MycG-RhFRED, (iii) MycG + RhFRED, (iv) MycG-RhFRED-Fdx, and (v) MycG + RhFRED-Fdx. The introduced hydroxyl and epoxy groups are labelled in red. The demethylated groups are highlighted in blue. (Abbreviations: M-I-V, mycinamicin I-V; M-IX, mycinamicin IX (*FEMS Microbiol. Lett.* 2010, 304, 148-156); dMe-M-I-V, *N*-demethylated mycinamicin I-V; d2Me-M-I, *N*-didemethylated mycinamicin I; d2Me-M-III, *N*-didemethylated mycinamicin III.)