

1 Supplemental Material for "

## 2 Informing Efforts to Develop Nitroreductase for 3 Amine Production

4 Anne-Frances Miller <sup>1,\*,#</sup>, Jonathan T. Park <sup>2,#</sup>, Kyle L. Ferguson <sup>3</sup>, Warintra Pitsawong <sup>1</sup> and  
5 Andreas S. Bommarius <sup>2,3,\*</sup>

6 <sup>1</sup> Department of Chemistry, University of Kentucky, Lexington, KY, USA; afmill3r2@gmail.com

7 <sup>2</sup> School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA, USA

8 <sup>3</sup> School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA, USA

9 \* Correspondence: afmill3r2@gmail.com or ;andreas.bommarius@chbe.gatech.edu Tel.: +1-859-257-9349

10

11 Present addresses:

12 J. T. Park: Janssen Pharmaceutical Companies of Johnson & Johnson, Malvern, PA

13 K. L. Ferguson: Dept. of Chemistry, Univ. of Michigan, Ann Arbor, MI.

14 W. Pitsawong: Dept. Biochemistry, Brandeis University Waltham, MA.

15

16 Academic Editor: W. Van Berkel

17 Received: date; Accepted: date; Published: date

### 18 1. Contents

#### 19 1.1. Tables

20 S1. Compounds used to calibrate  $E^\circ$  calculations, with ChemSpider ID numbers and measured  
21  $E^\circ$  values.

22 S2. Substrates used for kinetic analysis, with ChemSpider ID numbers and Hammett *para*  
23 constants ( $\sigma$ ).

24 S3. Surface area (per monomer) buried in dimer interface.

25 S4. List of the pdb files and amino acid sequences used in the Figures and the sequence  
26 alignment

#### 27 1.2. Figures

28 S1. Sequence similarity network of the different NR subgroups with ours colored and  
29 subgroups containing amine producers identified.

30 S2. Calibration of calculated  $E^\circ$  values vs. experimental values.

31 S3. Dependence of amine formation on molecular volume and calculated  $E^\circ$  or (b) on \*\* and  
32 calculated  $E^\circ$ .

33 S4. Comparison of *StNfsB* vs. *MsPnbA* via dependence of rates on Hammett *para* coefficient,  
34 showing unscaled data and experimental errors.

35 S5. Visualization of emergence of major subgroups from process of sequential identification of  
36 distinguishing elements of structure.

37 S6. Position numbering for flavin and nicotinamide ring of NADH.

38 S7. Comparison of the shapes and sizes of binding sites provided by NRs representing each of  
39 the four groups.

40 S8. Maps onto the structure of amino acid conservation within and between the subgroups.

41 S9. Alignment of the crystallographic sequences showing amine producers.

42

43

44 **2. Tables**

45 2.1. Table S1. Compounds used to calibrate  $E^\circ$  calculations, and resulting calibrated calculated  $E^\circ$   
 46 values, with ChemSpider ID numbers and measured  $E^\circ$  values.<sup>1</sup>

Calibration Compound	ChemSpider <sup>2</sup>	Uncalibrated Computed $E^\circ$ <sup>3</sup>	Measured $E^\circ$ <sup>1</sup> (mV vs. NHE)	Calculated $E^\circ_c$ <sup>4</sup> (Calibrated)
nitrobenzene	7138	-486	-486	-483
4-chloronitrobenzene	21106020	-231	-450	-428
4-nitroacetophenone	21106581	81	-356	-361
trinitrotoluene	8073	561	-253	-259
3-methyl nitrobenzene	21106146	-556	-475	-498
4-methyl nitrobenzene	13863774	-649	-500	-518
3-chloronitrobenzene	21106013	-188	-405	-419
1,4-dinitrobenzene	7211	551	-257	-261
2,4-dinitrotoluene	8150	41	-397	-370
3-nitroacetophenone	21106145	-185	-437	-419
<b>Experimental Compound</b>				
7 3-nitrofurazone	4566720		-270 <sup>5</sup>	-349
<b>Experimental Compound</b>		<b><math>\pi</math> system (# atoms)</b>	<b>Log(P)</b>	
1 nitrobenzene	7138	9	0.32	-483
3 4-nitrobenzene sulfonamide	21360	12	-1.17	-342
4 3-trifluoromethyl nitrobenzene	7108	9	1.24	-392
5 3-nitrophthalimide	11286	14	-0.96	-334
6 1-nitronaphthalene	6588	13	1.32	-427
7 3-nitrofurazone	4566720	14	-2.62	-349
8 4-nitro-1,8-naphthalic anhydride	73216	18	0.72	-170
9 BTZ043	24747357	15	2.01	-259 <sup>6</sup>
10 1,3-dinitrobenzene	7172	12	-1.39	-339
12 2,4,6-trinitrotoluene	8073	15	-2.61	-259
13 3,5-dinitrobenzotrifluoride	109682	12	-0.47	

47 <sup>1</sup> Measured values were obtained from [1, 2]. Plots are provided as Supplemental Figure S2.

48 <sup>2</sup>Information and links on each compound can be retrieved by entering the ChemSpider number as the  
 49 basis of a search at <http://www.chemspider.com/Default.aspx>

50 <sup>3</sup> *Ab-initio* computations were used to produce 1-electron *relative*  $E^\circ$ s using the functional  $\omega$ B97X-D and  
 51 the basis set 6-311+G\*\* as described in methods.

52 <sup>4</sup> Values calibrated to experiment using the linear dependence in Supplemental Figure S2(a): Calibrated =  
53 (Computed + 379)\*4.667.

54 <sup>5</sup> The measured  $E^\circ$  for 3-nitrofurazone at 7.45, which corresponds to  $\sim\text{NO}_2 + 2e^- + 2\text{H}^+ \rightarrow \sim\text{NO}_2 + \text{H}_2\text{O}$  [3].

55 <sup>6</sup> Computations of  $E^\circ$  retained the full aromatic system and attached amine but replaced the  
56 aliphatic extension with two Hs.

57

58 2.2. Table S2. Substrates used for kinetic analysis, with ChemSpider ID numbers and Hammett  
59 *para* constants ( $\sigma$ ). Plots are provided as Supplemental Figure S4.

60

Compound	Chem- Spider <sup>1</sup>	Hammett $\sigma$ para 2	StNfsB		MsPnbA	
			$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_M$ (s <sup>-1</sup> .mM <sup>-1</sup> )	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_m$ (s <sup>-1</sup> .mM <sup>-1</sup> )
nitrobenzene	7138	0	-	1.7 ± 0.1	3.4 ± 0.4	1.2 ± 0.3
4-nitrobenzoate	3620348	0	5.8 ± 0.3	5.1 ± 0.6	-	-
4-chloronitrobenzene	21106020	0.227	-	5.2 ± 0.2	4.5 ± 0.5	11 ± 4
methyl-4-nitrobenzoate	11586	0.45	-	84.5 ± 6.8	9.0 ± 0.3	75 ± 7
4-nitroacetophenone	21106581	0.5	166 ± 25	167 ± 54	-	-
4-nitrobenzenesulfonamide	21360	0.6	182 ± 11	108 ± 20	17.2 ± 0.9	91 ± 13
4-nitrobenzotrile	11593	0.66	290 ± 13	238 ± 22	16.0 ± 1.0	111 ± 22
methyl 4-nitrobenzenesulfonate	21173	0.9	275 ± 22	373 ± 90	19.1 ± 0.8	165 ± 18

61 <sup>1</sup> Information and links on each compound can be retrieved by entering the ChemSpider number as  
62 the basis of a search at <http://www.chemspider.com/Default.aspx>

63 <sup>2</sup> [4]

64

65

## 66 2.3. Table S3. Surface area (per monomer) buried in dimer interface.

Subgroup	Instance (PDB ID)	Buried Area ( $\text{\AA}^2$ ) <sup>1</sup>	Statistics ( $\text{\AA}^2$ )
NfsB	5J8D	4907, 5225	Avg = <b>4600</b> $\pm$ 400
	3HZN	4137	
	1YKI	4681, 4849	
	1VFR	4135	
	3OF4	4321	
	2HAY	4708	
NfsA	1F5V	5751	Avg = <b>5500</b> $\pm$ 300
	2BKJ	4892	
	3N2S	5599, 5593	
	3EOF	5787	
PnbA	2WZW	5014	Avg = <b>5080</b> $\pm$ 70
	3GR3	5146	
Frm2	2IFA	4050, 3957, 3806	Avg = <b>3900</b> $\pm$ 100
HUB	3E39	4168	Avg = <b>4200</b> $\pm$ 100
	3E10	4372	
	4DN2	4172	

67 <sup>1</sup> Average of areas buried of each of two participating monomers, multiple values correspond  
68 to area buried in AB dimer, CD dimer (if present), EF dimer (if present). Calculated in Chimera [5].

69  
70

71 2.4. Table S4. List of the pdb files and amino acid sequences used in the Figures and the sequence  
72 alignment in Figure S9.

73

Accession code <sup>a</sup> (ID in MSA)	Species source and annotated name	Subgroup, Amine production and ref.
Cons_NfsB	NONE consensus sequence for NfsB from SFLD <sup>b</sup>	
5J8D Ent_5J8D	<i>Enterobacter cloacae</i> NR UniProtKB AC: <a href="#">Q01234</a>	NfsB
1YKI Ecoli_1YKI	<i>Escherichia coli</i> NfnB UniProtKB AC: <a href="#">P38489</a>	NfsB
1VFR Vfisch_1VFR	<i>Vibrio fischeri</i> major NAD(P):FMN oxidoreductase UniProtKB AC: <a href="#">P46072</a>	NfsB
3QDL Helicobac_3QDL	<i>Helicobacter pylori</i> RdxA UniProtKB AC: <a href="#">Q25608</a>	NfsB
3OF4 Idiomar_3OF4	<i>Idiomarina loihiensis</i> L2TR Nitroreductase/ dihydropteridine reductase UniProtKB AC: <a href="#">Q5R179</a>	NfsB
3HZN Styph_3HZN	<b><i>Salmonella typhinurium dihydropteridine reductase</i></b> <b>UniProtKB AC: <a href="#">P15888</a></b>	<b>NfsB AMINE (this work)</b>
2HAY Spyo_2HAY	<i>Streptococcus pyogenes</i> M1 GAS Putative NAD(P)H-Flavin Oxidoreductase UniProtKB AC: <a href="#">Q9A120</a>	NfsB
3BEM Bsubt_3BEM	<i>Bacillus subtilis</i> putative nitroreductase YdfN (2632848) UniProtKB AC: <a href="#">P96692</a>	MhqN
4QLX Lplant_4QLX	<i>Lactobacillus plantarum</i> enone reductase UniProtKB AC: <a href="#">U6C5W9</a>	MhqN
3GBH Sepider_3GBH	<i>Staphylococcus epidermidis</i> ATCC 12228 putative NAD(P)H:FMN oxidoreductase (SE1966) UniProtKB AC: <a href="#">A0A0H2VHN8</a>	MhqN
3GE6 Esibir_3GE6	<i>Exiguobacterium sibiricum</i> 255-15 putative nitroreductase EXIG_2970 UniProtKB AC: <a href="#">B1YG32</a>	MhqN
AAP09971 AMINE_Bcer3024	<i>Bacillus cereus</i> ATCC 14579 NAD(P)H nitroreductase called BC_3024 and YdgI_Bc	MhqN, AMINE inferred [6]
3GAG Smutants_3GAG	<i>Streptococcus mutans</i> nitroreductase-like protein (smu.346) UniProtKB AC: <a href="#">Q8D VW4</a>	MhqN
2B67 Spneum_2B67	<i>Streptococcus pneumoniae</i> TIGR4 nitroreductase family protein UniProtKB AC: <a href="#">A0A0H2UP38</a>	MhqN
Cons_NfsA	NONE	

	consensus sequence for NfsA from SFLD	
1F5V Ecoli_1F5V	<i>Escherichia coli</i> NfsA UniProtKB AC: <a href="#">P17117</a>	NfsA
2BKJ Vharv_2BKJ	<i>Vibrio harveyi</i> Flavin reductase UniProtKB AC: <a href="#">Q56691</a>	NfsA
3N2S Bsubt_3N2S	<i>Bacillus subtilis</i> NfrA1 nitroreductase UniProtKB AC: <a href="#">P39605</a>	NfsA AMINE [7]
3EOF Bfrag_3EOF	<i>Bacteroides fragilis</i> Oxidoreductase (YP_213212.1) UniProtKB AC: <a href="#">Q5L9C9</a>	NfsA
KXV33494 AMINE_Gox0834	<i>Gluconobacter oxydans</i> Nitroreductase Gox0834	NfsA AMINE [8]
AAP08598 AMINE_Bcer1619	<i>Bacillus cereus</i> ATCC 14579 Oxygen-insensitive NADPH nitroreductase called BC_1619 or YfkO_Bc	NfsA, AMINE inferred [6]
010964037 AMINE_CaceNit A	<i>Clostridium acetobutylicum</i> 824 NitA WP_010964037 NADPH-dependent oxidoreductase	NfsA AMINE [9]
Cons_PnbA	NONE consensus sequence for PnbA from SFLD	
2WZW AMINE_Msmg_2 WZW	<i>Mycobacterium smegmatis</i> Nitroreductase NfnB UniProtKB AC: <a href="#">A0R6D0</a>	PnbA, AMINE [10]
3GR3 Bhen_3GR3	<i>Bartonella henselae</i> str. houston-1 PnbA UniProtKB AC: <a href="#">A0A0H3M323</a>	PnbA
WP_031019703.1 Strep	<i>Streptomyces</i> sp. NRRL WC-3795 Nitroreductase	PnbA
WP_055507769.1 Saur	<i>Streptomyces aurantiacus</i> Oxidoreductase	PnbA
WP_080719519.1 Coryn	<i>Corynebacterium jeikeium</i> Nitroreductase	PnbA
WP_014983339.1 Nocar	<i>Nocardia brasiliensis</i> Oxidoreductase	PnbA
WP_043519713.1 Kpneu	<i>Klebsiella pneumoniae</i> Nitroreductase	PnbA
WP_057084594.1 Dicke	<i>Dickeya solani</i> Nitroreductase	PnbA
WP_076001261.1 Halio	<i>Halioglobus pacificus</i> Nitroreductase	PnbA

KPQ07647.1 Rhodo	<i>Rhodobacteraceae bacterium</i> HLUCCA12 Nitroreductase	PnbA
ONF96818.1 Sphin	<i>Sphingomonas</i> sp. G39 Coenzyme F420:L-glutamate ligase	PnbA
WP_014992242.1 Actino	<i>Actinobacillus suis</i> Nitroreductase	PnbA
Cons_Frm2	NONE consensus sequence for Frm2 from SFLD	
2IFA Smut_2IFA	<i>Streptococcus mutans</i> Putative nitroreductase UniProtKB AC: <a href="#">Q8DW21</a>	Frm2
4URP Scere_4URP	<i>Saccharomyces cerevisiae</i> Nitroreductase UniProtKB AC: <a href="#">P37261</a>	Frm2
2WQF Llact_2WQF	<i>Lactococcus lactis</i> Nitroreductase CinD UniProtKB AC: <a href="#">Q9CED0</a>	Frm2
1YWQ Bcer_1YWQ	<i>Bacillus cereus</i> ATCC 14579 Nitroreductase family member UniProtKB AC: <a href="#">Q81EW9</a>	Frm2
Cons_HUB	NONE consensus sequence for HUB from SFLD	
3E39 Desulf_3E39	<i>Desulfovibrio desulfuricans</i> putative nitroreductase dde_0787 UniProtKB AC: <a href="#">Q314Q8</a>	HUB
3E10 Cace_3E10	<i>Clostridium acetobutylicum</i> putative NADH oxidase NP_348178.1 UniProtKB AC: <a href="#">Q97IT9</a>	HUB
WP_010966820 AMINE_CaceNit B	<i>Clostridium acetobutylicum</i> NitB nitroreductase	HUB AMINE [9]
3KWK Bthet_3KWK	<i>Bacteroides thetaiotaomicron</i> putative nitroreductase NP_809094.1 UniProtKB AC: <a href="#">Q8ABC9</a>	HUB
3GE5 Pging_3GE5	<i>Porphyromonas gingivalis</i> w83 Putative NAD(P)H:FMN oxidoreductase UniProtKB AC: <a href="#">Q7MX99</a>	HUB
3G14 Cnov_3G14	<i>Clostridium novyi</i> NT Nitroreductase family protein YP_877874.1 UniProtKB AC: <a href="#">A0PZS2</a>	HUB
4DN2	<i>Geobacter metallireducens</i> GS-15 putative	HUB



Gmetal_4DN2	nitroreductase UniProtKB AC: <a href="#">Q39RS1</a>	
3GFA Cdiff_3GFA	<i>Clostridium difficile</i> 630 putative nitroreductase YP_001089721.1 UniProtKB AC: <a href="#">Q17ZU8</a>	HUB
Tther_1NOX	<i>Thermus thermophilus</i> NADH oxidase UniProtKB AC: <a href="#">Q60049</a>	HUB
3EO8 Cdiff_3EO8	<i>Clostridium difficile</i> 630 BluB-like flavoprotein YP_001089088 UniProtKB AC: <a href="#">Q182R2</a>	HUB
Mtub_4XOM Mtubr_4XOM	<i>Mycobacterium tuberculosis</i> C-terminal domain of CoenzymeF420:L-glutamate ligase UniProtKB AC: <a href="#">P9WP79</a>	FbiB
4EO3 Tmar_4EO3	<i>Thermotoga maritima</i> Peroxireductase nitroreductase fusion enzyme. UniProtKB AC: <a href="#">Q9WYL7</a>	TdsD
2FRE Agro_2FRE	<i>Agrobacterium tumefaciens</i> (fabrum) Oxidoreductase UniProtKB AC: <a href="#">A9CKT4</a>	TdsD
AFW02496 Gox	<i>Gluconobacter oxydans</i> H24 Nitroreductase	TdsD
3GB5 Mmus_3GB5	<i>Mus musculus</i> iodotyrosine deiodinase IYD UniProtKB AC: <a href="#">Q9DCX8</a>	Iyd
4TTC Hsap_4TTC	<i>Homo sapiens</i> iodotyrosine deiodinase IYD UniProtKB AC: <a href="#">Q6PHW0</a>	Iyd
3EK3 Bfrag_3EK3	<i>Bacteroides fragilis</i> NCTC 9343 Nitroreductase YP_211706.1 UniProtKB AC: <a href="#">Q5LDN3</a>	Unknown 2
3PXV Dhafn_3PXV	<i>Desulfitobacterium hafniense</i> DCB-2 Nitroreductase UniProtKB AC: <a href="#">B8FRE0</a>	Unknown 2
2ISJ Smel_2ISJ	<i>Sinorhizobium meliloti</i> BluB Flavin conversion to 5,6-dimethylbenzimidazole. UniProtKB AC: <a href="#">Q92PC8</a>	BluB
3K6H Agro_3K6H	<i>Agrobacterium tumefaciens</i> Str. C58 (fabrum) Nitroreductase UniProtKB AC: <a href="#">A9CIP6</a>	Unknown 1

74 <sup>1</sup> PDB IDs are provided if applicable. For sequences not associated with a structure the NCBI  
75 accession number is provided and in parenthesis the abbreviation used in the Multiple Sequence  
76 Alignment. Accession codes can be associated with further information by using them as the basis  
77 for a search at <https://www.ncbi.nlm.nih.gov/home/proteins/> courtesy of the National Center for  
78 Biotechnology Information, U.S. National Library of Medicine 8600 Rockville Pike, Bethesda MD,  
79 20894 USA

80           <sup>2</sup> Subgroups are those of Akiva/Copp [11]. Enzymes for which we found literature reports of  
81 amine production are so indicated with the reference.

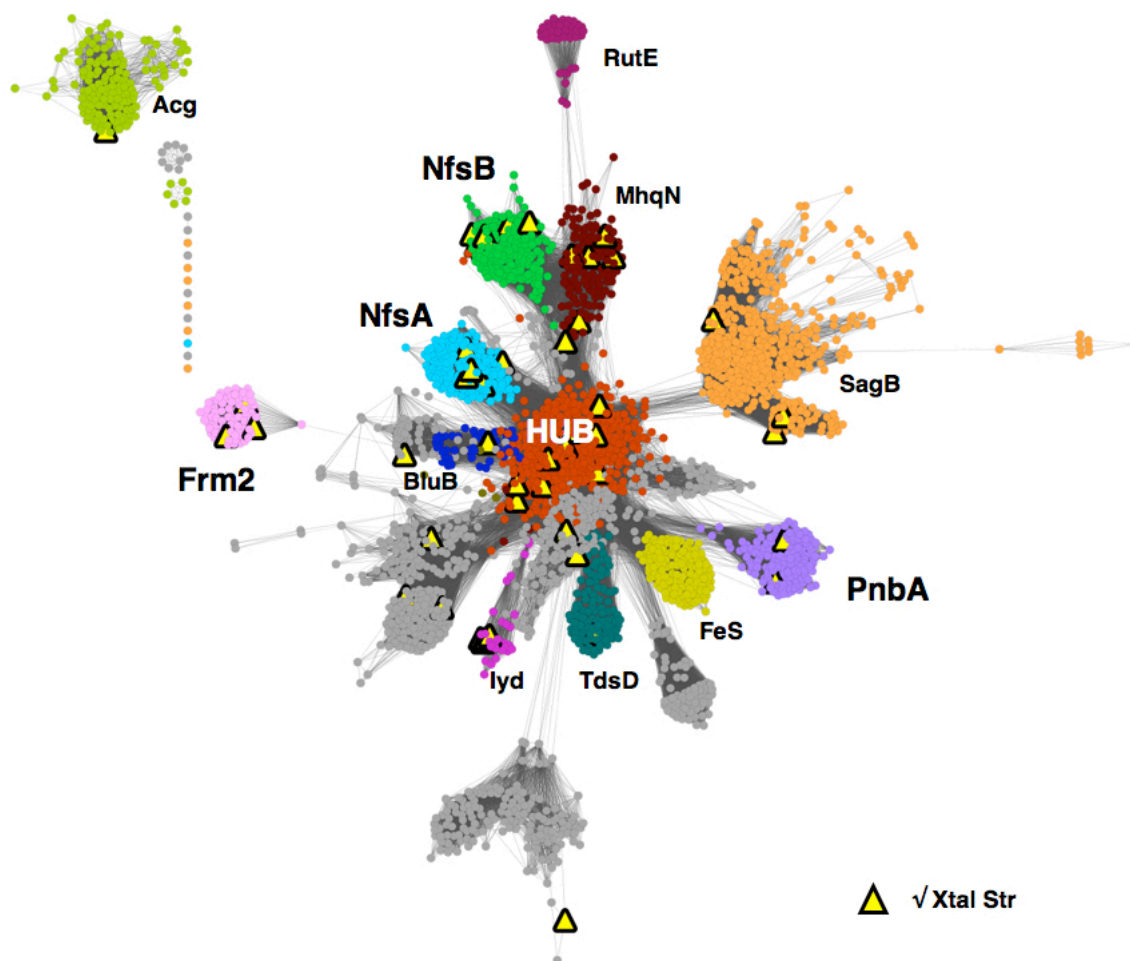
82           <sup>3</sup>SFLD is the Structure-Function Linkage Database [12].

83

84

## 85 3. Figures

86



87

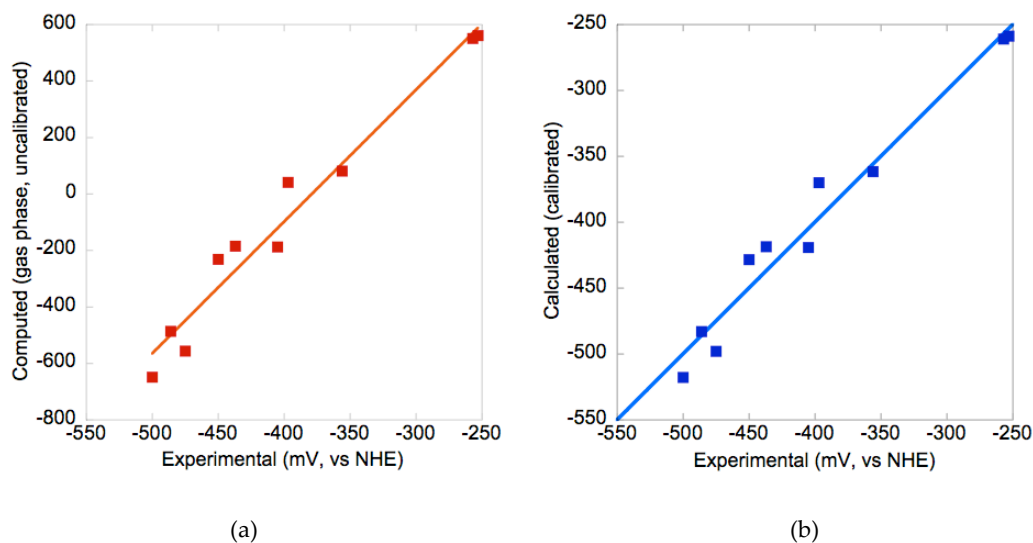
88 3.1. [Figure S1](#). Sequence similarity network of the different subgroups of the nitroreductase  
 89 superfamily coloured similarly to Akiva, Copp et al [11] but with unknown subgroups in grey.  
 90 Small differences in the distribution of the subgroups have resulted from the use of a different  
 91 software but the topology of the network is unchanged (compare with Figure 2 of [11]). Each small  
 92 coloured or grey dot is a node of amino acid sequences sharing 50% or higher identity. Any nodes  
 93 sharing homology scoring higher than  $e^{-16}$  with another node is connected to it with a line. Those  
 94 nodes that include a solved crystal structure are depicted as yellow triangles with heavy borders.  
 95 In each subgroup with ours colored and subgroups containing amine producers identified.  
 96 Otherwise each node is coloured according to the subgroup to which it is assigned. The colours  
 97 replicate those of Akiva/Copp except where needed to permit use of the same colours in structural  
 98 figures. Thus rust replaces the red of Akiva/Copp in order to allow CPK to be used (red = oxygen);  
 99 bright blue replaces light teal for NfsA, the pale mauve (PnbA) and rose (Frm2) of Akiva/Copp have  
 100 also been rendered in more saturated colours, to improve structural figures. For the origins of  
 101 subgroup names, see Akiva, Copp et al. [11].

102

103

104

105



106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

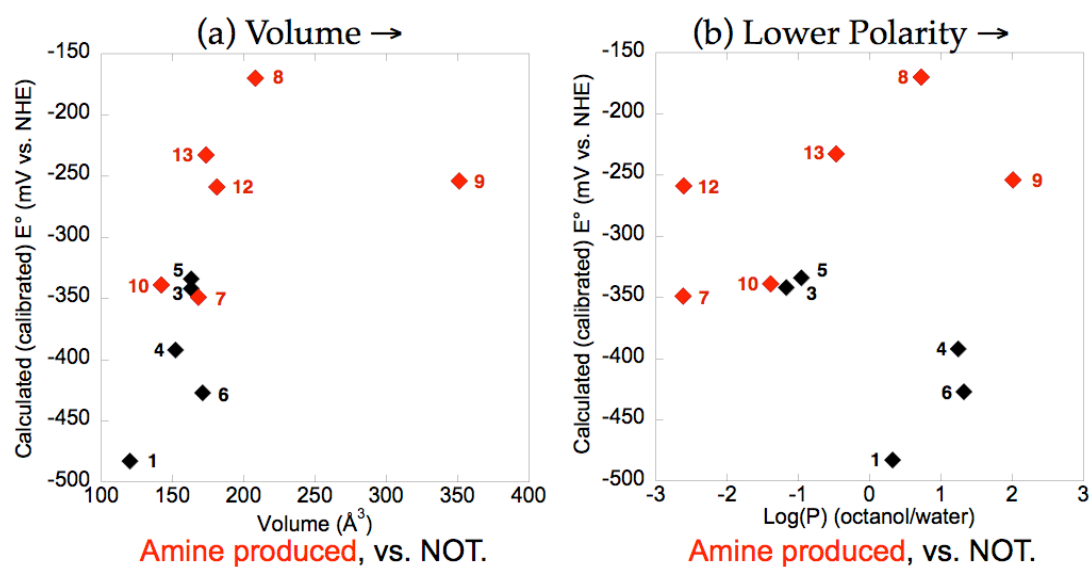
3.2. **Figure S2.** (a) Calibration of calculated  $E^\circ$  values *vs.* experimental values and (b) resulting correspondence between calibrated calculations and experiment yielding MAD = 14 mV. *Ab-initio* computations were used to produce 1-electron *relative*  $E^\circ$ s using the functional  $\omega$ B97X-D and the basis set 6-311+G\*\* as described in methods. These computed values of  $E^\circ$  show a linear trend *vs.* experimental values with  $R^2 = 0.96$  and  $\text{Comp} = (\text{Exp} + 379) \cdot 4.67$  where 'Comp' is the value from *ab initio* gas-phase calculation in mV. The 379 mV systematic offset and the multiplier of 4.67 both likely arise from our use of gas-phase calculations and a modest level of theory whereas the measured values derive from water (for another example see [1, 13]). Thus we emphasize that the *ab-initio* calculations at our level of theory in gas phase do not themselves yield values that replicate experiment in water. They do however produce a tight and linear correlation with all 10 experimental values that were most pertinent to the class of molecules and reaction under study, and therefore permit estimations of anticipated  $E^\circ$ s where they are not available from experiment. Thus we treated the straight-line correlation as a calibration curve, and applied the calibration to obtain calibrated values  $\text{Cal} = (\text{Comp}/4.67) - 379$  to produce calculated 1-e  $E^\circ$ s that are calibrated against experiment. This compromise has allowed us to apply the method to larger molecules than would have been accessible without access to specialized resources. The success of the calculated results can be assessed by comparing the plot *vs.* experimental data (dark blue squares) with an  $Y=X$  line. The mean absolute deviation (MAD) of the calculated values from the line is 14 mV.

126

127

128

Compounds used are tabulated in Supplemental Table S2 with measured values obtained from [1, 2].



129

130

131

132

133

134

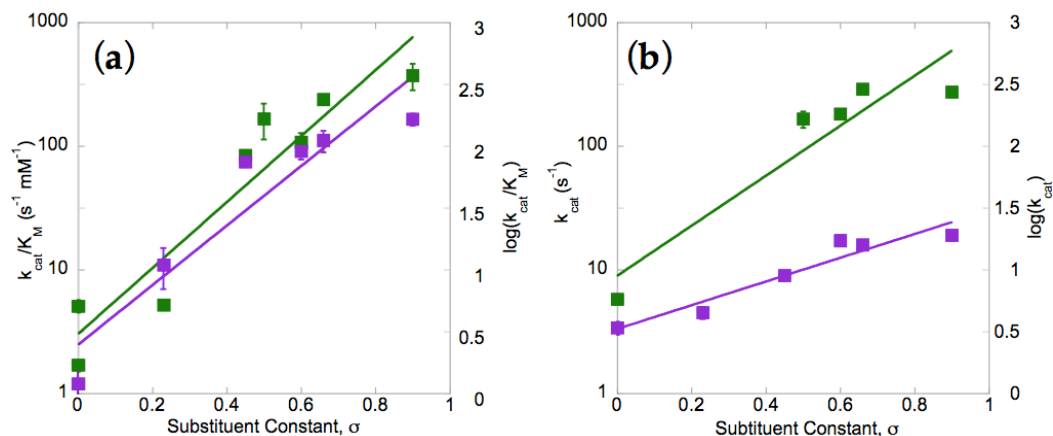
135

136

137

138

3.3. Figure S3. (a) Dependence of amine formation on molecular volume and calculated  $E^\circ$  or (b) on  $\log(P)$  and calculated  $E^\circ$ . Volume and  $\log(P)$  were calculated in Spartan based on CPK atomic volumes and an estimated octanol/water partition coefficient calculated by the method of Ghose and Crippen. While larger molecules (right hand side of panel (a)) seem slightly more prone to conversion to amines, molecular polarity does not appear to correlate with propensity to form amines (b). In both panels the vertical axis is calculated  $E^\circ$ , which correlates with formation of amine product.



139

140

141

142

143

144

145

146

147

148

149

150

151

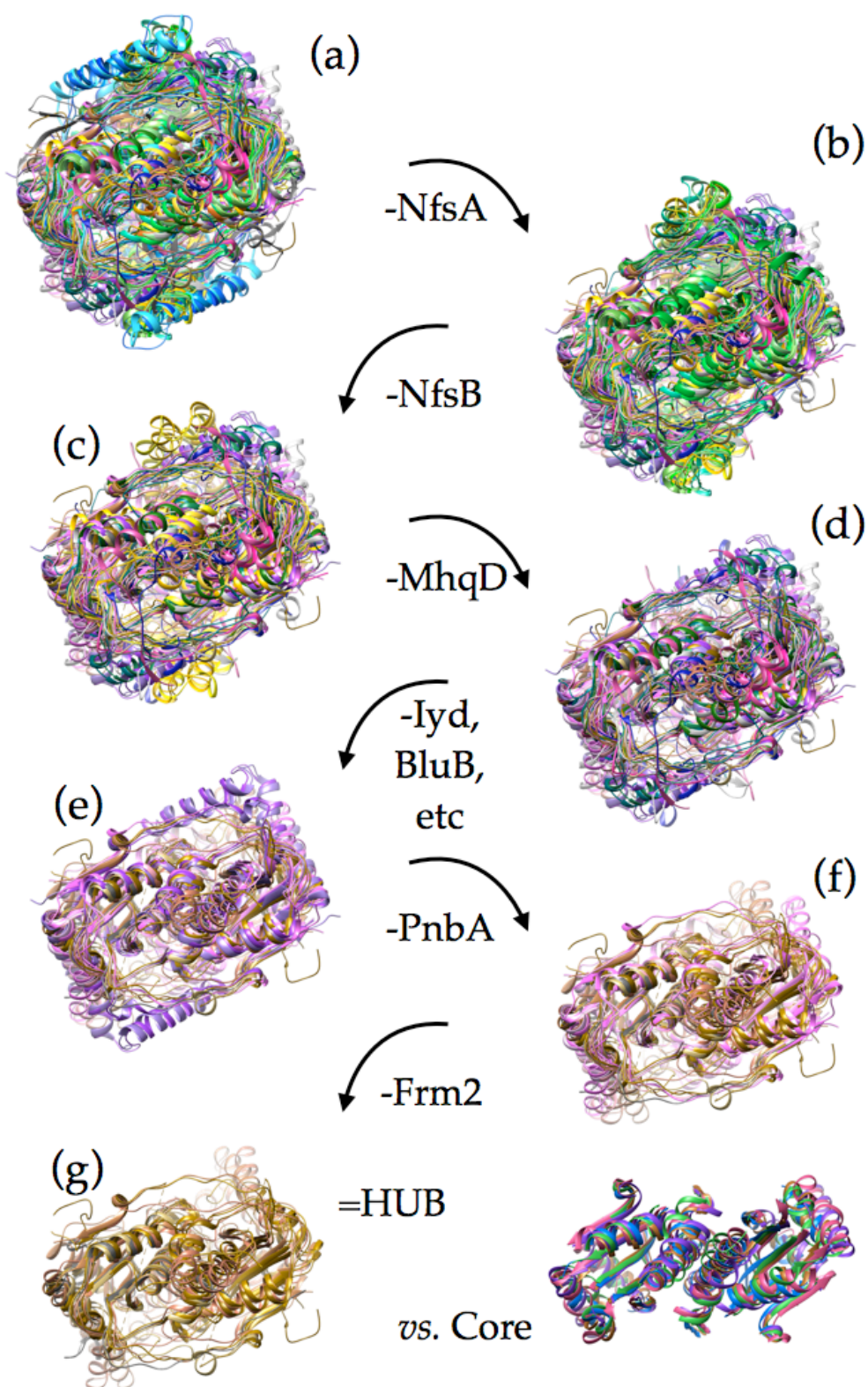
152

153

3.4. Figure S4. Plots of enzymatic reaction rates showing the errors evaluated via triplicate measurements for *StNfsB* (green) and *MsPnbA* (purple). Data are plotted with respect to sensitivity to Hammett para constant  $\sigma$  [4]. (a) shows dependence of the second-order rate constant and (b) shows that of the first-order rate constant. (b) here is the same as panel (b) of Figure 3, but is reproduced here to permit direct comparison with panel (a). Note that error bars are shown but rarely exceed the size of the symbol. Note also that these data are not normalized to the rate constant of the unsubstituted parent compound but presented as measured to permit direct comparison between the rates produced by the two different enzymes.  $\text{Log}(k_{\text{cat}}/K_M)$  data are fit with lines, with for *StNfsB* a slope of  $2.7 \pm 0.4$  and an intercept of  $0.5 \pm 0.2$  ( $R^2=0.89$ ), and for *MsPnbA* a slope of  $2.4 \pm 0.4$  and intercept of  $0.4 \pm 0.2$  ( $R^2=0.88$ ).  $\text{Log}(k_{\text{cat}})$  data are fit with lines, with for *StNfsB* a slope of  $2.0 \pm 0.4$  and an intercept of  $0.9 \pm 0.3$  ( $R^2=0.88$ ), and for *MsPnbA* a slope of  $1.0 \pm 0.1$  and intercept of  $0.5 \pm 0.08$  ( $R^2=0.92$ ). Values are tabulated in Supplementary Table S1.

154

155



156

157

158

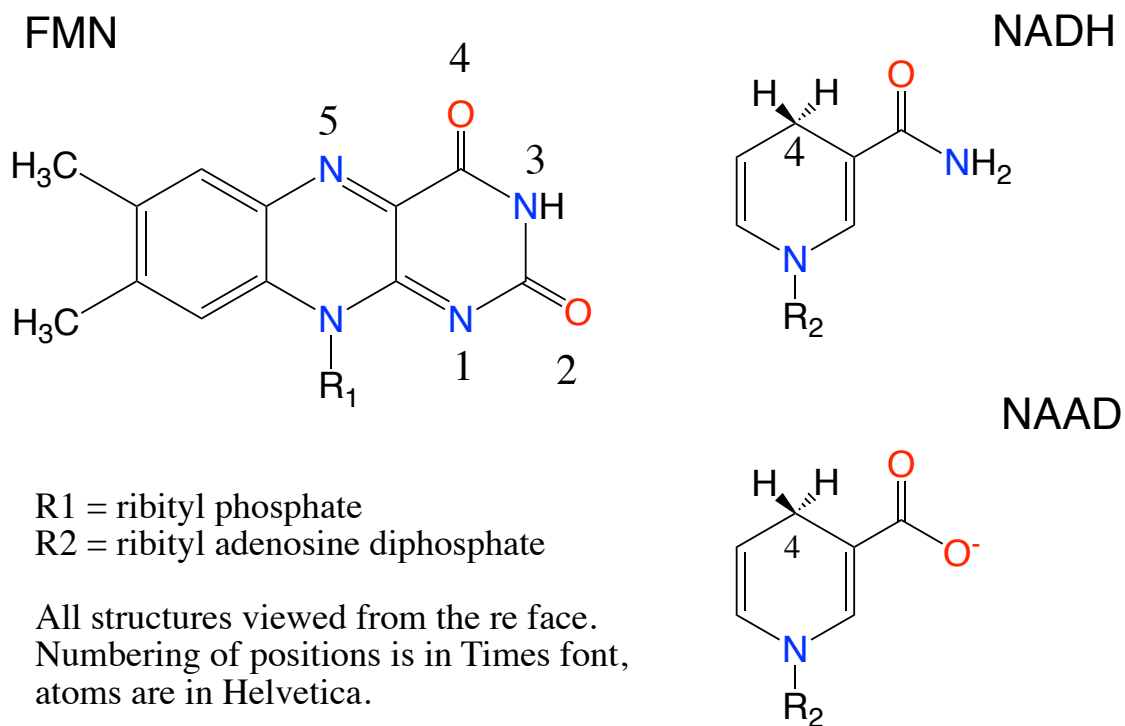
159

160

3.5. [Figure S5](#). View of steps in which successive subgroups can be identified and removed from the remaining large collection of NR-related enzyme structures. Panel A is the overlay of 36 structures shown in the manuscript, in B the NfsA structures have been removed, in C the NfsB structures have additionally been removed, D further lacks the structures of MhqD subgroup

161 members, and so forth. Finally only the structures of HUB subgroup members remain. PDB  
162 accession numbers of all the structures are in Supplementary Table S4.  
163





164

165

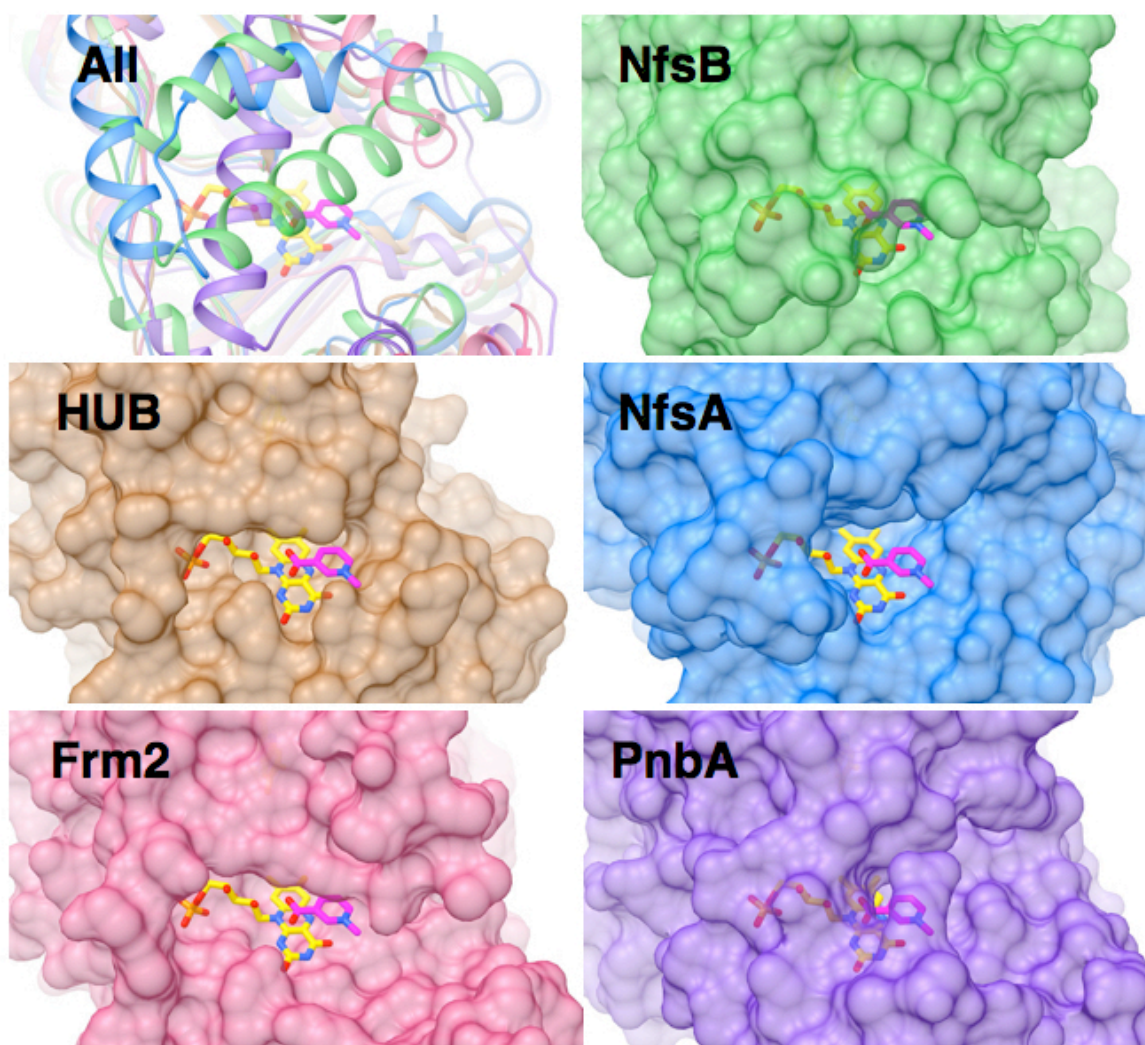
166

167

168

169

3.6. [Figure S6](#). Position numbering for flavin, nicotinamide ring of NADH and analogous nicotinic acid ring of the nicotinic acid adenine dinucleotide used as a model for NADH in crystallography of *EntNfsB*.



170

171

172

173

174

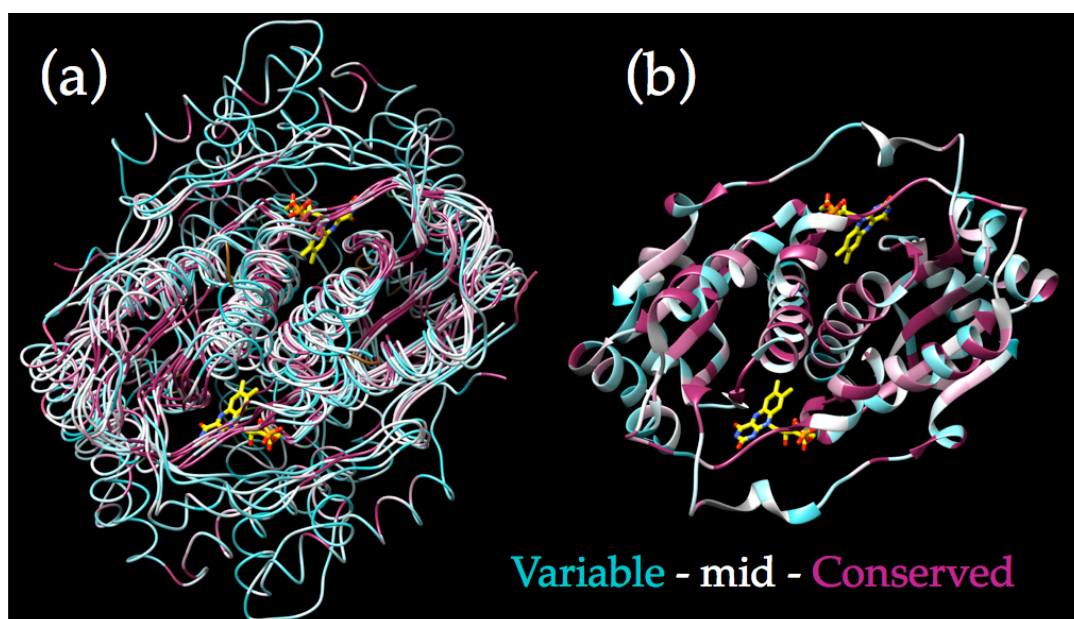
175

176

177

178

3.7. Figure S7. Comparison of the shapes and sizes of binding sites provided by NRs representing each of the four groups. Overlay of the ribbon structures of *MsPnbA* (purple, 2WZV), *StNfsB* (green, 5J8D), *E. coli NfsA* (blue, 1F5V), *Smut Frm2* (pink, 2IFA), *Desulf HUB* (brown, 3E39). In each case one of the fine structures' surface is shown at 50% transparency, and the ribbons of all structures are shown for reference in the top left panel. The FMN and the nicotinic acid ring of bound NAAD of 5J8D are shown in all cases to provide markers for the active site. NAAD is nicotinic acid adenine dinucleotide, and analog of NADH (Supplemental Figure S6).



179

180 3.8. Figure S8. Comparison of (a) distribution of conservation within each of the subgroups - each  
181 depicted as a cord, with (b) conservation between subgroups depicted on core sequence only  
182 (ribbons). Subgroups included are NsfA, NsfB, PnbA, Frm2, HUB. Conservation at each residue  
183 was provided from multiple sequence alignment of each subgroup provided by the Structure  
184 Function Linkage Database [12].

185

186

187

188

189 3.9. Figure S9. Following 5 pages:

190 Alignment of the crystallographic and supporting other amino acid sequences showing amine  
191 producers. The representative sequence from each subgroup is labeled on a coloured background  
192 with the theme that has been used throughout: NfsB (Ent\_5J8D, green), NfsA (Ecoli\_1F5V, blue),  
193 Frm2 (Smut\_2IFA, pink), PnbA (AMINE\_Msmg\_2WZW, purple) and HUB (Desulf\_3E39, rust). In  
194 the alignment, secondary structure is enclosed in yellow boxes (helices) and green (beta strands).

195

196

	1	11	21	31	41	51	61	71
Consensus	-	-	-	-	-	-	-	-
Conservation								
Cons_NsfB	1	-	-	-	-	-	-	-
Ent_5J8D	1	-	-	-	-	-	-	-
Ecoli_1YKI	1	-	-	-	-	-	-	-
Vfisch_1VFR	1	-	-	-	-	-	-	-
Helicobac_3QDL	1	-	-	-	-	-	-	-
Idiomar_3OF4	1	-	-	-	-	-	-	-
AMINE_Styph_3HZN	1	-	-	-	-	-	-	-
Spyo_2HAY	1	-	-	-	-	-	-	-
Bsubt_3BEM	1	-	-	-	-	-	-	-
Lplant_4QLX	1	-	-	-	-	-	-	-
Sepider_3GBH	1	-	-	-	-	-	-	-
Esibir_3GE6	1	-	-	-	-	-	-	-
AMINE_Bcer3024	1	-	-	-	-	-	-	-
Smutants_3GAG	1	-	-	-	-	-	-	-
Spneum_2B67	1	-	-	-	-	-	-	-
Cons_NfsA	1	-	-	-	-	-	-	-
Ecoli_1F5V	1	-	-	-	-	-	-	-
Vharv_2BKJ	1	-	-	-	-	-	-	-
AMINE_Bsubt_3N2S	1	-	-	-	-	-	-	-
Bfrag_3EOF	1	-	-	-	-	-	-	-
AMINE_Gox0834	1	M	S	G	T	S	S	E
AMINE_Bcer1619	1	-	-	-	-	-	-	-
AMINE_CaceNitA	1	-	-	-	-	-	-	-
Cons_PnbA	1	-	-	-	-	-	-	-
AMINE_Msmg_2WZW	0	-	-	-	-	-	-	-
Bhen_3GR3	1	-	-	-	-	-	-	-
Strep	1	-	-	-	-	-	-	-
Saur	1	-	-	-	-	-	-	-
Coryn	1	-	-	-	-	-	-	-
Nocar	1	-	-	-	-	-	-	-
Kpneu	1	-	-	-	-	-	-	-
Dicke	1	-	-	-	-	-	-	-
Halio	1	-	-	-	-	-	-	-
Rhodo	1	-	-	-	-	-	-	-
Sphin	1	-	-	-	-	-	-	-
Actino	1	-	-	-	-	-	-	-
Cons_Frm2	1	-	-	-	-	-	-	-
Smut_2IFA	1	-	-	-	-	-	-	-
Scere_4URP	1	-	-	-	-	-	-	-
Llact_2WQF	1	-	-	-	-	-	-	-
Bcer_1YWQ	1	-	-	-	-	-	-	-
Cons_HUB	1	-	-	-	-	-	-	-
Desulf_3E39	0	-	-	-	-	-	-	-
Cace_3E10	1	-	-	-	-	-	-	-
AMINE_CaceNitB	1	-	-	-	-	-	-	-
Bthet_3KWK	1	-	-	-	-	-	-	-
Pging_3GE5	1	-	-	-	-	-	-	-
Cnov_3G14	1	-	-	-	-	-	-	-
Gmetal_4DN2	1	-	-	-	-	-	-	-
Cdiff_3GFA	1	-	-	-	-	-	-	-
Tther_1NOX	1	-	-	-	-	-	-	-
Cdiff_3EO8	1	-	-	-	-	-	-	-
Mtub_4XOM	1	-	-	-	-	-	-	-
Tmar_4EO3	1	-	-	-	-	-	-	-
Agro_2FRE	1	-	-	-	-	-	-	-
Gox	1	-	-	-	-	-	-	-
Mmus_3GB5	28	W	Q	E	A	E	E	S
Hsap_4TTC	34	W	Q	E	S	E	E	N
Bfrag_3EK3	1	-	-	-	-	-	-	-
Dhafn_3PXV	1	-	-	-	-	-	-	-
Smel_2ISJ	1	-	-	-	-	-	-	-
Agro_3K6H	1	-	-	-	-	-	-	-

	81	91	101	111	121	131	141	151	f - g a -
Consensus	v v s g	- - - k - r l a l a l - a a	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - f - g a -
Conservation									
Cons_NsfB	50	v v e n - p e l k e k - l k - - - p a a w -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - n q p q i t d a S
Ent_5J8D	50	V A S T - E E G K A R V A K S A - A G T Y -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - V F N E R K M L D A S
Ecoli_1YKI	50	V A S T - E E G K A R V A K S A - A G N Y -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - V F N E R K M L D A S
Vfisch_1VFR	52	V I E S - D A A K Q R M - H D S - F A N M -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - H Q F N Q P H I K A C S
Helicobac_3QDL	56	M V T D - K D L K K Q I A - A - - H S Y F -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - N E E M I K S A S
Idiomar_3OF4	50	V I R N - K G L R E Q L V N - - - H S F -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - G Q Q K V A D S S
AMINE_Styph_3HZN	53	V A S T - E E G K A R V A K S A - A G N Y -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - T F N E R K M L D A S
Spyo_2HAY	58	V L D N - K P I K E E I - K P - - F A W G -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - A Q Y Q L E T A S
Bsubt_3BEM	51	T V L D Q D V - K E K L K Q - - A A N G -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Q Y K V V S S S
Lplant_4QLX	62	V V D T P E A - K A K F K Q A V M K F N -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Y P Q V D S A S
Sepider_3GBH	58	V V Q S D E M - K E K V K E - S F G F N -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - S R Q L T T S S
Esibir_3GE6	57	V I D S E E G - K A T L A P - L A K F N -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Q V Q V E T S S
AMINE_Bcer3024	57	V I E S D E A - K A T L A P - L A K F N -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Q S Q V E T S S
Smutants_3GAG	53	V V K N K N K - Q E D L K K - L A A L Q -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - P Q V A T A S
Spneum_2B67	52	V V R E K - - - N A E L A K L A Y G S N -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - F E Q V S S
Cons_NfsA	50	r v t d - - - p e k k a r l a e l a g -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - n Q p y v a q a p - -
Ecoli_1F5V	50	R I T D - - - K A L R E E L V T L T G -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - G Q K H V A Q A A - -
Vharv_2BKJ	50	R V T D - - - S E K R N E L A Q F A G -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - N Q A Y V E S A A - -
AMINE_Bsubt_3N2S	50	G V S D - - - P E K K R E L S V L A G -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - N Q P Y V E K N G - -
Bfrag_3EOF	48	V T R D - - - A E K K E I L S P A H - F -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - N Q P M V K E A P - -
AMINE_Gox0834	76	A V E N - - - P E T R A R L A E L C G -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - N Q K H I E V A P - -
AMINE_Bcer1619	50	Y V T D - - - Q D L K I K L A E L S G -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - N - R H V D C A - -
AMINE_CaceNitA	50	V V R D - - - K K K R E K L A E L V G -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - N Q E Y V A K A P - -
Cons_PnbA	49	v v t G - - e a l d r l r a a l l e a a a a g e p s p d f p y p r - e y - r g p y r e R r r e v g w q l y g a l G l a r g D k e a r a r q a l r n f r f F g A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
AMINE_Msmg_2WZW	64	V V S G - - A A R D R L A E A L V T A H A E E R V T V D F P Y R E G L F Q G - V L Q E R R A D F G S R L Y A A L G I A R D Q T D L L Q G Y N T E S L R F Y G A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Bhen_3GR3	55	V L T G - - K I L Q K V G Q E L S Q L V L S G I K G E R E Y H Y Y P R Q W R E P Y L S R R R K V G L D L Y K S L G I Q K G D Q E K M L H Q K A K N F L F Y G A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Strep	54	I V S G - - A A - - R D A L S K E L L K A Y E D G R D S A D F T G D Y G D G - V Y L E R S Y Q H G A R V Y G A R G I A R S D M E G R N E V I R E N L R F Y D A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Saur	55	V A S G - - A A R D N L A E A L V A A H R A G R R T V D F P F S E D T F T P - V H Q H R R A A M G A A L Y G A L G I D R D D H E L R A A Y D V Q S L S F Y G A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Coryn	49	V I G G - - Q K L - K D L S A A L I E E F D T N G L N P D F T I D Y G Q G V - H P Q - R S M E L A A K M Y G L L G I A R E D K E A R T E F V R E N L R F F G A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Nocar	62	V V S G - - P A R D R L S A A L L T A H T Q K Y L S V D F P Y R D D L Y T D - L H R R R R E D F G E T V Y R A L G I A R D D H P A R E A H S T E S L R F Y G A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Kpneu	53	I V S G - - S K K D V L S K A I L A A D D A G Q L T P D F S F G I N D F P G - V Y R E R Y K E Q G A A Y Y Q A I G V K R D D K D L R R E A S R R N L E F F G A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Dicke	65	I V S G - - K K L T E L A K I L H E K N N A G E L T P D F T F D M D H F Y G - A Y G E R K N A Q G K A Y Y E A M N I A R E D K E G R H R A V A Y N Y S F F N A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Halio	54	V A S G - - D A V E T L R K E M P S R F M Q G E M S L D F P Y D G V Y - E G - V Y K Q R Q H G S A Q A L Y D A A G I A R E D K A R R H E V Q F M E N F T F F G A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Rhodo	50	V V S R - - A V L R A L G H A M T E A A R S G Q P P A P D V P M I S R Y E G - P F R H R Q I D A A K A L Y G A M G I A R N D H D G R A A A F L R N L D A F G A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Sphin	61	V V S G - - E A A E R M R T V L H D A A K S G T K M T P D F P L T G A Y P G - H Y R G R Q I D A A K S L F A A T N V A R D D V P A R T E S F L R N F R F F D A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Actino	50	I V S G - - E A L - K R L S D K L V D G F R T G N L A P D F I F D Q H A L T G V Y E D R M R D S Y K R L Y D A F N I L R E D K A G R K S F A E E N A R F Y G A P	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Cons_Frm2	53	v l f g - - e e h d k l w d - i t k e t l r - k i v p a e a f - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - a a T e e K i d s F k a g -
Smut_2IFA	52	V L F G - - Q D S Q D F W N K I A Y S E L E - K V T P A E A F - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - A G T K A K L E S F A A G -
Scere_4URP	57	I I V G - - D T H K R I W D A V A - S A M P - - - T - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - A E A K K R P E S C R D E A
Llact_2WQF	50	I L T G - - D A Q D K L W D E I V A P E L K - - A A M E A Q G - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - V P E S A W D N T R A K L D G F K - A A
Bcer_1YWQ	55	V L M D - - G E H E K F W D - I V K E T L R A R V - P A E N F - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - E A T V E R L K G F H A G -
Cons_HUB	49	v v t - - - d - k - e l l - e - - k l - - - a - - - - - k p - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - y g - a m f k
Desulf_3E39	53	V I R A D - D - P R C D I L A A - - H T R - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Y G - H I V R
Cace_3E10	47	V L E - D - R - E N I D K L S N - - F S K - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Y A - N S L K
AMINE_CaceNitB	47	a v t n - k n l i s e l s t a - - - - - - - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - r k m g s q - - - - l k
Bthet_3KWK	58	V V S - D - R - A K L D S M A A - A L P - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Y A - K M L T
Pging_3GE5	64	V V R - E - E - E G R K R L A P - - C S - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Q Q P W - F P
Cnov_3G14	50	V V E - S - D - K L K L D I A N - A I E N K - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - A A S E A V L
Gmetal_4DN2	57	C W R F V - V - - V E - D Q A T - K V Q I S E L S Y V E A Y F G P - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Cdiff_3GFA	49	V I T - - - E - K E K E S M L K - A M S K G I Q N E I N D N G L L P G S R - Q - H I A G A N - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Y T V E I M K
Tther_1NOX	56	V V R - D - P - A T K R A L R E - A A F G - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - Q A - H V - E
Cdiff_3EO8	50	V I K - R - R - D L M E K I A D - V I T K K Q Q E I L V E M D K V S V D K A N R F R K F V K - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - N F T L F Y L
Mtub_4XOM	58	W L - Q T P A I - R A R L L D R M K D K W R S D L T S D G L P A D A I E R R - V A R G Q I L Y - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - - D A P
Tmar_4EO3	52	V V D E E E L - - - - - - - L K K I H E - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - G N Y W M
Agro_2FRE	59	Y A H K D S E - - - - - - - D W P L F - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - G N Q K W A
Gox	59	Y A R R N T P D - - - - - - - W D R F - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - N R S W A
Mmus_3GB5	103	V V K D P D M - - - - - - - K H K I R E I I E E E E E I N Y M K R M G K R W V T D L K K L R T N W I K E Y L D - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Hsap_4TTC	109	V V K D P D V - - - - - - - K H K I R K I I E E E E E I N Y M K R M G H R W V T D L K K L R T N W I K E Y L D - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Bfrag_3EK3	53	A V C N T V K L E E L - - - - - N E R I K G A F A K S D D K H L R E R G H S E T Y C C Y Y H A - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - -
Dhafn_3PXV	52	V V K N K E L - - - - - - - M Q E M E A - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - E G L A Y L A G M E D Q S S Y N R I M E R G G R - - - - L F Y
Smel_2ISJ	73	L V R Q D E T - - - - - - - R E K V - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - W Q - - - A F - - - - - Q R A
Agro_3K6H	59	V Y R G E E - - - - - - - R V R L - - - - - - - - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - S E A A L R I A L E K N P D L D L Q Q Q E A E R T - - - - R F T



	241	251	261	271	281	291	301	311
Consensus	I a a q n l l l a a a a l g l g s c p i g g f d - - - a e v v r e - - - g l e - - - k l l f g i a f G y - p d - - - - n g - p k p r - p - l e e v v h f -							
Conservation								
Cons_NsfB	137 I A I G N I I T A A A E I G I D S C P M E G F D - - - p e k v d e i L - - g l d e k g l r s v v m l a I G Y r d - e e - d d v - l p K v R k p - l e e v v t e i							
Ent_5J8D	145 L N V G N F L L G V G A M G L D A V P I E G F D - - - A A I L D E E F - - G L K E K G F T S L V V V P V G H H S V E D F N - A T L P K S R L P - L S T I V T E C							
Ecoli_1YK1	145 L N V G N F L L G V A A L G L D A V P I E G F D - - - A A I L D A E F - - G L K E K G Y T S L V V V P V G H H S V E D F N - A T L P K S R L P - Q N I T L T E V							
Vfisch_1VFR	145 L A L G N A L H T L A R L N I D S T M E G I D - - - P E L L S E I F - - A D E L K G Y E C H V A L A I G Y H H P S E D Y N A S L P K S R K A - F E D V I T I L							
Helicobac_3QDL	142 I A V G Q I C M G V S L M G L D S C I I G G F D - - - P L K V G E V L - - E E R I N K P K I A C L I A L G K - - - R V A E A S Q - K S R K S - K V D A I T W L							
Idiomar_3OF4	137 I G L G T F L L A A A E L E V D S C P M E G I E - - - H D A Y D N I L - - S L K D L G L S T V F A C P V G Y R S E A D T T - Q F Q K K V R Q P - L S R F K V V L							
AMINE_Styph_3HZN	148 L N V G N F L L G V A A M G L D A V P I E G F D - - - A E V L D A E F - - G L K E K G Y T S L V V V P V G H H S V E D F N - A G L P K S R L P - L E T T L T E V							
Spyo_2HAY	150 I A L G N M M M T A A L L G I D T C P I E G F H - - - Y D K V N H I L A K H N V I D L E K E G I A S M L S L - - G Y R L R D P K H A Q V R K P - K E E V M S V V							
Bsubt_3BEM	137 L S A M M F M L S A A A G W D T C P M I G F D A E - - - - A V K R - I - L N I D D Q F E V M M I T I G K E K T E S - - - R R P R G Y R K P V N E F V E Y M -							
Lplant_4QLX	150 V V G M Q L L L V A R A H G Y D A N A F S G I D F E - - - - K M I P - T L G L D P K R Y V P V M G I A I G K - - - A A Q E P L H T T R Y D A K T Q T D F L A - - -							
Sepider_3GBH	145 L M A M Q L M L T A K A H G Y D T N P I G G F D K E - - - - N I A D - I I G Y D S D R Y V P V L A I A I G K - - - K A Q - - - D A H D S V R L P I D D V R E F L -							
Esibir_3GE6	144 L V S M Q L M L V A R A H G Y D T N P I G G Y E K D - - - - Q I A E A F G M E K D R Y V P M L L S I G K - - - A V D - - - A G Y P S V R L P I N D I A D W K -							
AMINE_Bcer3024	144 L V A M Q F M L A A R A H G Y D T C P I G G F E K D - - - - Q I A E A F G L D K E S Y V P V M L I S I G K - - - A A D - - - S G Y Q S V R L P I E K V A E W K -							
Smutants_3GAG	138 L F A M N L M Q V V R V Y G Y D S V P M R G V D F D - - - - - A I K T Y L D M P N G W E P I L M L P V G K - - - - A L Q - A G N P H V R K S V A E F A E I E							
Spneum_2B67	137 L V A M N L V L A L T D Q G I G S N I L G F D K S - - - K V N - E - V L E I E D R - F R P E L L I T V G Y - T D E - K L E - P S Y - - R L P V D E I I E K R -							
Cons_NfsA	113 I a a Q n a v v a A e s I G L G i v y i G g i r n - n p e e v a e l l e L P e y - - - v f P v f G l c v G y P d q - - - e p a - v K P R L P - l e a v l h e e							
Ecoli_1F5V	110 M M A Q N A L V I A A E S L G L G G V Y I G G I R N - N I E A V T K L L K L P Q H - - - V L P L F G L C L G W P A D - - - N P D - L K P R L P - A S I L V H E N							
Vharv_2BKJ	110 I M A Q N C L L A A E S M G L G G V Y I G G L R N S A A Q - V D E L L G L P E N - - - S A V L F G M C L G H P D Q - - - N P E - V K P R L P - A H V V V H E N							
AMINE_Bsubt_3N2S	116 L A A Q N M S I A A E S M G L G I C Y I G G I R N - E L D K V T E V L Q T P D H - - - V L P L F G L A V G H P A N L - - - S G - K K P R L P - K Q A V Y H E N							
Bfrag_3EOF	112 L V A Q T F C T L A E E A G L G I C Y L G T T Y - N P Q M I D A L H L P E L - - - V F P I T V T V G Y P A E - - - - S P K Q V D R L P - I E G I I H E E							
AMINE_Gox0834	141 I A A Q N A V T A F E S L G L G T V Y I G G L R N - H P E E V S L L S L P P R - - - S V A V V G L C V G K P D P A A - - - Q N G - Y K P R L G - Q D V L H R E							
AMINE_Bcer1619	113 L F A Q N L A L A A E S L G Y G I C Y I G G I R N - N P R E V S E L L H L P D K - - - V Y P V F G M T V G V P D E - - - D H G - V K P R L P - V A A V L H E N							
AMINE_CaceNitA	115 I A L G A S V V A A E S L G L G T V P I G G I R K - N P E E V I E I L G L P K Y - - - T F P M V G L V V G Y P A D - - - E S H - K K P R V P - F E S F R H N E							
Cons_PnbA	145 m f l q n f m l a a r a r G l g t c p Q a a l a m - - y s p v v r e h l g i p d d r - - m i v c g i s I G y - a D p d a p a n s f r t e R a p - v e e f v t f v							
AMINE_Msmg_2WZW	161 I Y A Q T L M L A M T A H G I A S C P Q A L L S F - - Y A D T V R A E L G V - E N R - - K L L M G I S F G Y - A D D T A A V N G V R I P R A G - L S E T T R F S							
Bhen_3GR3	153 M F M Q T I M L A A R G F G L D T C A Q A A F A D - - Y H K Q I R S L L S V P S D R - - H I I C G M A L G Y - R D M N A P E N N F E T E R E P - I D N F V H F I							
Strep	150 M Y A Q N F L L S L T A R G F H G I P Q A V L G V - - Y A D T V R A S L G V R D D L - - K L L F G I S F G I - G D K T S P M Y D I G M G R I P - L Q Q S V V L H							
Saur	152 M Y A Q T L L L A L N A Y G V A S C P Q A L M S F - - F A D T V R E T L G L T G G - - - K L L F G I A F G Y - A D P A A A V N T V A V G R E P - L R E T T R F H							
Coryn	145 T Y A E N F L L S L T A H G Y H G V P Q G M I S L - - L A P T V R E F L G V D E E Y - - K L V T A L T F G R - A D E S S P V Y K T A P G R V A - L S E T V T V H							
Nocar	159 M Y G Q T L L L A M T A Y G I A S C P Q G L G F - - Y G E T I R H Q L G I T D G - - - K V L F G V S F G Y - A N T S A A V N R I R V A R A K - L C E T T R F H							
Kpneu	151 M Y A Q T L L L S L T A H G L G G V P Q T S L G F - - Y G E T I R K E L G V D P S L - - K L L F G I S L G Y - P D I T H P A N Q Y R I N K V S - L G E S V T F H							
Dicke	163 M Y G Q T F L L S L A A H G L G G I P Q T S L G F - - F A D P I R H F L G I P D N L - - K L L F G I S F G Y - P D E A A P G N R L R M D R V N - I S D S V T F H							
Halio	150 M Y A Q T L M L A M A A N G L G S C P Q T A L S F - - Q A D L V R E V L E I D A S N - - K L M F G I S F G H - P D P A P V N R C V T D R A L - L S D T V H F H							
Rhodo	147 G Y A Q T L M L A M A S R G I G S C A Q G A L S L - - Y P E I V R A H L G L V Q G P - - R L I M G I A F G Y - K D P D H P A N A A H T D R A A - L A D I V H H H							
Sphin	158 M Y A Q S F M L A L T A N G L A S C A Q G A L S H - - H A E I V H R E L G V G E D Q - - K L L F G I A F G Y - - - E D A A H P A N A A R T V - R D K D H V L H							
Actino	148 M L A Q N F M L S L T A H G F G S V P Q I A L T Y - - S P D V V R Q E F N V P D N Y - - K L V L G I S F G R - P K P N S T A N S Y I Q P R A P - L S E T V V F H							
Cons_Frm2	129 m l q f a v w t a l e e e G I G A s L Q H Y N P I I d e a v a k e - - - w n i P a s w k L r A Q m p F G s i e a p - p - g e K e f q p - - - l e e R v k v f							
Smut_2IFA	129 I A L Y A I W L A L A E Q N I G M S V Q H Y N P L V D A Q V A E K - - - Y D L P T N W K M R A Q I P F G S - I E A - P A G E K E F M A - - - D Q E R F K V F							
Scere_4URP	127 A V Q I Q S W T A L E L L G L G A N L Q H Y N D Y V K S A L P - Q - - - D V P I A W T V Q S Q L V F G - - - V - P T A L P E E K - - - T F I N N V I N V							
Llact_2WQF	132 I I S V N V W T A L A E L G L G A N L Q H Y N P L I D E A V A K E - - - W N L P E S W K L R G Q L V F G S - I E A - P A G E K T F M D - - - D A D R F I V A							
Bcer_1YWQ	131 M L Q H T V W M L L S A E G I G A S L Q H Y N P I V D A E V K E T - - - W N I P A E W S L V G Q M P F G E - P N E Q P A - E R T F L P - - - T E D V V K Y F							
Cons_HUB	93 I a a e n m l l a a t e I G L G s c w v g i f - - d e - e k v r e l l - g l P - e n l e p v a l i p I g y - p a - - - e e p k p p d r k - p - l e e i v y n -							
Desulf_3E39	102 A A V Q N M L L A A H A L Q L G A V W L - - G E I N Q A A T L L P A L - A L D P A R L S F E A I A A G H - P A - - - Q N G S S - R R - P - L A E L L L E E							
Cace_3E10	95 A A A E N I L L E A A Y L D L G A V W L G A Q P I E E R V K N L K E M F - N L K - S N I K P F C V I S V G Y - P E - - N S E N K F I D R F - D - A K R I H I E K							
AMINE_CaceNitB	93 I a s t l i q v t a h s m g l g s c w - - - - - c h v r n r d r i - d y d t t - e a f i k t l i l n i - p e n l k v e c m l g i g y - p d e e r k a y t d							
Bthet_3KWK	103 A A A Q N I L L A A E S M G L G A V W T A A Y P Y E D R M E V V R K Y T - H L P E N - I L P L C V I P F G Y - P A - - - T K E Q P K Q K Y - D - E K K I H Y N Q							
Pging_3GE5	114 I A M T Y M M L E A H S L G L G C T W V C A F - - - D Q A L C S E I F - D I P S H - M T P V S I L A L G Y - G D P T V P P R E A F N R K - T - I E E V V S F E							
Cnov_3G14	106 I A M E H I V L G A T D E G Y G T C W I A A F - - - - N E N K I K E A L - K I P D N - L R V V A L T P L G V - P K D S A E D E P H H P K K - D - M D E Y L Y I D							
Gmetat_4DN2	115 L R G Q Q Y Y L T - - D V G I A A Q N L M L A A H D L G L G - - - S V F V G V F - D E - - Q Q L G E L L G I - P A E - - L R I V G L F P L - G - Y P L - - E G P							
Cdiff_3GFA	131 A A I Q N M S L T A V E L G L G S L W I C D V - Y - L F A Y R E L C E W L - N T - - D - S Q L V A A I S L G Y - P D - - - E E P S R R P R L - Q - L S D V T E W R							
Tther_1NOX	138 I L L G Y L L L L L E A Y G L G S V P M - - - - L G F D P E R V R A I L - G L P S R - A A I P A L V A L G Y - P A - - - E E G Y P S H R L - P - L E R V L W R							
Cdiff_3EO8	145 A I E N F T L S A I E L G Y G S C W L T S Q N Y A A D E I E A V L E A E T G F E K G E Y F L G A M L A L G V - P E - - - D N L K S P S K K - P - V E E I C T F I							
Mtub_4XOM	138 A A V Q A L L V A L A V R G L G S C W I G S T I F A A D L V R D E L D L P V - - - - D W E P L G A I A I G Y A D E P S G L - - R D P V P A A - - - - D L L I L K							
Tmar_4EO3	100 L F D T G L A V - - - - G N L L V Q A T Q M -							
Agro_2FRE	121 L A M Q A H L L G Y H A H G M G G I F K D -							
Gox	120 V L I Q I - - - Q A S Q A G W A V H P I S G - - - - - F D R D L A H A G L E F P E D Y Q V H A A L V I G R Q G D A G A L S E E L R R E R E Q P S G - - - - - R F							
Mmus_3GB5	185 C G L L L A A L Q N A G L V T V T T T P L N C G P - - - R L R V - L - - - L G R P S H E K L L V L L P V G Y - P S R - - - - - D A - T V P D L K R K A L - - - -							
Hsap_4TTC	191 C G I L L A A L Q N A G L V T V T T T P L N C G P - - - R L R V - L - - - L G R P A H E K L L M L P V G Y - P S K - - - - - E A - T V P D L K R K P L - - - -							
Bfrag_3EK3	122 L A A - - - - T S L G I A S C W I N Q L G T T C D D P E V R A Y L T S - L G V P E N H K V Y G C V A L G Y -							
Dhafn_3PXV	122 L A A - - - - T S L G I A S C W I N Q L G T T C D D P E V R A Y L T S - L G V P E N H K V Y G C V A L G Y -							
Smel_2ISJ	149 V C A V Q N L W L A R A E G V G V G W V - - - - S I F H E S E I K A I L - - G I P D H V E I V A W L C L G F - - - V D - - - - R L Y Q E P E L A A K G W R Q R							
Agro_3K6H	125 A V C L N V I F A A N A S G F A A N W L T E W L A F D P A - - - - F L A E I G V - - - - S A E - E K V A G Y I H I G S -							



321 331 341 351 361 371 381 391

Consensus	321	331	341	351	361	371	381	391
Conservation	-	-	-	-	-	-	-	-
Cons_NsfB	207	-	-	-	-	-	-	-
Erit_5J8D	217	-	-	-	-	-	-	-
Ecoli_1YK1	217	-	-	-	-	-	-	-
Vfisch_1VFR	218	-	-	-	-	-	-	-
Helicobac_3QDL	210	-	-	-	-	-	-	-
Idiomar_3OF4	209	-	-	-	-	-	-	-
AMINE_Styph_3HZN	220	-	-	-	-	-	-	-
Spyo_2HAY	224 K	-	-	-	-	-	-	-
Bsubt_3BEM	206	-	-	-	-	-	-	-
Lplant_4QLX	219	-	-	-	-	-	-	-
Sepider_3GBH	213	-	-	-	-	-	-	-
Esibir_3GE6	212	-	-	-	-	-	-	-
AMINE_Bcer3024	212	-	-	-	-	-	-	-
Smutants_3GAG	206	-	-	-	-	-	-	-
Spneum_2B67	204	-	-	-	-	-	-	-
Cons_NfsA	183 t Y - - - t d e - - - e a l a a Y d e t l s e y y q e r t s n q r - - t t w s e q i a a r l s k e g r - - - p h i l l e f l k k q G f l l r							
Ecoli_1F5V	180 S Y Q P L D - - - - K G A L A Q Y D E Q L A E Y Y L T R G S N N R R - D T W S D H I R R T I I K E S R - - - P F I L D Y L H K Q G W A T R							
Vharv_2BKJ	180 Q Y Q E L N - - - - L D D I Q S Y - Q T M Q A Y Y A S R T S N Q K L - S T W S Q E V T G K L A G E S R S K - - G L A K R - - - - - - -							
AMINE_Bsubt_3N2S	186 T Y N V N T D D - F - R H T M N T Y D K T I S D Y Y R E R T N G K R E - E T W S D Q I L N F M K Q K P R - - - T Y L N D Y V K E K G F N K N							
Bfrag_3EOF	182 S Y H D Y T A E D I N R L Y A Y K E S L P E N K L F I E E N Q K E T L P Q V F T - D V R Y T K K D N - E F M S E N L L K - V L R R Q G F M D -							
AMINE_Gox0834	213 R Y - - S T E N E A - K G I A A Y - D R I A D D Y K E Q G L T S R V - - - W S E T V A K R V D S F K G L S G R H V I R T V L H R L G F P L R							
AMINE_Bcer1619	183 G Y D E Q K Y D E - - - L L N E Y D E T M N A Y Y K E R S S N K K N - V T W T E S M S S F M S K E - K R - - - M H M K E F L S E K G L N K K							
AMINE_CaceNitA	185 S Y D I K A V E D - - - S I N V Y E Q - M N K Y L K E I G R A E K E - I N W - S T F T S T I Y Q - - - - S V Y Y K Q G L K T K - - - - -							
Cons_PnbA	219 g -							
AMINE_Msmg_2WZW	234 R -							
Bhen_3GR3	227 K S Y P -							
Strep	224 D T P G V L D E Q -							
Saur	225 R -							
Coryn	219 G I E G L G L -							
Nocar	232 T -							
Kpneu	225 Q -							
Dicke	237 N -							
Halio	224 D -							
Rhodo	221 G D T -							
Sphin	229 G -							
Actino	222 E -							
Cons_Frm2	199 k -							
Smut_2IFA	199 G D L E -							
Scere_4URP	192 Y H -							
Llact_2WQF	202 K -							
Bcer_1YWQ	200 -							
Cons_HUB	159 -							
Desulf_3E39	172 P F P Q P E -							
Cace_3E10	168 Y -							
AMINE_CaceNitB	160 e d l a l n k l h y d - k y s k -							
Bthet_3KWK	175 Y -							
Pging_3GE5	185 K L -							
Cnov_3G14	177 K W G T S F M E S N V K I L E K N -							
Gmetal_4DN2	180 K A G P S R K P L D E I V H Y G K Y Q A -							
Cdiff_3GFA	198 -							
Tther_1NOX	205 -							
Cdiff_3EO8	219 K -							
Mtub_4XOM	207 -							
Tmar_4EO3	147 L I A V G Y L G D E S E L S E K H R E L E R S E R V R K E L S E I V R W N L -							
Agro_2FRE	179 V P S K R V P L A D V A F E G R F T G K A D -							
Gox	186 P L A - E I A F E G R F R S N E -							
Mmus_3GB5	247 - D Q I M V T V -							
Hsap_4TTC	253 - D Q I M V T V -							
Bfrag_3EK3	179 K T V K - - - - A G T I T I V E -							
Dhafn_3PXV	171 - - - - A - N T T K P P H V P D - - - - K D K I T Y V E -							
Smel_2ISJ	216 L - - - - - - - - - - - P L E D L V F E E G W G V R -							
Agro_3K6H	185 P - - - - - - - - - - - E L A D V V T W V G D V -							

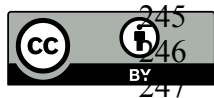
197

198 **4. References**

- 199 1. Phillips, K. L.; Sandler, S. I.; Chiu, P. C., A method to calculate the one-electron  
200 reduction potentials for nitroaromatic compounds based on gas-phase quantum  
201 mechanics. *J. Comput. Chem.* **2011**, *32*, (2), 226-239.
- 202 2. Salter-Blanc, A. J.; Bylaska, E. J.; Johnston, H. J.; Tratnyek, P. G., Predicting  
203 reduction rates of energetic nitroaromatic compounds using calculated one-electron  
204 reduction potentials. *Environ. Sci. Technol.* **2015**, *49*, 3778-3786.
- 205 3. La-Scalea, M. A.; Menezes, C. M. S.; Juliao, M. S. S.; Chung, M. C.; Serrano, S. H.  
206 P.; Ferreira, E. I., Voltammetric Behavior of Nitrofurazone and its Hydroxymethyl  
207 Prodrug with Potential Anti-Chagas Activity. *J. Braz. Chem. Soc.* **2005**, *16*, (4),  
208 774-782.
- 209 4. Hansch, C.; Leo, A.; Taft, R. W., A survey of Hammett substituent constants and  
210 resonance and field parameters. *Chem. Rev.* **1991**, *91*, (2), 165–195.
- 211 5. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.;  
212 Meng, E. C.; Ferrin, T. E., UCSF Chimera - a visualization system for exploratory  
213 research and analysis. *J. Comput. Chem.* **2004**, *25*, (13), 1605-1612.
- 214 6. Gwenin, V. V.; Poornima, P.; Halliwell, J.; Ball, P.; Robinson, G.; Gwenin, C. D.,  
215 Identification of novel nitroreductases from *Bacillus cereus* and their interaction  
216 with th CB1954 prodrug. *Biochem. Pharmacol.* **2015**, *98*, 392-402.
- 217 7. Chaignon, P.; Cortial, S.; Ventura, A. P.; Lopes, P.; Halgand, F.; Laprevote, O.;  
218 Ouazzani, J., Purification and identification of a *Bacillus* nitroreductase: Potential  
219 use in 3,5-DNBTF biosensing system. *Enz. Microb. Technol.* **2006**, *39*,  
220 1499-1506.
- 221 8. Yang, Y.; Lin, J.; Wei, D., Heterologous overexpression and biochemical  
222 characterization of a nitroreductase from *Gluconobacter oxydans* 621H. *Mol.*  
223 *Biotechnol.* **2016**, *58*, 428-440.
- 224 9. Kutty, R.; Bennett, G. N., Biochemical characterization of trinitrotoluene  
225 transforming oxygen-insensitive nitroreductases from *Clostridium acetobutylicum*  
226 ATCC 824. *Arch Microbiol* **2005**, *184*, 158-167.
- 227 10. Manina, G.; Bellinzoni, M.; Pasca, M. R.; Neres, J.; Milano, A.; Ribeiro, A. L. J. L.;  
228 Buroni, S.; Skovierova, H.; Dianiskova, P.; Mikusova, K.; Marak, J.; Makarov, V.;  
229 Giganti, D.; Haouz, A.; Lucarelli, A. P.; Degliacomini, G.; Piazza, A.; Chiarelli, L.  
230 R.; De Rossi, E.; Salina, E.; Cole, S. T.; Alzari, P. M.; G., R., Biological and  
231 structural characterization of the *Mycobacterium smegmatis* nitroreductase NfnB,  
232 and its role in benzothiozine resistance. *Mol. Microbiol.* **2010**, *77*, (5), 1172-1185.
- 233 11. Akiva, E.; Copp, J. N.; Tokuriki, N.; Babbitt, P. C., Evolutionary and molecular  
234 foundations of multiple contemporary functions of the nitroreductase superfamily.  
235 *Proc Natl Acad Sci U S A* **2017**, *114*, (45), E9549–E9558.
- 236 12. Akiva, E.; Brown, S.; Almonacid, D. E.; Barber, A. E.; Custer, A. F.; Hicks, M. A.;  
237 Huang, C. C.; Lauck, F.; Mashiyama, S. T.; Meng, E. C.; Mischel, D.; Morris, J. H.;  
238 Ojha, S.; Schnoes, A. M.; Stryke, D.; Yunes, J. M.; Ferrin, T. E.; Holliday, G. L.;

- 239 Babbitt, P. C., The Structure-Function Linkage Database. *Nucl. Acids Res.* **2014**, *42*,  
240 D521-D530.
- 241 13. Zubatyuk, R. I.; Gorb, L.; Shishkin, O. V.; Qasim, M.; Leszczynski, J., Exploration  
242 of density functional methods for one-electron reduction potential of nitrobenzenes.  
243 *J. Comput. Chem.* **2-1-**, *31*, 144-150.

244



© 2017 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).

248