Supplemental Information

In vitro reconstitution of the biosynthetic pathway of 3-hydroxypicolinic acid

Xuan Yun‡[1], Qian Zhang‡[1], Meinan Lv‡[1], Hai Deng[3], Zixin Deng[1,*], Yi Yu[1,2,*]

[1]Key Laboratory of Combinatorial Biosynthesis and Drug Discovery (Ministry of Education), School of Pharmaceutical Sciences, Wuhan University, 185 East Lake Road, Wuhan 430071, P. R. China. Email: yu_yi@whu.edu.cn; zxdeng@sjtu.edu.cn

[2]Hubei Key Laboratory of Industrial Microbiology, Hubei Provincial Cooperative Innovation Center of Industrial Fermentation, Key Laboratory of Fermentation Engineering (Ministry of Education), Hubei University of Technology, Wuhan 430068, P. R. China

[3]Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Aberdeen AB24 3UE, Scotland, UK

‡ These authors contributed equally to this work

* Corresponding authors

Experimental Procedures

Chemicals, Biochemicals and culture conditions

All chemical reagents and anhydrous solvents were purchased from commercial sources unless otherwise specified. Kanamycin, apramycin, thiostrepton, chloramphenicol, nalidixic acid, benzoic acid, pyridine acid, 3-hydroxypicolinic acid (3-HPA), L-lysine, 2-oxobutyric acid, NADH, riboflavin, FMN, FAD, and pyridoxal phosphate (PLP) were purchased from Sigma-Aldrich Co. Ltd (USA). Restriction enzymes were purchased from Takara Biotechnology (Dalian, China). Primers were synthesized at TSINGKE Biological Technology (Wuhan, China).

E. coli DH10B and *E. coli* ET12567/pUZ8002 were cultured in Luria-Bertani (LB) or LB agar medium (tryptone 1%, yeast extract 0.5%, NaCl 1%, Agar 1.2%) at 37 °C for cloning, proteins overexpression and intergeneric conjugation, respectively. *Streptomyces lividans* TK24 was cultured in TSBY medium (tryptone soya broth 3%, yeast extract 0.5%) as seed liquid, in MS medium (soybean 2%, D-mannitol 2%, Agar 2%) at 28 °C for spores production, in YEME medium (glucose 1%, tryptone 0.5%, yeast extract 0.3%, malt extract 0.3%, sucrose 10.3%) for proteins overexpression.

Construction of plasmids for overexpressing N-terminally hexa-His-tagged BomA/B/D/E in *Streptomyces*

The *bomA/B/D/E* genes were amplified from the *Streptomyces sp*. NRRL 12068 genomic DNA using primer pairs pGM1190-BomA-F and pGM1190-BomA-R, pGM1190-BomB-F and pGM1190-BomB-R, pGM1190-BomD-F and pGM1190-BomD-R, and pGM1190-BomE-F and pGM1190-BomE-R, respectively (Table S2). The PCR products were digested with NdeI and HindIII, purified using a QIAGEN PCR purification kit, and inserted into the same restriction site of the expression vector pGM1190. Chemically competent *E. coli* DH5α cells were transformed with the ligation mixture and plated on LB-agar containing apramycin (50 μg/mL) to screen for positive clones (Table S1), which were confirmed by DNA sequencing.

Overexpression of BomA/B/D/E in *S. lividans* TK24

E. coli ET12567/pUZ8002 containing pWDY1201, pWDY1202, pWDY1203 and

pWDY1204 were acquired by transformation and grown in 2 mL LB medium supplemented with 50 μg/mL kanamycin, 25 μg/mL chloramphenicol and 50 μg/mL apramycin at 37 °C shaker until an OD_{600}~0.6-0.8 was reached. The *E. coli* cells were washed twice to remove antibiotics. 500 μL cells were added to 0.5 mL pretreated *S. lividans* TK24 spores. After being mixed and spun briefly, most of the supernatant was poured off and the residual mixture was plated out on MS agar + 10 mM $MgCl_2$ at 28 °C for 16-20 h incubation. Then the plate was overlaid with 1 mL water containing 0.5 mg nalidixic acid and 1 mg apramycin and the potential exconjugants were selected for further verification after 3-5 days incubation at 28 °C.

S. lividans TK24 containing BomA/BomB/BomD/BomE overexpression plasmids were grown in 50 mL TSBY medium supplemented with 50 μg/mL apramycin at 28 °C with shaking at 200 rpm until an OD_{600}~0.6-0.8 was reached. The seed culture was transferred to 500 mL YEME medium supplemented with 50 μg/mL apramycin at 28 °C with shaking at 200 rpm until an OD_{600}~0.6-0.8 was reached. Protein expression was induced by addition of 0.5 mM (final concentration) thiostrepton with further cultivation at 28 °C for 72 h.

Protein purification of BomA/B/D/E

The *E. coli* or *Streptomyces* cell pellet collected by centrifugation was resuspended in 40 mL ice-cold lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, and 10% (*v/v*) glycerol, pH 8.0), and further disrupted by a Nano Homogenize Machine (ATS Engineering Inc, AH100B). For the purification of BomE, 10 μM pyridoxal phosphate was supplemented in the above lysis buffer. Cell debris was removed via centrifugation at 12000 rpm for 30 min at 4 °C. The supernatant was loaded onto 2 mL Ni-NTA-affinity column (GE Healthcare) pre-equilibrated with the lysis buffer. The desired elution fractions were combined and concentrated using a Centrifugal Filter Units (Milipore, 10,000 MWCO), and the concentrated protein solution was desalted using a PD-10 Column (GE Healthcare) pre-equilibrated with the elution buffer (20 mM Tris-HCl, 100 mM NaCl, and 10% (*v/v*) glycerol, pH 8.0). The protein fraction was collected and concentrated, analyzed by SDS-PAGE (12% Tris-glycine gel), and was used directly for *in vitro* assay or stored at −80 °C upon further use. Protein concentration was

determined using a Bradford assay kit (Promega) with bovine serum albumin (BSA) as the standard.

LC-MS parameters of Enzymatic Reaction Assays

LC-MS analysis was carried out in positive mode by using a Thermo Scientific LTQ XL Orbitrap mass spectrometer equipped with a Thermo Scientific Accela 600 pump (Thermo Fisher Scientific Inc). LC conditions for *in vitro* reaction assays: C18 column was pre-equilibrated with 95%A/5%B and developed at a flow rate of 0.8 mL/min, 0-5 min, constant with 95%A/5%B; 5-13 min, a linear gradient from 95%A/5%B to 80%A/20%B; 13-15 min, a linear gradient from 80%A/20%B to 5%A/95%B; 15-20 min, constant with 5%A/95%B; 20-22 min, a linear gradient to 95%A/15%B; 22-27 min, constant with 5%A/95%B; UV monitored at 260 nm and 450 nm. Solvent A was 10 mM heptafluorobutyric acid (HFBA) in H_2O and solvent B was 0.1% formic acid in CH_3CN. All MS analysis parameters were set as 45 V capillary voltage, 45 °C capillary temperature, auxiliary gas flow rate 10 arbitrary units, sheath gas flow rate 40 arbitrary units, 3.5 kV spray voltage, and 50-1000 Amu mass range.

Enzymatic Assay of BomA

The BomA reaction was performed by incubating 2 μM BomA with 1 mM NADH, and 1 mM FMN in 100 mM MOPS (pH 7.83) in a 150 μL total volume. The reaction was carried out at 30 °C for 0 h, 0.5 h, 1 h, 3 h, 8 h and 24 h separately in anaerobic environment. The reaction mixture was measured by UV-Vis absorption spectra immediately.

Enzymatic Assay of BomE

The BomE reaction was performed by incubating 2 μM BomE with 5 mM L-lysine, 100 mM 2-oxobutyric acid, 100 μM pyridoxal phosphate, and 100 mM MOPS (PH 7.83) in a 100 μL total volume. The reaction was carried out at 30 °C for 3 h and quenched by equal volume methanol to precipitate proteins. The clarified supernatant was subjected to LC-ESI-HRMS analysis. The boiled BomE was used in the negative control reaction.

***In vitro* assay of BomA/B/D/E one-pot reaction**

The one-pot reaction was performed by incubating 2 μM enzymes (BomA, BomB, BomD, BomE) with 5 mM L-lysine, 100 mM 2-ketobutyric acid,100 μM PLP, 1 mM NADH，100 μM FAD, 100 μM FMN and 100 mM MOPS (pH 7.83) in a 100 μL total volume. The mixture was incubated at 30 °C for 3 h and quenched by equal volume methanol to precipitate proteins. The clarified supernatant was subjected to LC-ESI-HRMS analysis. The boiled enzymes were used in the negative control reaction. The L–lysine and 2-ketobutyric acid were substituted to pyridine acid in the substrate tolerance reactions (Figure 3(vi and vii)).

Table S1. Strains and plasmids used in this study

Strain/plasmid	Relevant genotype/comments	Source/Ref.
Plasmids		
pWDY1201	BomA overexpression construction	This study
pWDY1202	BomB overexpression construction	This study
pWDY1203	BomD overexpression construction	This study
pWDY1204	BomE overexpression construction	This study
Strains		
Streptomyces		
WDY1201	BomA overexpression strain	This study
WDY1202	BomB overexpression strain	This study
WDY1203	BomD overexpression strain	This study
WDY1204	BomE overexpression strain	This study

Table S2. PCR primers used in this study

Primer	Sequence *	Function
pGM1190-BomA-F	TGCCGCGCGGCAGCCATATGACGTCGACACCCACTTCCCAGCAGGCGCCGCTG	BomA overexpression
pGM1190-BomA-R	TCGAGTGCGGCCGCAAGCTTTCATGTCAGGGCGTGGAAGGCCCGGCGGTACCAC	BomA overexpression
pGM1190-BomB-F	TGCCGCGCGGCAGCCATATGGCGCAGTGGGACGCGGAGGTCGCGGTG	BomB overexpression
pGM1190-BomB-R	TCGAGTGCGGCCGCAAGCTTTCAGCCGAAGCGGTTCGGGGACATG	BomB overexpression
pGM1190-BomD-F	TGCCGCGCGGCAGCCATATGAAATTCGGCGTGAATTTCTTTCCGGTG	BomD overexpression
pGM1190-BomD-R	TCGAGTGCGGCCGCAAGCTTTTATTATCACTCCGTGAAGGAGGGGGCCAC	BomD overexpression
pGM1190-BomE-F	TGCCGCGCGGCAGCCATATGCGCGGGGCCTACAAAGTTCCGTACGCGAGTCGCG	BomE overexpression
pGM1190-BomE-R	TCGAGTGCGGCCGCAAGCTTTCAATGAGCGGCATGCGAGGTCTCCTTGATGCTCGG	BomE overexpression

***** Restriction sites (NdeI and HindIII) are underlined.

Table S3. Accession number of proteins used for bioinformatics analysis in this study

Accession number	Source	Enzymes
AGN74901	*Streptomyces griseoviridis*	SgvS
AGN74878	*Streptomyces griseoviridis*	SgvF
AGN74881	*Streptomyces griseoviridis*	SgvL
3HZL A	*Streptomyces tendae*	NikD
CAB46533	*Streptomyces tendae*	NikC
AEF33084	*Streptomyces pyridomyceticus*	PyrK
AEF33077	*Streptomyces pyridomyceticus*	PyrD
AEF33076	*Streptomyces pyridomyceticus*	PyrC
AEF33097	*Streptomyces pyridomyceticus*	PyrB
BAF50714	*Streptomyces virginiae*	VirM
BAF50713	*Streptomyces virginiae*	VirN
BAB83671	*Streptomyces virginiae*	VisA
AAA83566	*Streptomyces pristinaespiralis*	SnaC
CBW45735	*Streptomyces pristinaespiralis*	SnaO
CBW45733	*Streptomyces pristinaespiralis*	SnaQ
CBW45759	*Streptomyces pristinaespiralis*	HpaA
2D36_A	*Sulfurisphaera tokodaii str. 7*	HpaC
ARZ67054	*Streptomyces albireticuli*	ActVB
O87008	*Burkholderia cepacia*	TftC
ABS30826	*Rhodococcus erythropolis*	PheA2
1LUC_A	*Vibrio harveyi*	LuxA
1LUC_B	*Vibrio harveyi*	LuxB
ABX76821	*Photobacterium leiognathi subsp. mandapamensis*	LuxF

Table S4. Source of the proteins in Figure 1 and homologous proteins with the BomA/B/D/E in the A33853 biosynthetic pathway

source	Antibiotic	Identity with BomA	Identity with BomB	Identity with BomD	Identity with BomE
Streptomyces griseoviridis	Viridogrisein	_	SgvS (64%)	SgvF (74%)	SgvL (72%)
Streptomyces tendae	Nikkomycin D	_	NikD (26%)	_	NikC (59%)
Streptomyces pyridomyceticus	Pyridomycin	PyrK (38%)	PyrD (56%)	PyrC (66%)	PyrB (64%)
Streptomyces virginiae	Virginiamycin	wp033226799(34%)	VirM (55%)	VirN (63%)	VisA (63%)
Streptomyces pristinaespiralis	Pristinamycin	SnaC (51%)	SnaO (53%)	SnaQ (63%)	HpaA (65%)

Figure S1. Genetic organization of the A33853 biosynthetic gene cluster. The proteins involved in this study are labeled in blue color

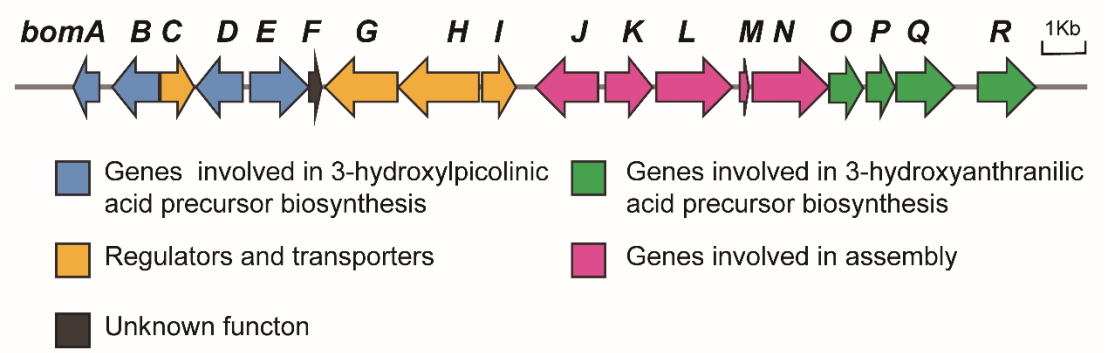


Figure S2. SDS-PAGE and UV-Vis spectra analysis of the purified proteins.

(A) Purified 6× His-tagged BomA expressed in *S. lividans* TK24 and its UV-Vis spectra.
(B) Purified 6× His-tagged BomB expressed in *S. lividans* TK24 and its UV-Vis spectra.
(C) Purified 6× His-tagged BomD expressed in *S. lividans* TK24 and its UV-Vis spectra.
(D) Purified 6× His-tagged BomE expressed in *S. lividans* TK24 and its UV-Vis spectra.

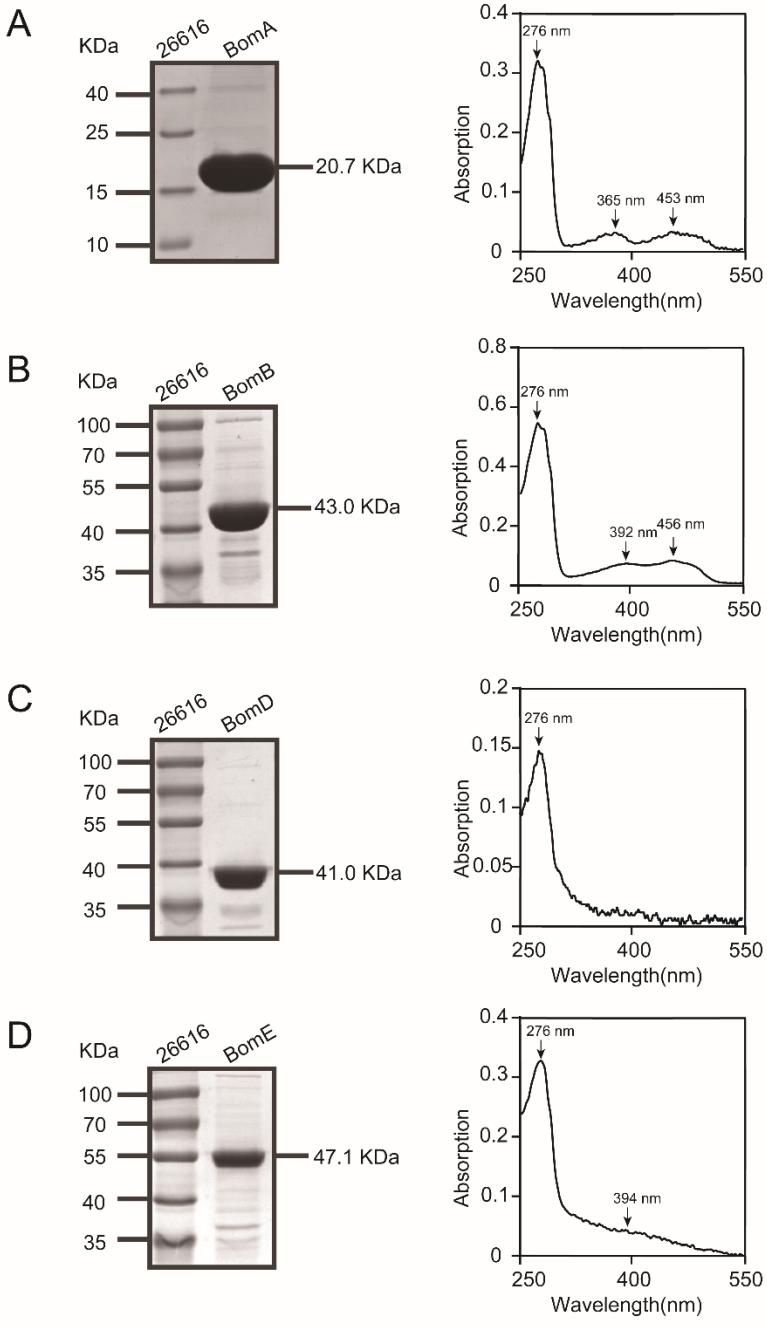


Figure S3. Sequence alignment of FAD dependent oxidoreductase BomB with its homologues NikD[1], SgvS[2], VirM[3], SnaO[4]. Secondary structure assignment based on the NikD structure is shown above the alignment as helices for α-helices and arrows for β-strands. Strictly conserved residues among the proteins are indicated in white on a red background. Relatively conserved residues among the proteins are indicated in red. The triangle represents the FAD covalent binding site and stars represent active sites as electron donor or receptor. The figure was generated from CLUSTALW.

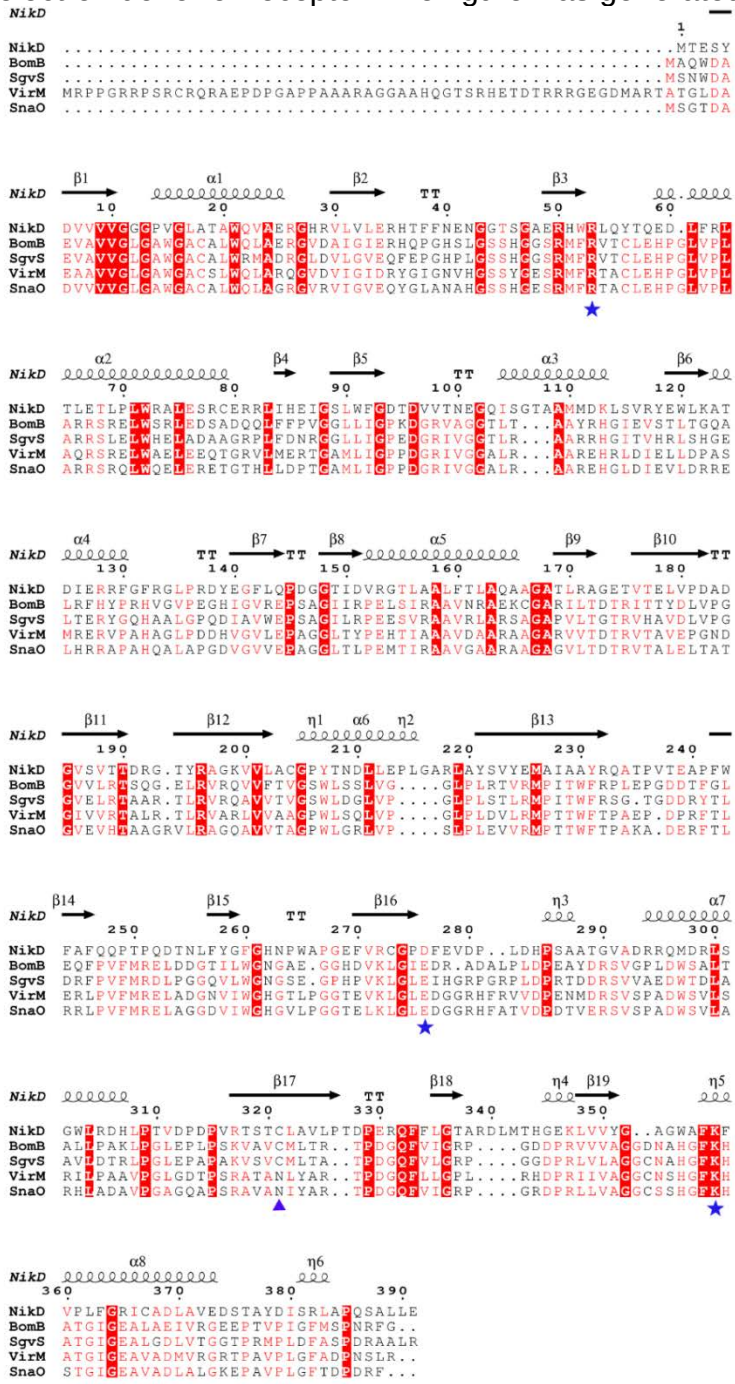


Figure S4. Determination of the flavin content of BomA. UV-Vis spectra of the BomA reactions in the present of (A) NADPH and Riboflavin (for 3 h), (B) NADPH and FAD (for 3 h), and (C) NADPH and FMN (for 3 h). The arrows point to the characteristic UV absorption wavelength of riboflavin, FAD, and FMN (450 nm).

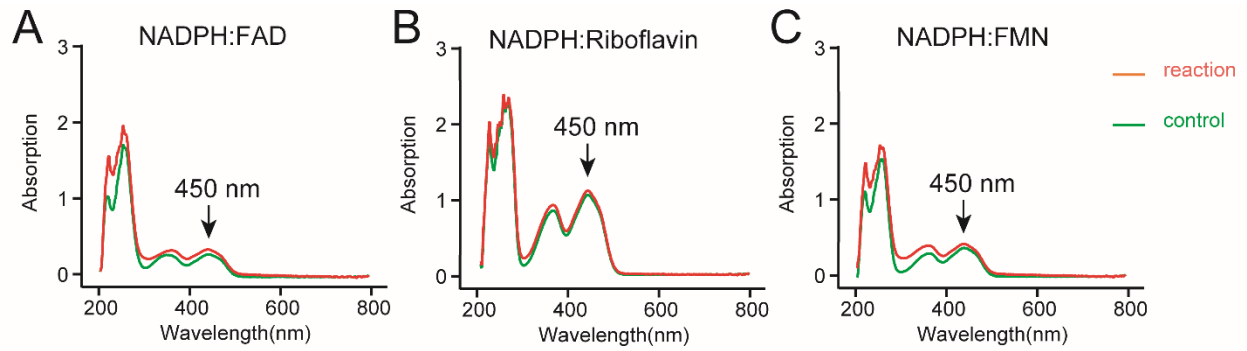


Figure S5. Sequence alignment of NAD(P)H-flavin oxidoreductases BomA with its homologues HpaC[5], ActVB[6], TftC[7], PheA2[8]. Secondary structure assignment based on the structure of HpaC is shown above the alignment as helices for α-helices and arrows for β-strands. Strictly conserved residues among the proteins are indicated in white on a red background. Relatively conserved residues among the proteins are indicated in red. The triangle represents the FMN binding site and stars represent NADH binding sites. The figure was generated from CLUSTALW[9].

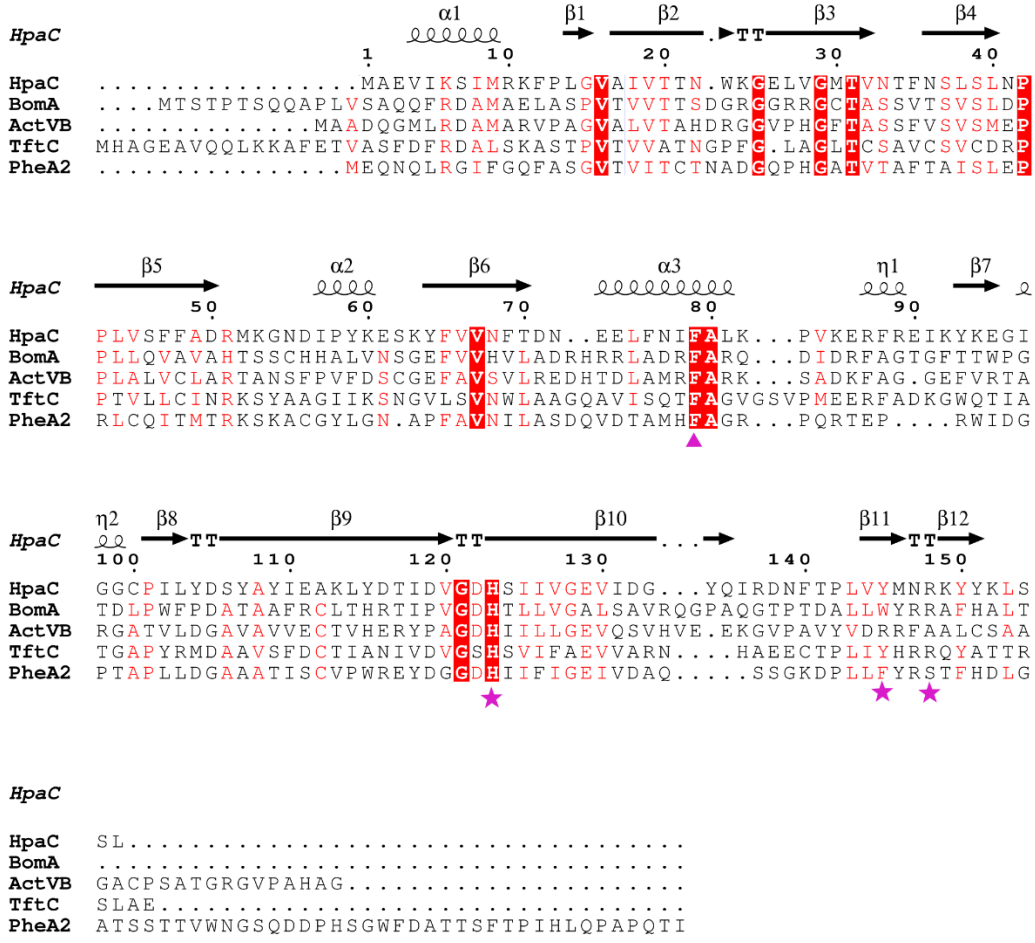
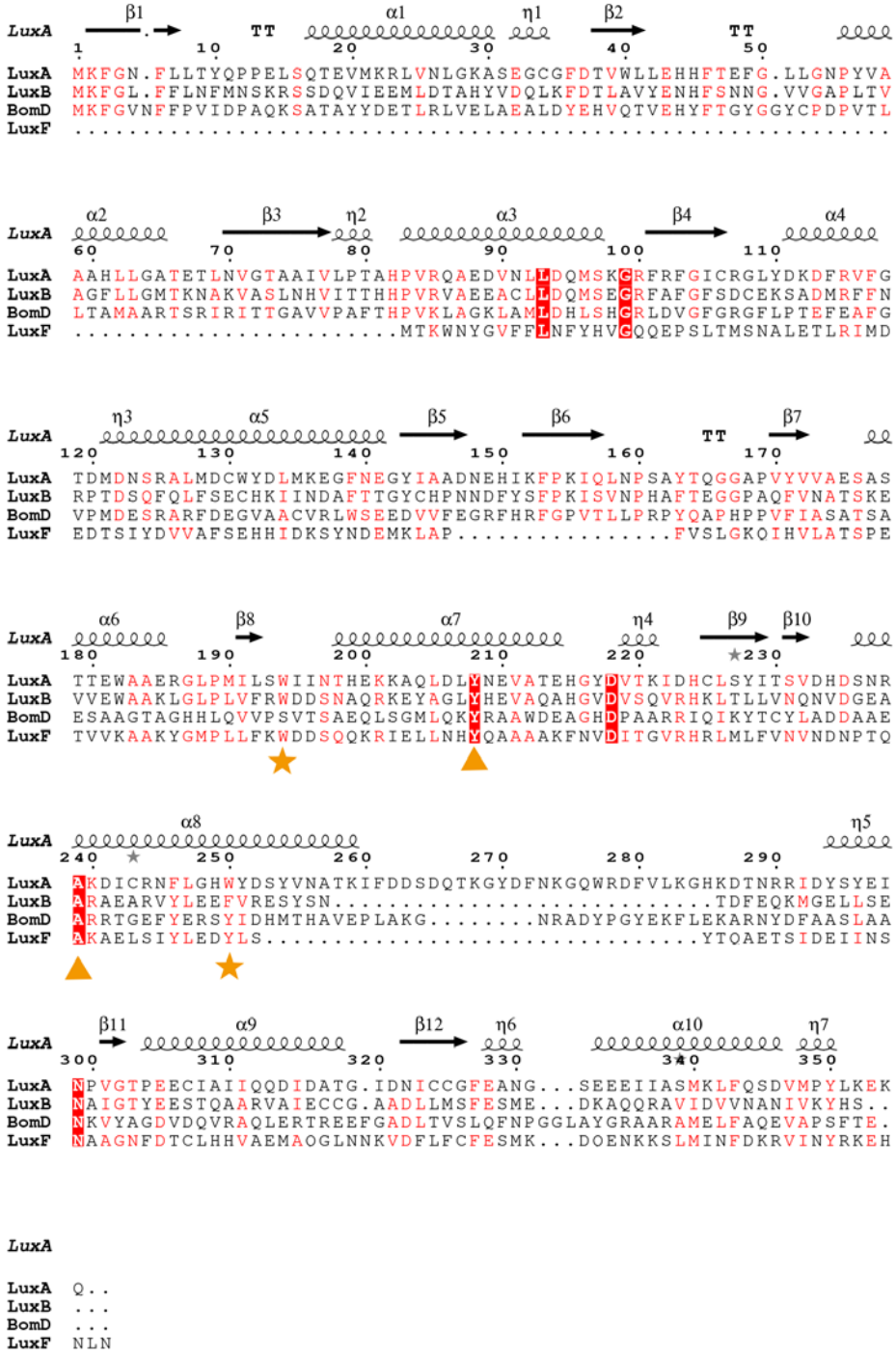


Figure S6. Sequence alignment of flavin monooxygenase BomD with its homologues LuxA[10], LuxB[10], LuxF[11]. Secondary structure assignment based on the structure of LuxA is shown above the alignment as helices for α-helices and arrows for β-strands. Strictly conserved residues among the proteins are indicated in white on a red background. Relatively conserved residues among the proteins are indicated in red. Stars represent FMN binding sites and the triangles represent active sites as electron donor or receptor. The figure was generated from CLUSTALW.



Reference

1. C. J. Carrell, R. C. Bruckner, D. Venci, G. Zhao, M. S. Jorns and F. S. Mathews, *Structure*, 2007, **15**, 928-941.
2. Y. Xie, B. Wang, J. Liu, J. Zhou, J. Ma, H. Huang and J. Ju, *Chembiochem : a European journal of chemical biology*, 2012, **13**, 2745-2757.
3. N. Pulsawat, S. Kitani and T. Nihira, *Gene*, 2007, **393**, 31-42.
4. Y. Mast, T. Weber, M. Golz, R. Ort-Winklbauer, A. Gondran, W. Wohlleben and E. Schinko, *Microbial biotechnology*, 2011, **4**, 192-206.
5. M. Okai, N. Kudo, W. C. Lee, M. Kamo, K. Nagata and M. Tanokura, *Biochemistry*, 2006, **45**, 5103-5110.
6. L. Filisetti, M. Fontecave and V. Niviere, *The Journal of biological chemistry*, 2003, **278**, 296-303.
7. B. N. Webb, J. W. Ballinger, E. Kim, S. M. Belchik, K. S. Lam, B. Youn, M. S. Nissen, L. Xun and C. Kang, *The Journal of biological chemistry*, 2010, **285**, 2014-2027.
8. L. Saa, A. Jaureguibeitia, E. Largo, M. J. Llama and J. L. Serra, *Applied microbiology and biotechnology*, 2010, **86**, 201-211.
9. J. D. Thompson, D. G. Higgins and T. J. Gibson, *Nucleic acids research*, 1994, **22**, 4673-4680.
10. A. J. Fisher, T. B. Thompson, J. B. Thoden, T. O. Baldwin and I. Rayment, *The Journal of biological chemistry*, 1996, **271**, 21956-21968.
11. S. A. Moore, M. N. James, D. J. O'Kane and J. Lee, *The EMBO journal*, 1993, **12**, 1767-1774.