Electronic Supplementary Material (ESI) for RSC Chemical Biology.

Biochemical and structural characterization of *Haemophilus influenzae* nitroreductase in metabolizing nitroimidazoles

Dake Liu,[‡a] Thisuri N. Wanniarachchi,[‡b] Guangde Jiang,[a] Gustavo Seabra,[a] Shugeng Cao,[c] Steven D. Bruner,[*b] Yousong Ding[*a]

[a]Department of Medicinal Chemistry, Center for Natural Products, Drug Discovery and Development (CNPD3), University of Florida, Gainesville, Florida, 32610, USA

[b]Department of Chemistry, University of Florida, Gainesville, Florida, 32611, USA

[c]Department of Pharmaceutical Sciences, University of Hawai'i at Hilo, Hilo, Hawaii, 96720, USA.

Table S1. HiNfsB homologs highlighted in a black dashed circle shown in **Figure 2**.

Sequence ID	Predicted Function	Species	Phylum
WP_135675601.1	NDOR[a]	Actinobacillus indolicus	Proteobacteria
WP_005817727.1	NDOR	Actinobacillus minor	proteobacteria
SUT88287.1	nitroreductase	Actinobacillus rossii	proteobacteria
WP_012073277.1	NDOR	Actinobacillus succinogenes	proteobacteria
EFX91759.1	nitroreductase	Actinobacillus ureae ATCC 25976	proteobacteria
WP_150539066.1	NDOR	Actinobacillus vicugnae	proteobacteria
WP_150539758.1	NDOR	Actinobacillus vicugnae	proteobacteria
WP_205774381.1	nitroreductase	Aggregatibacter actinomycetemcomitans	proteobacteria
WP_111395789.1	NDOR	Aggregatibacter aphrophilus	proteobacteria
WP_109849318.1	NDOR	Aggregatibacter kilianii	proteobacteria
WP_109078874.1	NDOR	Aggregatibacter kilianii	proteobacteria
WP_111295653.1	NDOR	Aggregatibacter segnis	proteobacteria
ERH28656.1	nitroreductase	Aggregatibacter sp. oral taxon 458str. W10330	proteobacteria
WP_034294564.1	NDOR	Alysiella crassa	proteobacteria
WP_103853092.1	NDOR	Avibacterium	proteobacteria
WP_115249331.1	NDOR	Avibacterium avium	proteobacteria
WP_046098299.1	NDOR	Avibacterium paragallinarum	proteobacteria
WP_011201269.1	NDOR	Basfia	proteobacteria
WP_218015861.1	nitroreductase	Bergeriella denitrificans	proteobacteria
WP_015431668.1	NDOR	Bibersteinia trehalosi	proteobacteria
MBP7302631.1	NDOR	Brachymonas sp.	proteobacteria
WP_115072860.1	NDOR	Canicola haemoglobinophilus	proteobacteria
WP_048713719.1	NDOR	Cardiobacterium hominis	proteobacteria
RKW17646.1	NDOR	Cardiobacterium sp.	proteobacteria
WP_100295698.1	NDOR	Caviibacterium pharyngocola	proteobacteria
WP_034613865.1	NDOR	Chelonobacter oris	proteobacteria
WP_100289829.1	NDOR	Conservatibacter flavescens	proteobacteria
WP_131976547.1	NDOR	Cricetibacter osteomyelitidis	proteobacteria
WP_012031187.1	NDOR	Dichelobacter nodosus	proteobacteria
WP_049258825.1	NDOR	Eikenella	proteobacteria
WP_126983001.1	NDOR	Eikenella corrodens	proteobacteria
WP_153714917.1	NDOR	Eikenella corrodens	proteobacteria
WP_067593869.1	NDOR	Eikenella longinqua	proteobacteria
WP_197904047.1	NDOR	Eikenella sp. S3360	proteobacteria
WP_123957534.1	NDOR	Frederiksenia canicola	proteobacteria
WP_010945332.1	NDOR	Haemophilus ducreyi	proteobacteria
WP_078236793.1	NDOR	Haemophilus paracuniculus	proteobacteria
WP_181874609.1	NDOR	Haemophilus parahaemolyticus	proteobacteria
WP_115912646.1	NDOR	Haemophilus parainfluenzae	proteobacteria
WP_197555712.1	NDOR	Haemophilus parainfluenzae	proteobacteria
	HiNfsB	Haemophilus influenzae	proteobacteria
WP_095177020.1	NDOR	Haemophilus pittmaniae	proteobacteria
WP_215767787.1	nitroreductase	Haemophilus sp. SZY H8	proteobacteria
WP_207799970.1	nitroreductase	Haemophilus sputorum	proteobacteria
WP_192574124.1	NDOR	Haemophilus sputorum	proteobacteria
WP_075293801.1	NDOR	Histophilus somni	proteobacteria
EGC17376.1	nitroreductase	Kingella denitrificans ATCC 33394	proteobacteria
WP_042949531.1	NDOR	Kingella kingae	proteobacteria
WP_032136559.1	NDOR	Kingella negevensis	proteobacteria
WP_132300857.1	NDOR	Lonepinella koalarum	proteobacteria
WP_188156245.1	NDOR	Mannheimia bovis	proteobacteria

Continued

Sequence ID	Predicted Function	Species	Phylum
WP_192574124.1	NDOR	Haemophilus sputorum	proteobacteria
WP_075293801.1	NDOR	Histophilus somni	proteobacteria
EGC17376.1	nitroreductase	Kingella denitrificans ATCC 33394	proteobacteria
WP_042949531.1	NDOR	Kingella kingae	proteobacteria
WP_032136559.1	NDOR	Kingella negevensis	proteobacteria
WP_132300857.1	NDOR	Lonepinella koalarum	proteobacteria
WP_188156245.1	NDOR	Mannheimia bovis	proteobacteria
WP_042803014.1	NDOR	Mannheimia granulomatis	proteobacteria
WP_027073341.1	NDOR	Mannheimia granulomatis	proteobacteria
WP_126301985.1	NDOR	Mannheimia haemolytica	proteobacteria
WP_126301237.1	NDOR	Mannheimia haemolytica	proteobacteria
WP_044470289.1	NDOR	Mannheimia massilioguelmaensis	proteobacteria
WP_176812525.1	NDOR	Mannheimia pernigra	proteobacteria
WP_025236058.1	NDOR	Mannheimia sp.USDA-ARS-USMARC-1261	proteobacteria
WP_025235183.1	NDOR	Mannheimia sp.USDA-ARS-USMARC-1261	proteobacteria
WP_138316925.1	NDOR	Mannheimia varigena	proteobacteria
WP_025218067.1	NDOR	Mannheimia varigena	proteobacteria
WP_133545162.1	NDOR	Mesocricetibacter intestinalis	proteobacteria
WP_065256401.1	NDOR	Moraxella	proteobacteria
WP_019518835.1	NDOR	Moraxella boevrei	proteobacteria
WP_078275161.1	NDOR	Moraxella bovis	proteobacteria
WP_029103779.1	NDOR	Moraxella caprae	proteobacteria
WP_152704388.1	NDOR	Moraxella catarrhalis	proteobacteria
WP_079324653.1	NDOR	Moraxella equi	proteobacteria
WP_067006831.1	NDOR	Moraxella nonliquefaciens	proteobacteria
WP_066802655.1	NDOR	Moraxella oblonga	proteobacteria
WP_063514818.1	NDOR	Moraxella ovis	proteobacteria
WP_047977655.1	NDOR	Muribacter muris	proteobacteria
WP_040976402.1	nitroreductase	Necropsobacter massiliensis	proteobacteria
WP_123803883.1	NDOR	Neisseria	proteobacteria
WP_123796250.1	NDOR	Neisseria animalis	proteobacteria
WP_049260955.1	NDOR	Neisseria bacilliformis	proteobacteria
WP_089036383.1	NDOR	Neisseria chenwenguii	proteobacteria
WP_123806104.1	NDOR	Neisseria chenwenguii	proteobacteria
WP_204812176.1	NDOR	Neisseria elongata	proteobacteria
WP_106743254.1	NDOR	Neisseria iguanae	proteobacteria
WP_049322340.1	NDOR	Neisseria meningitidis	proteobacteria
6WT2_A	Nitroreductase	Neisseria meningitidis	proteobacteria
EFM04285.1	Nitroreductase[b]	Neisseria meningitidis ATCC 13091	proteobacteria
WP_016687204.1	NDOR	Neisseria sicca	proteobacteria
WP_203026330.1	NDOR	Neisseria sp. HSUH001	proteobacteria
WP_003778288.1	NDOR	Neisseriaceae	proteobacteria
MBQ9601000.1	NDOR	Neisseriaceae bacterium	proteobacteria
WP_132500780.1	NDOR	Nicoletella semolina	proteobacteria
WP_121123079.1	NDOR	Otariodibacter oris	proteobacteria
WP_111750553.1	NDOR	Glaesserella	proteobacteria
VEG69038.1	Nitroreductase	Pasteurella aerogenes	proteobacteria
WP_211597867.1	Nitroreductase	Pasteurella atlantica	proteobacteria
WP_071523279.1	NDOR	Pasteurella multocida	proteobacteria
WP_126373477.1	NDOR	Pasteurella multocida	proteobacteria
QIM62210.1	NDOR	Pasteurellaceae bacterium Orientalotternb1	proteobacteria
OOH90305.1	NDOR	Pasteurellaceae bacterium 15-036681	proteobacteria

Continued

Sequence ID	Prediction Function	Species	Phylum
TNH03914.1	NDOR	Pasteurellaceae bacterium Phil11	proteobacteria
QIW15646.1	NDOR	Pasteurellaceae bacterium RH1A	proteobacteria
SUB34095.1	Putative nitroreductase	Pasteurella mairii	proteobacteria
WP_159990264.1	NDOR	Pelistega ratti	proteobacteria
WP_115314965.1	NDOR	Phocoenobacter uteri	proteobacteria
VTM25380.1	NDOR	Pseudomonas stutzeri	proteobacteria
WP_077543984.1	NDOR	Rodentibacter genomosp.2	proteobacteria
WP_077582427.1	NDOR	Rodentibacter heylii	proteobacteria
WP_077462997.1	NDOR	Rodentibacter heylii	proteobacteria
WP_077424279.1	NDOR	Rodentibacter myodis	proteobacteria
WP_077584658.1	NDOR	Rodentibacter pneumotropicus	proteobacteria
WP_077499819.1	NDOR	Rodentibacter rarus	proteobacteria
WP_077495209.1	NDOR	Rodentibacter ratti	proteobacteria
WP_194811986.1	NDOR	Rodentibacter sp. DSM 111151	proteobacteria
WP_077421316.1	NDOR	Rodentibacter trehalosifermentans	proteobacteria
MBF1647295.1	NDOR	Rothia dentocariosa	actinobacteria
WP_135008398.1	NDOR	Rothia dentocariosa	actinobacteria
MBS6433272.1	NDOR	Rothia mucilaginosa	actinobacteria
MBF1659027.1	NDOR	Rothia mucilaginosa	actinobacteria
WP_204863289.1	NDOR	Rothia mucilaginosa	actinobacteria
WP_012902713.1	NDOR	Rothia mucilaginosa	actinobacteria
WP_070680901.1	NDOR	Rothia sp.HMSC072E10	actinobacteria
WP_002641763.1	NDOR	Simonsiella muelleri	proteobacteria
WP_157394976.1	NDOR	Ursidibacter arcticus	proteobacteria
WP_124211380.1	NDOR	Vespertiliibacter pulmonis	proteobacteria

[a]: NDOR: NAD(P)H-dependent oxidoreductase; [b]: PDB: 6WT2

Table S2. Primers used in this work.

Primer Name	Sequence (5' to 3')
HinfsbFNde1	ACTCATATGACTCAACTTACTCGTGAA
HinfsbRHindIII	ACTAAGCTTCCCCACCCATTTCACCACTTCA
nfsbR20A-F	GCTCAACAGCGTATTACGACCC
nfsbR20A-R	GGGTCGTAATACGCTGTTGAGC
nfsbR20K-F	GCTCAACAAAATATTACGACCC
nfsbR20K-R	GGGTCGTAATATTTTGTTGAGC
nfsbW71A-F	ACCTTTTAGCGCGGGAATGATAAATCAG
nfsbW71A-R	TTCATTTTTTCGCGTAAGGTTTT
nfsbW71F-F	ACCTTTTAGCTTCGGAATGATAAATCAG
nfsbW71F-R	TTCATTTTTTCGCGTAAGGTTTT
nfsbK119A-F	CCCTCACAGCATACAAAGCCCTG
nfsbK119A-R	CAGGGCTTTGTATGCTGTGAGGG
nfsbY120A-F	CCTCACAAAAGCCAAAGCCCTGCAAGAAG
nfsbY120A-R	GCGGCTTGTTGTTGCTCTGCG
EcnfsbFNde1	ACTCATATGGATATCATTTCTGTCGCCTTAAAGCG
EcnfsbRHindIII	ACTAAGCTTTTACACTTCGGTTAAGGTGATGTTTTGCG

Table S3. The refinement statistics for the structure of HiNfsB in a complex with 1-methyl-5-nitroimidazole (**12**) and nicotinic acid.

	HiNfsB_12	**HiNfsB_nicotinic acid**
Resolution range	45.92 - 2.255 (2.336 - 2.255)	45.79 - 2.301 (2.384 - 2.301)
Space group	P 31	P 31
Unit cell	57.3148 57.3148 121.007 90 90 120	57.084 57.084 121.459 90 90 120
Total reflections	108609 (10407)	182492 (9637)
Unique reflections	20949 (2086)	19464 (1838)
Multiplicity	5.2 (5.0)	9.4 (5.2)
Completeness (%)	99.90 (99.05)	99.13 (92.50)
Mean I/sigma(I)	7.82 (3.60)	19.15 (2.89)
Wilson B-factor	41.71	49.71
R-merge	0.1776 (0.8612)	0.06491 (0.5064)
R-meas	0.1983 (0.9648)	0.06848 (0.5643)
R-pim	0.087 (0.429)	0.02163 (0.2393)
CC1/2	0.976 (0.636)	0.999 (0.807)
CC*	0.994 (0.882)	1 (0.945)
Reflections used in refinement	20946 (2083)	19461 (1838)
Reflections used for R-free	970 (107)	1020 (95)
R-work	0.1784 (0.2637)	0.1846 (0.2543)
R-free	0.2224 (0.3560)	0.2261 (0.2892)
CC(work)	0.964 (0.803)	0.966 (0.827)
CC(free)	0.935 (0.636)	0.960 (0.842)
Number of non-H atoms	3684	3647
macromolecules	3516	3516
ligands	83	90
solvent	85	41
Protein residues	438	438
RMS(bonds)	0.009	0.016
RMS(angles)	0.98	1.26
Ramachandran favored (%)	96.54	96.77
Ramachandran allowed (%)	3.46	3.23
Ramachandran outliers (%)	0.00	0.00
Rotamer outliers (%)	0.00	1.56
Clashscore	5.19	5.76
Average B-factor	46.29	54.77
macromolecules	46.43	54.81
ligands	39.44	56.08
solvent	46.84	48.37

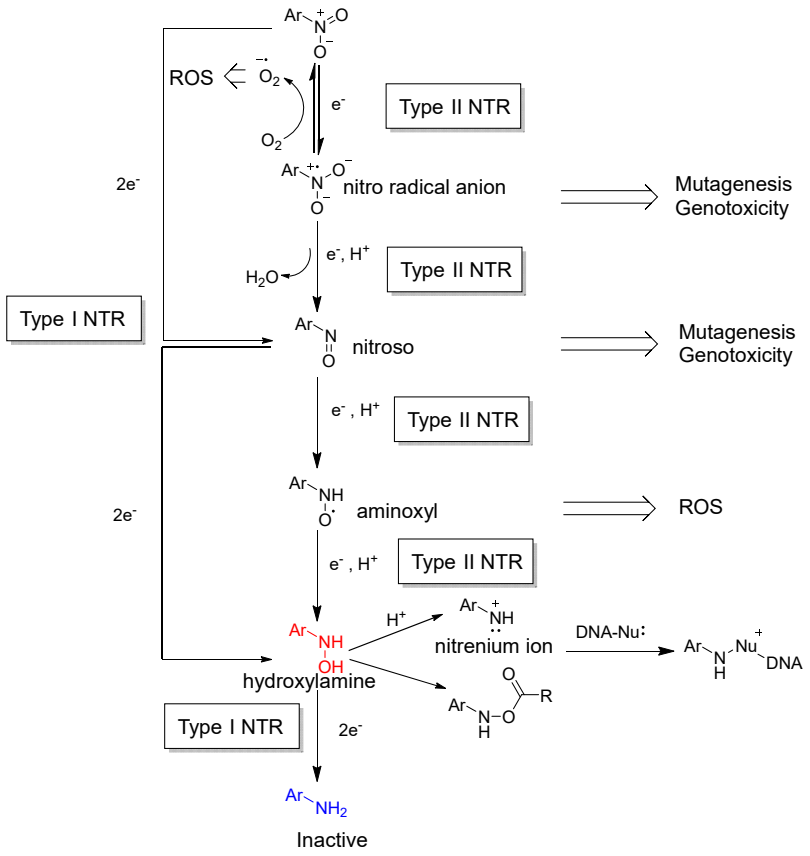
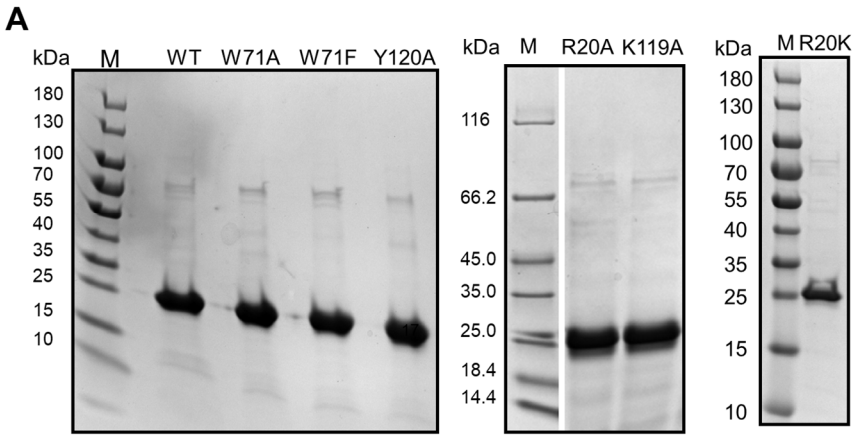


Figure S1. Nitroreduction catalyzed by Type I and Type II nitroreductases (NTRs).



B

	FMN/protein
WT	0.52 ± 0.03
EcNfsB	0.58 ± 0.06
R20A	0.09 ± 0.02
R20K	0.56 ± 0.05
W71A	0.54 ± 0.02
W71F	0.53 ± 0.02
Y120A	0.52 ± 0.01
K119A	0.47 ± 0.02

Figure S2. (**A**) SDS-PAGE analysis of recombinant HiNfsB and its mutants. The calculated molecular weight of recombinant HiNfsB is about 24 kDa. (**B**) Flavin content of HiNfsB, EcNfsB and its mutants. FMN content in each protein sample was quantitated in the HPLC analysis and then normalized again the protein concentration. Each sample was analyzed at least twice. Value is expressed as mean ± standard deviation.

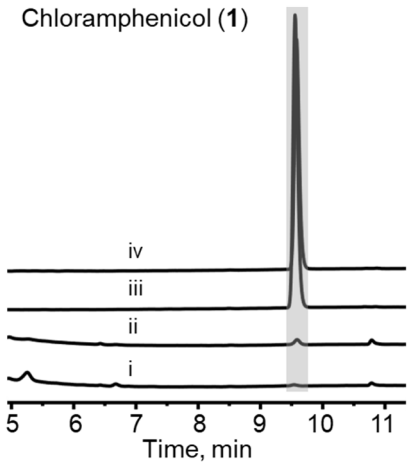


Figure S3. LC analysis of the HiNfsB reaction with chloramphenicol as substrate. The reactions contained 2 mM chloramphenicol, 1 μM enzyme, 1 mM $NADP^+$, 20 mM glucose, and 16 μM GDH in 100 mM phosphate buffer (pH 7.4) and were incubated at room temperature for 10 (ii) or 30 (i) min. Negative controls contained no enzyme (iii: 30 min; iv: 10 min). Chloramphenicol peak is highlighted with a grey bar.

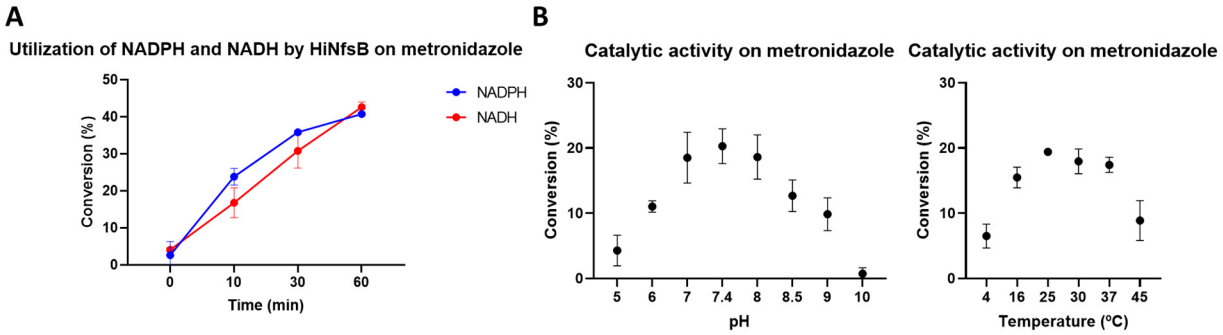


Figure S4. Biochemical characterization of HiNfsB. (**A**) HiNfsB uses both NADPH and NADH in conversing MTZ. The reactions contained 2 mM MTZ, 0.1 μM HiNfsB and 4 mM NADH or NADPH in 100 mM phosphate buffer (pH 7.4) and were incubated at room temperature. Samples were collected at 0 min, 10 min, 30 min and 60 min for the HPLC analysis. (**B**) pH and temperature dependence of HiNfsB in conversing MTZ. The reactions contained 2 mM MTZ, 0.1 μM enzyme, 1 mM $NADP^+$, 20 mM glucose, and 16 μM GDH in 100 mM buffer with pH ranging from 5 to 10 and were incubated at 4 to 45 °C for 10 min. MTZ consumption was quantitated at 320 nm. The data represent means ± s.d. of three independent experiments.

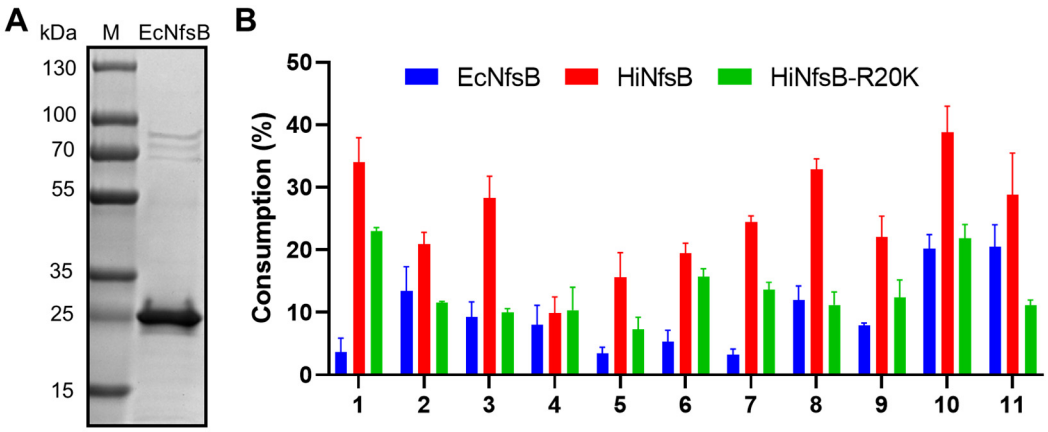


Figure S5. (**A**) SDS-PAGE analysis of recombinant EcNfsB. (**B**) Consumption of clinically used nitroimidazoles by EcNfsB, HiNfsB and its R20K mutant under the same conditions. Recombinant EcNfsB catalyzes limited consumption of selected nitroimidazoles. The enzymatic reaction contained 2 mM substrate, 1 μM enzyme, 1 mM $NADP^+$, 20 mM glucose, and 16 μM GDH in 100 mM phosphate buffer (pH 7.4) and were incubated at room temperature for 10 min. The data represent means ± s.d. of three independent experiments.

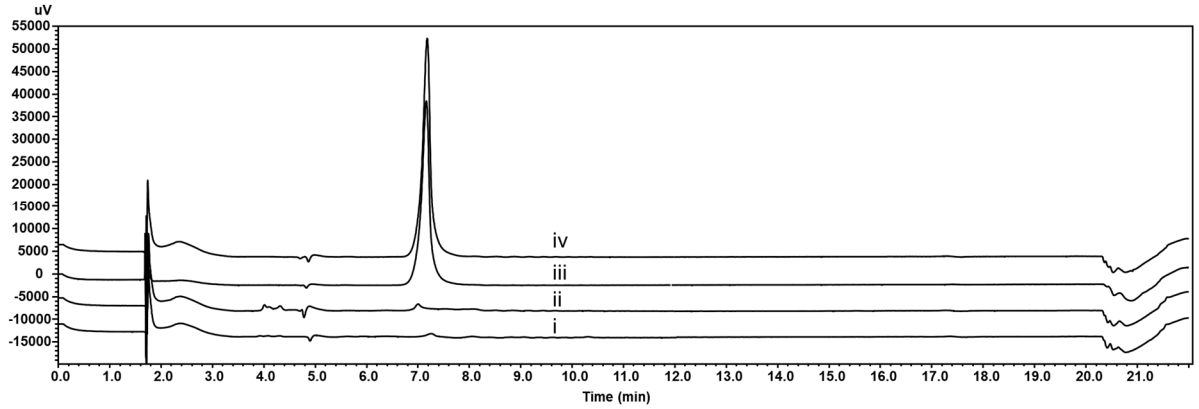


Figure S6. Full LC traces of the HiNfsB reactions with metronidazole (**2**) as substrate. The reactions contained 2 mM substrate, 1 μM enzyme, 1 mM $NADP^+$, 20 mM glucose, and 16 μM GDH in 100 mM phosphate buffer (pH 7.4) and were incubated at room temperature for 3 hours under aerobic (i) or anaerobic (ii) condition. Negative controls lacked $NADP^+$ (iii) or enzyme (iv).

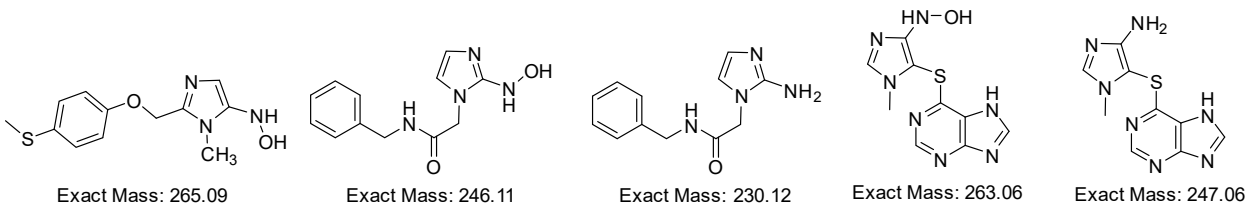


Figure S7. Chemical structures and exact masses of hydroxylamine and amine metabolites of compounds **8**, **10** and **11**.

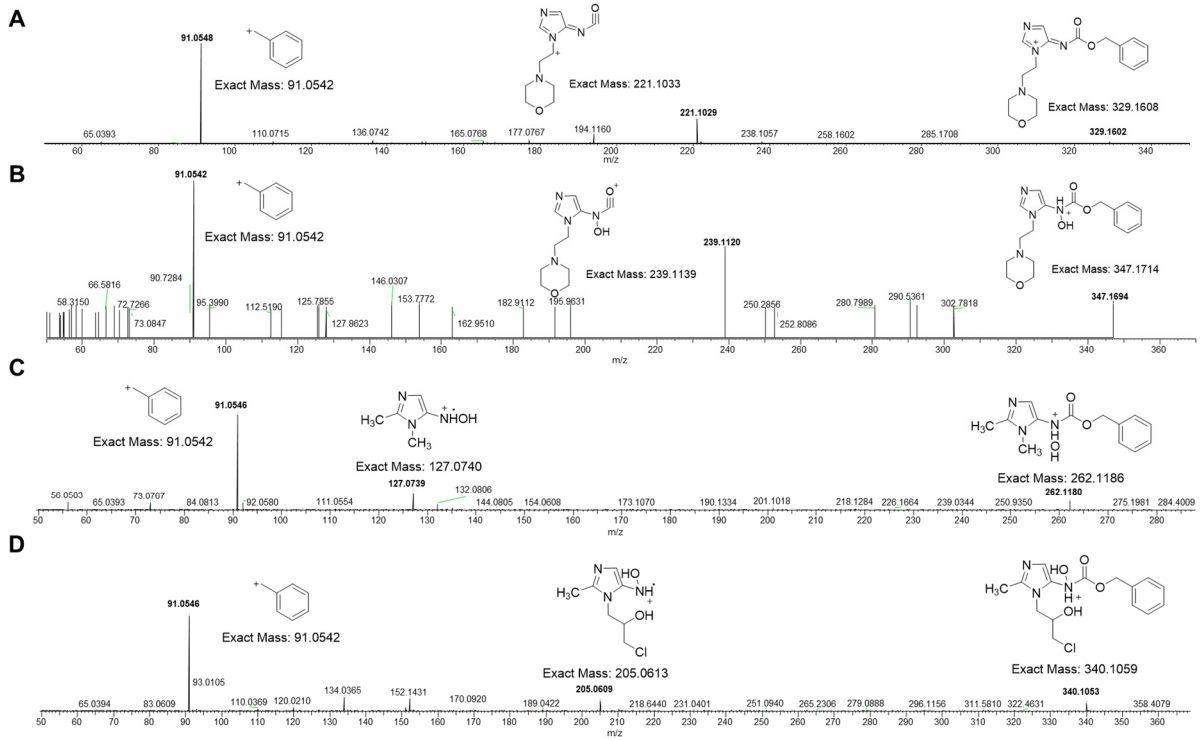


Figure S8. HRMS/MS fragmentation of CBZ derivatives in the HiNfsB reactions with compounds **4** (**A**, **B**), **5** (**C**) and **7** (**D**) as substrates.

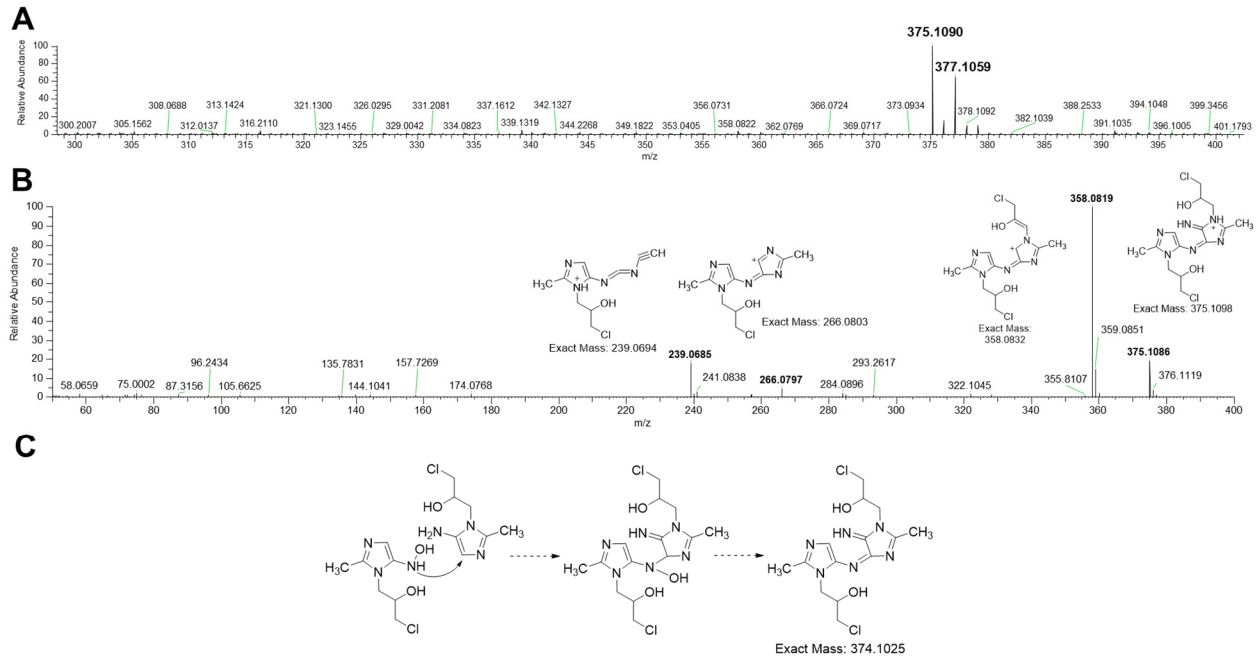


Figure S9. The formation of a potential dimeric metabolite from ornidazole (**5**) by HiNFsB. HRMS (**A**) and MS/MS (**B**) analysis provided structural information of this metabolite. Putative fragment structures were shown. (**C**): A putative way to form the dimeric metabolite from one molecule hydroxylamine and one molecule amine ornidazole species.

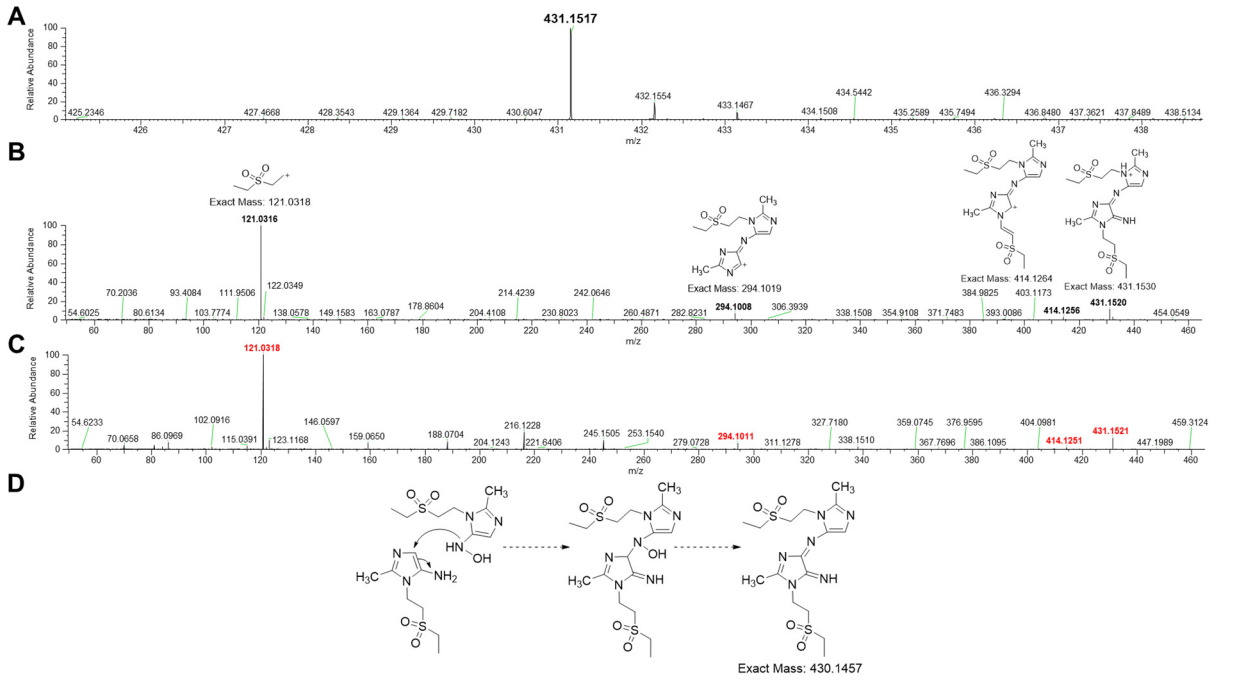


Figure S10. The formation of a potential dimeric metabolite from tinidazole (**6**) by HiNFsB. HRMS (**A**) and MS/MS (**B**) analysis provided structural information of this metabolite. Putative fragment structures were shown. (**C**): HRMS/MS trace of the metabolite produced from tinidazole (**6**) by HiNfsB expression *E. coli* cells. (**D**): A putative way to form the dimeric metabolite from one molecule hydroxylamine and one molecule amine tinidazole species.

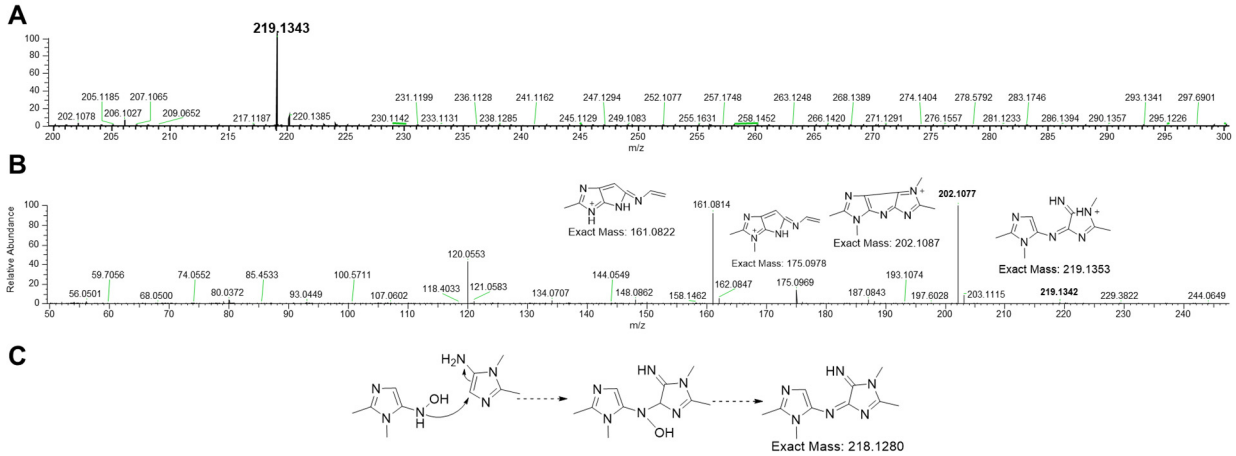


Figure S11. The formation of a potential dimeric metabolite from dimetridazole (**7**) by HiNFsB. HRMS (**A**) and MS/MS (**B**) analysis provided structural information of this metabolite. Putative fragment structures were shown. (**C**): A putative way to form the dimeric metabolite from one molecule hydroxylamine and one molecule amine dimetridazole species.

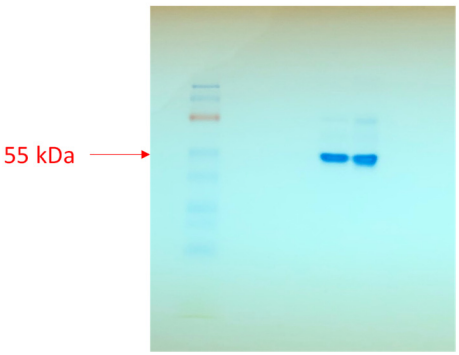


Figure S12. Native gel analysis revealed the homodimer of HiNfsB in solution.

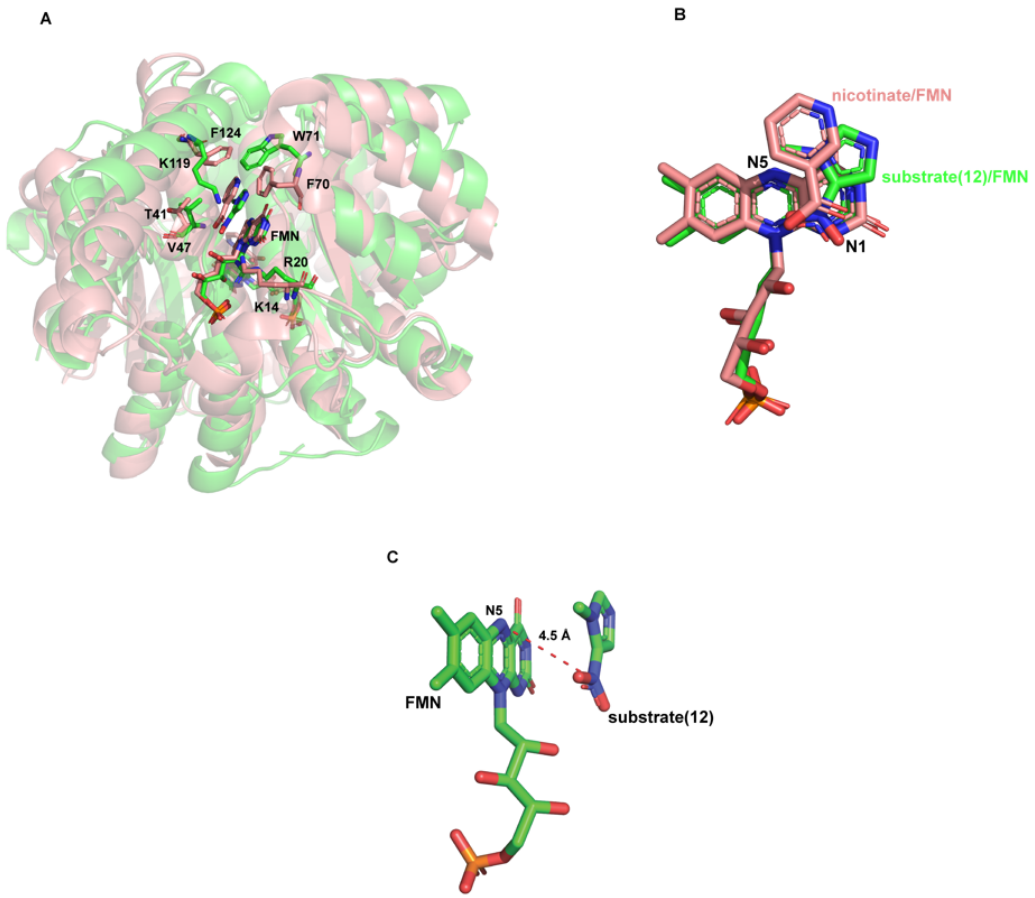


Figure S13. (**A**) Structure alignment of HiNfsB (green) with EcNfsB (PDB 1ICU, light gold). The secondary structure elements of the two enzymes are conserved. Active site residues R20, V47, W71 and K119 of HiNfsB and the corresponding residues of EcNfsB (K14, T41, F70, and F124) are shown as sticks. (**B**) Alignment of bound substrates and cofactors of HiNfsB (green) and EcNfsB. The substrate (**12**) of HiNFsB is rotated relative to nicotinate of EcNfsB to allow similar positioning for the nitro group of substrate (**12**) and the carboxylate of nicotinate. (**C**) The substrate (**12**) is stacked against the isoalloxazine ring of FMN placing the substrate nitro group at a distance of 4.5 Å from the N5 of FMN.

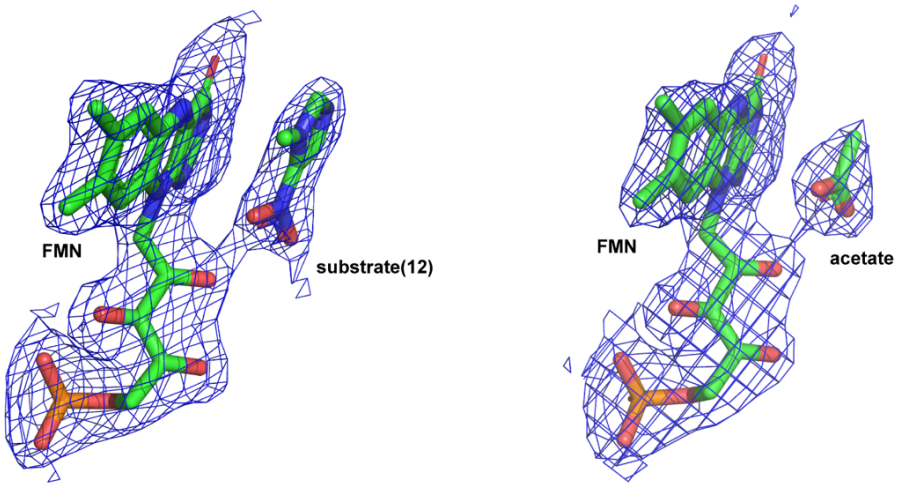


Figure S14. 2Fo-Fc electron density map contoured at 1.0σ around the cofactor FMN, bound substrate in active site 1 (**A**) and acetate in active site 2 (**B**).

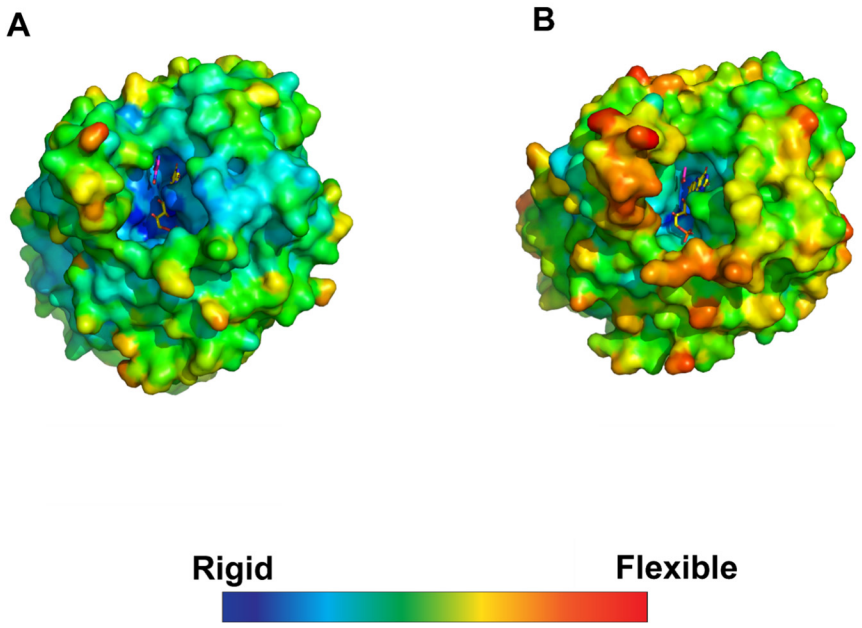


Figure S15. Surface view of the two active sites of HiNfsB. Substrate (**12**) is bound is the active site 1 (**A**) where residues are more ordered than acetate bound active site 2 (**B**).

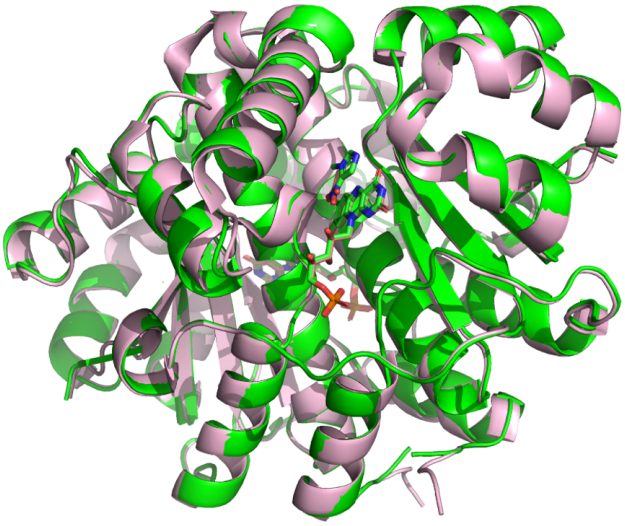


Figure S16. Structure alignment of HiNfsB/nicotinate (green) with recently deposited crystal structure of putative NAD(P)H-flavin oxidoreductase from *Haemophilus influenzae* R2846 (PDB 7LDQ, light pink). The secondary structural elements are conserved in the two structures except for the residue range Lys67-Ile74. The calculated RMSD between HiNfsB and PDB structure 7LDQ is 0.958 Å on all atoms.

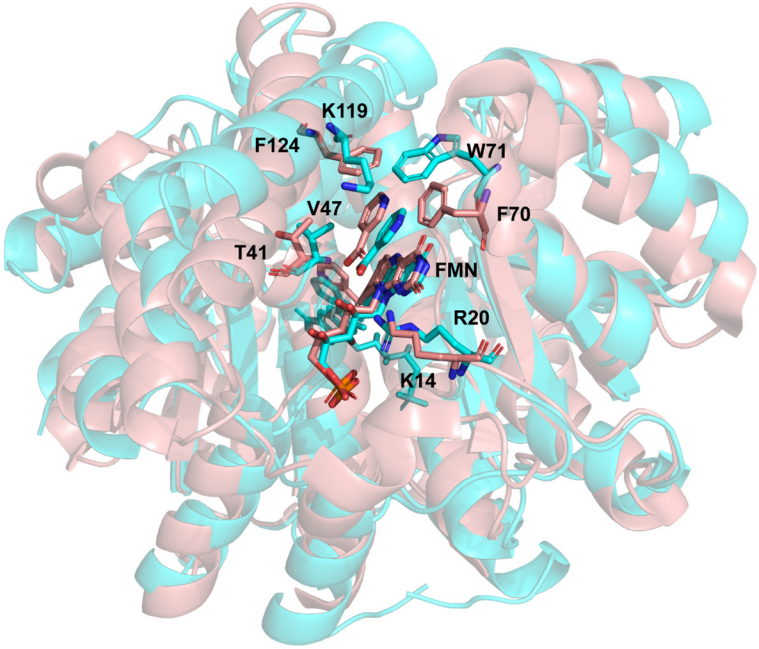


Figure S17. Structure alignment of HiNfsB/nicotinate (blue) with EcNfsB/nicotinate (PDB 1ICU, light gold).

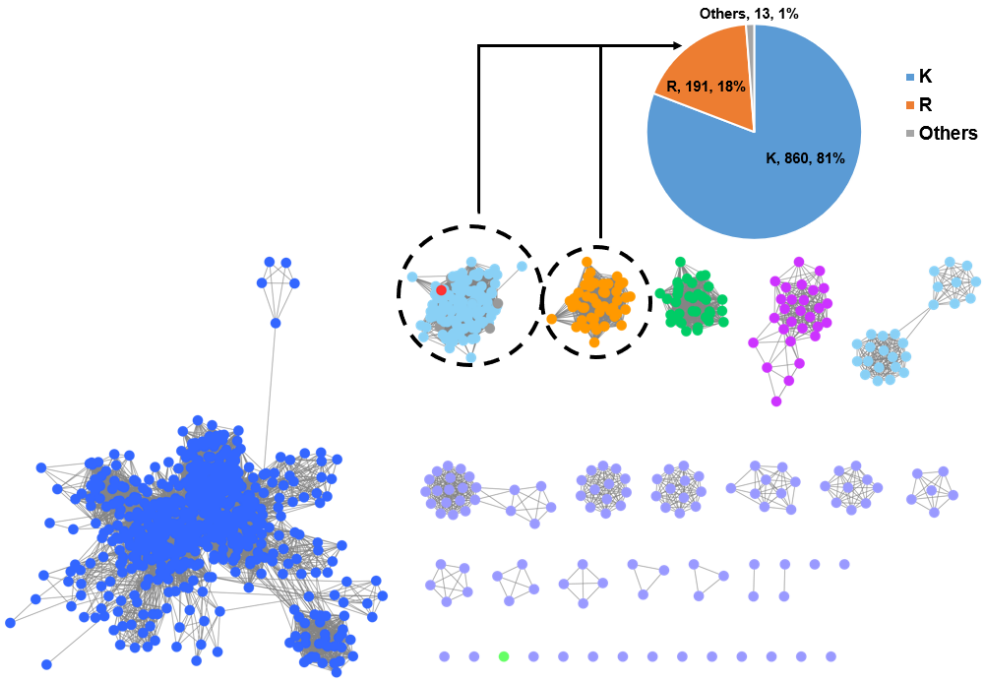


Figure S18. Structure alignment of all 1,046 HiNfsB homologs shown in Figure 2 identified conserved Arg or Lys in the residue 20 (numbered by HiNfsB sequence). All 191 homologs in the HiNfsB cluster and the cluster in orange (homologs from *Streptococcus*, Figure 2) carry Arg in this position.

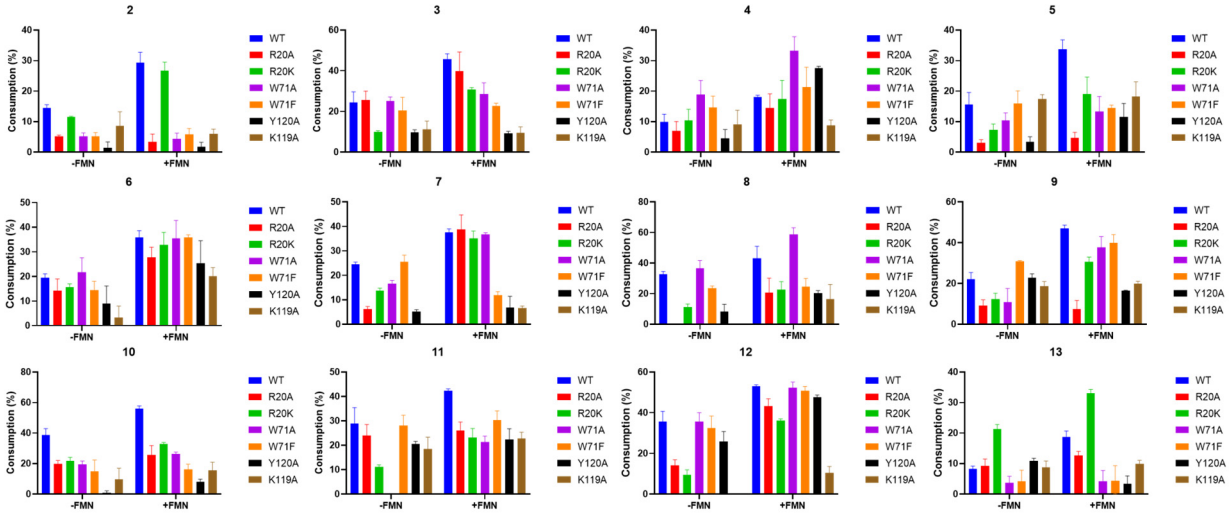


Figure S19. Consumption of compounds **2-13** by WT HiNfsB and its R20A, R20K, W71A, W71A, Y120A and K119A mutants, with or without FMN supplement. The reaction was conducted using 2 mM substrate mixing with (Full) or without (NC) 0.1 μM enzyme, along with NADPH/GDH electron regeneration system. The reaction was incubated for 10 min. Catalytic activities of the enzyme are represented by substrate conversion ratio based on the peak areas of the substrates detected under 320 nm or 276 nm in HPLC. The data represent means ± s.d. of three independent experiments.

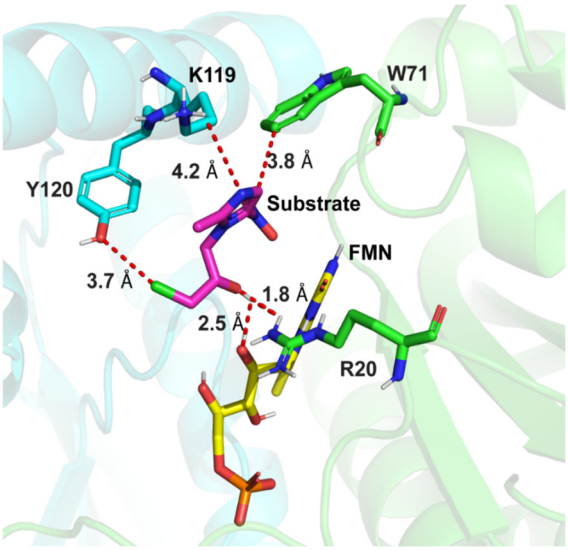


Figure S20. Molecular docking of ornidazole identified Arg20, Trp71, Lys119 and Tyr120 as critical residues for substrate binding.

ATGACTCAACTTACTCGTGAACAAGTTCTTGAACTCTTCCATCAACGCAGCTCAACACGTTA
TTACGACCCAACAAAAAAAATCAGTGATGAAGATTTTGAATGTATTTTAGAGTGCGGTCGAT
TATCGCCGAGTTCTGTAGGCTCTGAGCCTTGGAAATTTTTAGTGATTCAAAATAAAACCTTA
CGCGAAAAAATGAAACCTTTTAGCTGGGGAATGATAAATCAGCTTGATAATTGCAGTCATCT
TGTGGTAATTCTCGCGAAGAAAAATGCCCGTTATGATAGTCCGTTTTTTGTGGATGTGATG
GCACGCAAAGGCTTGAACGCAGAGCAACAACAAGCCGCCCTCACAAAATACAAAGCCCTG
CAAGAAGAAGATATGAAATTACTCGAAAACGACCGCACTTTATTTGATTGGTGCAGCAAACA
AACTTATATCGCCCTTGCAAATATGCTTACTGGAGCTTCAGCCCTTGGCATCGACTCTTGC
CCAATTGAAGGTTTTCATTACGACAAAATGAATGAATGCCTCGCCGAAGAAGGATTATTCGA
TCCTCAAGAATATGCGGTTTCTGTCGCCGCAACCTTTGGCTATCGCTCACGCGATATTGCG
AAAAAATCCCGTAAAGGATTGGATGAAGTGGTGAAATGGGTGGGGTAA

Figure S21. DNA sequence of codon-optimized HiNFsB.