frontiers

Supplementary Material

1 Additional Information

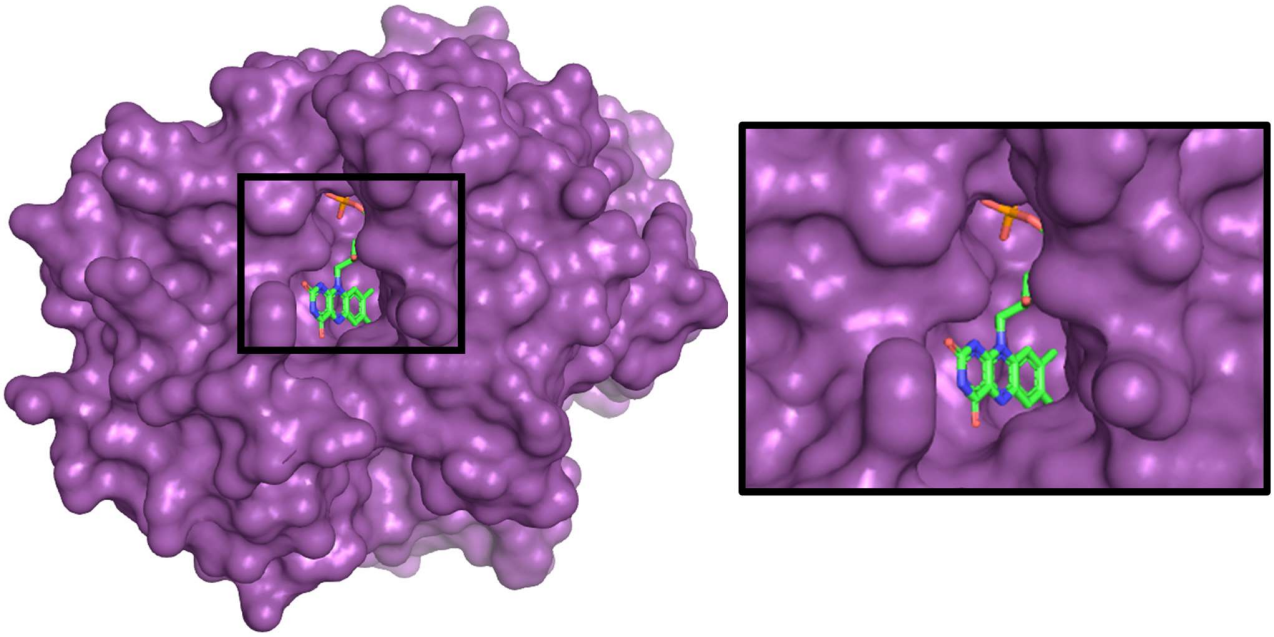
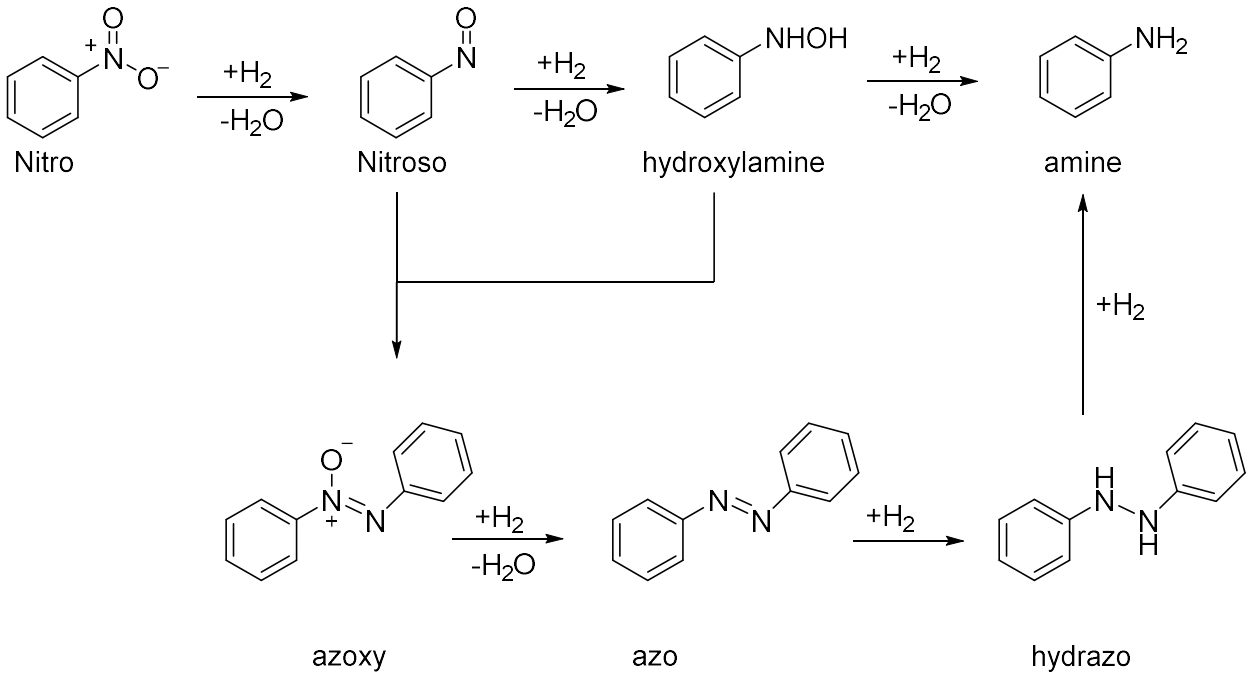


Figure S1. Prosthetic flavin (green, blue, red) in nitroreductase is bound, but surface-exposed. Space-filling representation (purple) of *E. coli* NfsB nitroreductase (prepared using PyMOL 2.3.4, PDB: 3X22).



Scheme S1. Reduction of benzonitrile to aniline and the possible intermediates and side products.

2 Additional Experiments and Data

2.1 Hydrogenase 1 overexpression and purification

Table S1

Plasmids	**Description**	**Source**
pQE80_Hyd1	pQE80L vector encoding the *HyaA-F* operon for the synthesis	This work
pACYCDuet_HypF	pACYCDuet vector encoding the HypF gene in one of the two multiple cloning sites	This work
pACYCDuet_hypAE-HypF	pACYCDuet vector encoding the HypA-F genes in two different multiple cloning sites	This work
Strains		
DH5α	fhuA2 Delta(argF-lacZ)U169 phoA glnV44 Φ80 Delta(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	ThermoFisher Scientific UK
MC4100	*E. coli* K-12. F^-,λ^- $[ara\text{D}139]_{B/r}$, Δ(*argF-lac*)U169, e14-, flhD5301, Δ(*fruK-yeiR*) 725(*fruA25*), *relA1*, *rpsL150* (Str^R), *rbsR22*, Δ(*fim*B-*fim*E)632(::*IS1*), *deoC1*	(Casadaban and Cohen, 1979. Peters *et al.* 2003.)
FTH004	*E. coli* MC4100, $hyaA^{6His}$	(Dubini *et al.*, 2002)
FTD147(DE3)J	*E. coli* MC4100, Δ*hyaB*, Δ*hybC*, Δ*hycE* Δ*MetJ*, DE3	(Skibinski *et al.*, 2002; Redwood *et al.*, 2008; Weyman *et al.*, 2011); F. Sargent (unpublished)

Table S2. Primers used in this study

Hyd1_F	ACAGAATTCATTAAAGAGGAGAAATTAACTATGAATAACGAGGAAAC ATTTTACCAGGCC
Hyd1_R	GATCTATCAACAGGAGTCCAAGCTCAGCTATTACGTCGGTGCAGCTTC CGCCAGCCACTG
pQE80_F	CAGTGGCTGGCGGAAGCTGCACCGACGTAATAGCTGAGCTTGGACTC CTGTTGATAGATC
pQE80_R	GGCCTGGTAAAATGTTTCCTCGTTATTCATAGTTAATTTCTCCTCTTTA ATGAATTCTGT
NdeI_HypF_F	TGTAATCATATGGCAAAAAACACATCTTG
XhoI_HypF_R	ACTAGACTCGAGTTATCCGTTCTGGACTTCAC
Gibs_HypA_F	GTTTAACTTTAATAAGGAGATATACCATGGGGATGCACGAAATAACC CTCTGCCAACGGG
Gib_HypE_R	TCGACTTAAGCATTATGCGGCCGCAAGCTTTTAGCATATACGCGGAAG CGGTTCGGCGTG
Gibs_PACYC_F	CACGCCGAACCGCTTCCGCGTATATGCTAAAAGCTTGCGGCCGCATAA TGCTTAAGTCGA
Gibs_PACYC_R	CCCGTTGGCAGAGGGTTATTTCGTGCATCCCCATGGTATATCTCCTTAT TAAAGTTAAAC

2.1.1 Methods for overexpression of the *E. coli* Hydrogenase-1 enzyme

Plasmid Cloning

To increase production yields of the *E. coli* Hydrogenase-1 (Hyd1) enzyme, we developed an overexpression strategy using two plasmids encoding the structural and maturation proteins involved in the synthesis of Hyd1. Genomic DNA from *E. coli* strain FTH004 was used as template for the amplification of the operon encoding the *HyaA, HyaB, HyaC, HyaD, HyaE* and *HyaF* genes using primers Hyd1_F and Hyd1_R (amplicon size 5578 nucleotides). The strain FTH004 was engineered to encode a 6xHis tag at the C-terminus of the small structural subunit HyaA, resulting in a 6xHis tag also encoded in the overexpression plasmid. The plasmid backbone DNA was generated by amplification of the circular pQE80L plasmid (QIAGEN) with primers pQE80_F and pQE80_R. After gel isolation of the two PCR products, a circular plasmid was generated by Gibson Assembly in a one-step isothermal in vitro assembly reaction (Gibson *et al.*, 2009) resulting in plasmid pQE80_Hyd1. Circular plasmids were transformed into *E. coli* DH5α cells and verified by sequencing.

A second plasmid encoding the *hyp* genes, required for the [NiFe] co-factor assembly (Sargent, 2016), was generated as follows. First, the gene encoding the hypF protein was amplified directly from

genomic DNA of *E. coli* strain FTH004 with primers NdeI_HypF_F and XhoI_HypF_R. After PCR amplification, the product was digested with NdeI and XhoI restriction enzymes (New England Biolabs) and inserted into one of the two multiple-cloning-sites of the pACYCDuet-1 plasmid (Novagen) that was previously digested with the NdeI and XhoI restriction enzymes. The ligation product, plasmid pACYCDuet-HypF, was transformed into *E. coli* DH5α cells and verified by sequencing. Second, the operon containing the *hypA, hypB, hypC, hypD,* and *hypE* genes (3680 nucleotides) was amplified with primers Gibs_HypA_F and Gib_HypE_R, also from *E. coli* strain MGH004. The backbone DNA was generated by amplification of the previously cloned pACYCDuet_HypF plasmid with primers Gibs_PACYC_F and Gibs_PACYC_R (6199 nucleotides). After gel isolation of the two PCR products, a circular plasmid was generated by Gibson Assembly in a one-step isothermal *in vitro* assembly reaction resulting in plasmid pACYCDuet_HypAE-HypF. This plasmid was transformed into *E. coli* DH5α cells and verified by sequencing.

Bacterial Growth and over-expression

Protein overexpression was carried out using the bacterial strain *E. coli* FTD147(DE3)J (Table S1). This strain does not encode for the large structural subunits of the *E. coli* Hydrogenases 1, 2 and 3. The two plasmids encoding the accessory and structural genes of Hyd1 (pQE80_Hyd1 and pACYCDuet_hypAE-HypF) were sequentially transformed into this strain. After transformation, a pre-inoculum was grown from a single colony in 10 mL of Luria-Bertani (LB) medium containing 100 μg mL^{-1} ampicillin and 34 μg mL^{-1} chloramphenicol. The tube was incubated for 15 hours at 37 ˚C, 220 rpm. This pre-culture was used to inoculate 2.4 L of LB medium containing 0.5% wv^{-1} glycerol, 0.4% wv^{-1} sodium fumarate, 100 μM $FeCl_3$, 30 μM $NiCl_2$, 100 μg mL^{-1} ampicillin, and 34 μg mL^{-1} chloramphenicol inside a glass bottle filled to the rim. The culture was incubated without shaking anaerobically at 37 ˚C. Once the bacterial growth reached the stationary phase (approximately 15 hours later) protein over-expression was initiated by adding isopropyl *β*-D-1-thiogalactopyranoside (IPTG) (final concentration 0.2 mM). The bottle was further incubated anaerobically at 37 ˚C for an additional 8 hours. Subsequently, cells were harvested aerobically by centrifugation at 6000 × g, 4 ˚C for 20 minutes, flash-frozen in liquid N_2 and stored at -80 ˚C.

Purification

Protein purification was done as previously reported with some modifications (Rowbotham *et al.*, 2020). Frozen cell pellets were thawed at room temperature and resuspended in Buffer A (100 mM Tris, 350 mM NaCl, 3% TritonX-100, pH 7.5). The buffer was supplemented with Complete EDTA-free protease inhibitor tablets (Roche Molecular Biochemicals). Cells were disrupted by sonication (Fisherbrand™ Q500 Sonicator fitted with standard 0.5-inch probe; Amp 30, 2-second pulse, 5-second pause, total sonication of 5 minutes) and centrifuged at 18,000 × g for 1 hour at 4 ˚C. After centrifugation, filtered supernatant (0.45 μm syringe filter unit) was passed through a previously equilibrated 5 mL Ni-NTA HisTrap FF column (GE Healthcare Life Sciences) using an ÄKTA Start protein purification system (GE Healthcare Life Sciences). The column was washed with 50 mL of Buffer B (20 mM Tris, 350 mM NaCl, 60 mM imidazole, 0.02% Triton X-100, 1 mM DTT, at pH 7.2). The target protein was eluted with an imidazole gradient between 60 mM and 750 mM imidazole in Buffer B. Fractions containing the target protein were pooled and concentrated to 1 mL volume using an Ultra Centrifugal Filter Units (50,000 kDa cut-off molecular weight, Amicon™) and further purified using size exclusion chromatography in a HiLoad 16/600 Superdex 200PG column (GE Healthcare Life Sciences) using Storage Buffer (50 mM Tris, 350 mM NaCl, 1 mM DTT, at pH 7.2). Pure Hyd1 was collected and concentrated. Aliquots were flash-frozen in liquid nitrogen and stored at −80 ˚C until use.

2.2 Method for determining conversion of 1a to 1b using GC-FID to evaluate different NRs with the H_2/Hyd1/FMN system

Gas chromatography analysis for the reduction of **1a** to **1b** was not straightforward due to differences in the response factor of the GC-FID (see Figure S2). To resolve this issue, a standard solution containing equimolar concentrations of **1a** and **1b** was prepared to calculate a valid ratio of peak integrations that can reflect the original amounts of compounds in the chromatogram.

GC analysis samples were prepared exactly as how a reaction mixture sample would be given for analysis: 300 µL each of 10 mM **1a** and **1b** were made in potassium phosphate buffer (100 mM, pH 7.0) with 5 vol% DMSO, then 5 µL **1a** + 5 µL **1b** were taken, mixed with 200 µL EtOAc, and 150 µL of the organic layer was removed, dried over Na_2SO_4, and given for analysis in GC. This procedure was repeated thrice to determine an average value of the required ratio of corrected peak area.

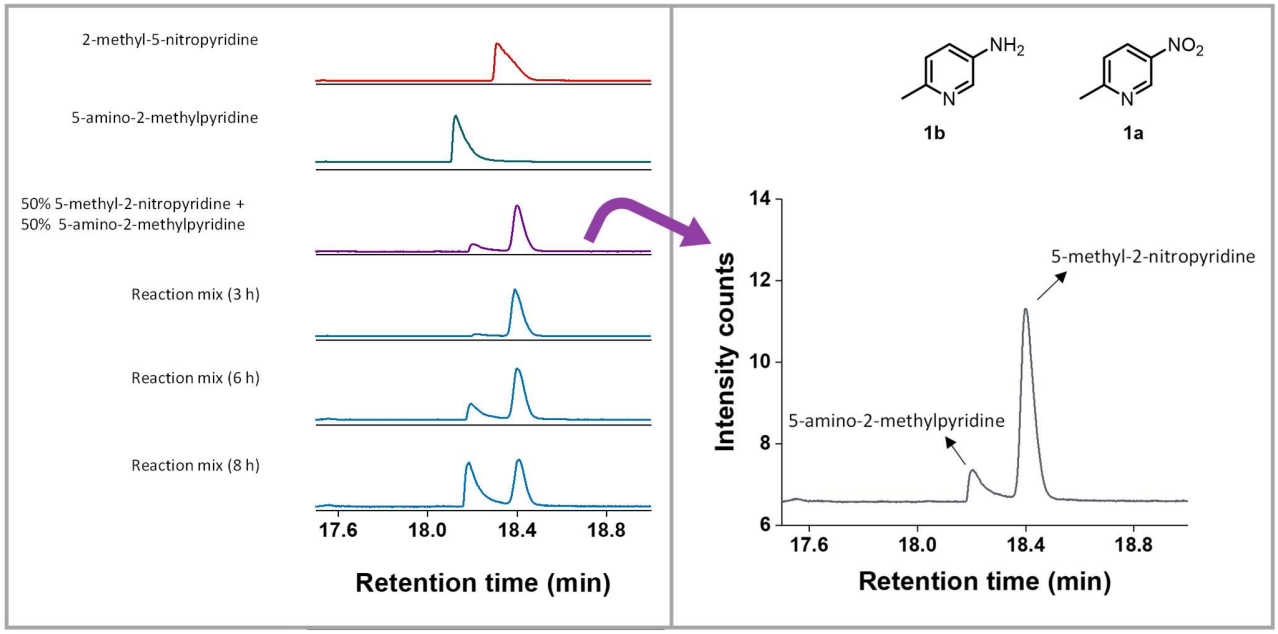


Figure S2. i) Gas chromatograms of 2-methyl-5-nitropyridine (**1a**), 5 amino-2-methylpyridine (**1b**), equimolar concentrations of **1a** and **1b**, and the reaction mixture corresponding with entry 5 in Table 1 at 3, 6 and 8 h. **ii)** Zoom in of the gas chromatogram of equimolar **1a** and **1b**.

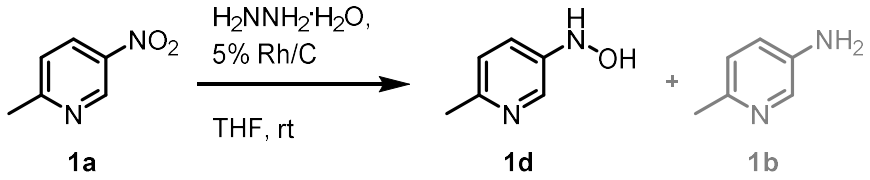
The peak areas in the spectrum shown in **Figure S2**(ii) recorded for equimolar **1a** and **1b** are 0.2977 respectively and 0.0635, respectively. Therefore,

$$\frac{peak\ area\ of\ \mathbf{1b}}{peak\ area\ of\ \mathbf{1a}} = 0.213\ \frac{1b}{\mathbf{1a}} \qquad \textbf{(equation S1)}$$

2.3 Using ^{1}H NMR spectroscopy to determine conversions of 1a to 1b and 1d

^{1}H NMR spectroscopy was selected for quantifying the compound composition of nitroreduction reactions as it did not require an extraction step which likely complicates the detection of **1b** and **1d**.

2.3.1 Preparation of an ^{1}H NMR spectroscopy standard for 1d



***N*-(6-methylpyridin-3-yl)hydroxylamine** (**1d**) was synthesized in accord with Feng *et al.*, 2016: 2-methyl-5-nitropyridine (**1a**, 200 mg, 1.45 mmol) and 5 wt% Rh/C (24 mg) were stirred in THF (7.5 ml) under an N_2 atmosphere at room temp. To this suspension, hydrazine monohydrate (0.1 mL, 2.05 mmol) was added dropwise and the reaction was allowed to stir at room temperature. After 2 h 45 min, the suspension was filtered through a pad of celite, the filtrate was extracted into EtOAc, then concentrated *in vacuo* to obtain 20.7 mg (11% yield) of a yellow solid. ^{1}H NMR spectroscopy showed a 1 : 0.7 mixture of amine **1b** and hydroxyl amine **1d** (see Figure S3). The amine (**1d**) spectrum was confirmed against the commercially obtained product standard.

^{1}H NMR of **1d** (400 MHz, 9:1 H_2O/D_2O): δ 7.93 (d, J = 3.5 Hz, 1H), 7.13 (dd, J = 8.4, 2.6 Hz, 1H), 7.02 (d, J = 8.4 Hz, 1H), 2.67 (s, 3H).

^{1}H NMR of **1b** (400 MHz, 9:1 H_2O/D_2O): δ 7.72 (d, J = 3.5 Hz, 1H), 7.24 (dd, J = 8.4, 2.8 Hz, 1H), 7.07 (d, J = 8.4 Hz, 1), 2.25 (s, 3H).

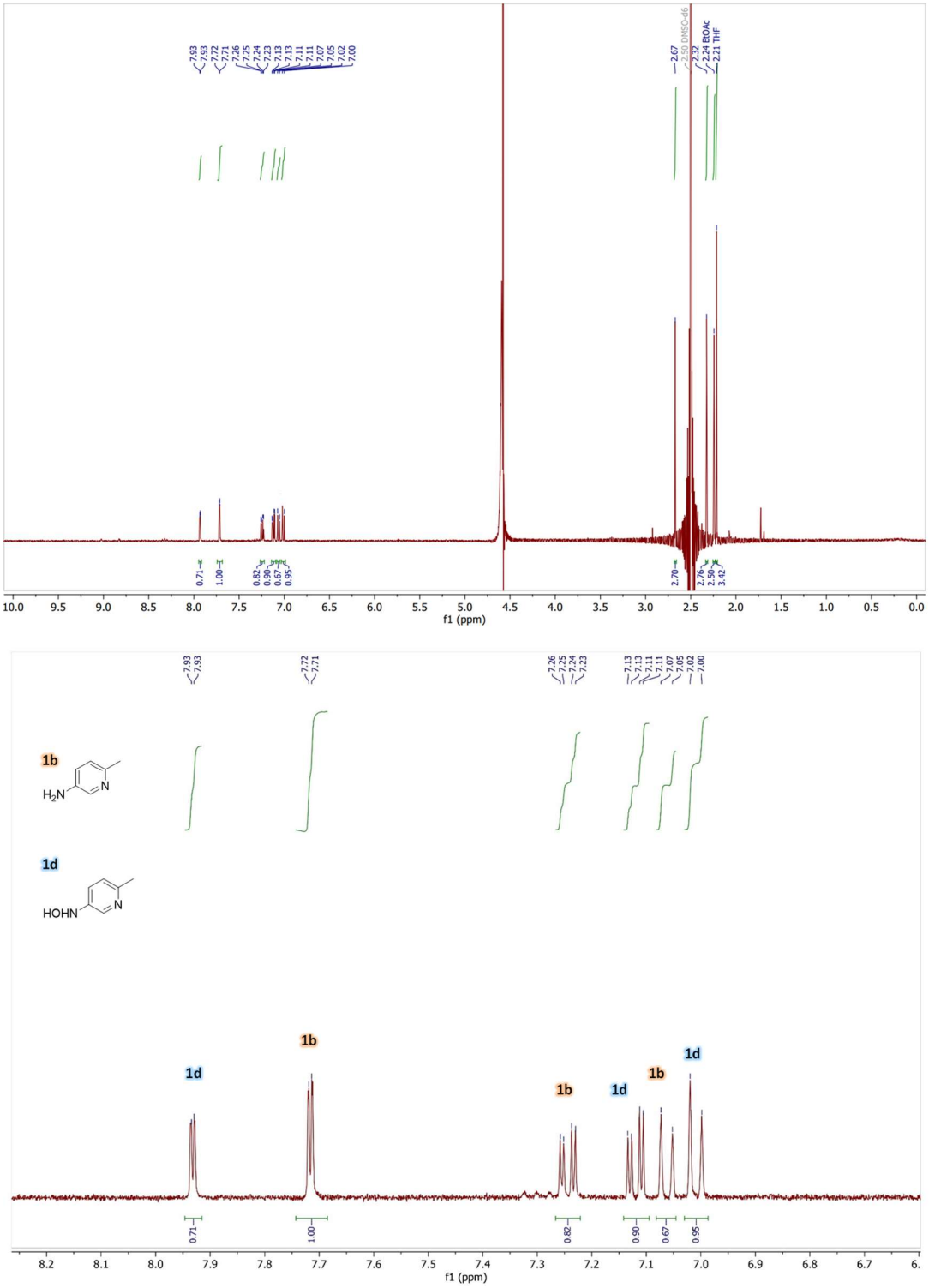


Figure S3. ^{1}H NMR spectrum of the isolated material in 9:1 H_2O/D_2O, which shows *ca.* 1 : 0.7 ratio **1b** to **1d** (peaks labelled b). Top: full spectrum. Bottom: zoom in of aromatic region.

2.3.2 Using ^{1}H NMR spectroscopy to determine conversions summarised in Table 2

Chemical compositions of the control experiments in Table 2 were calculated using ^{1}H NMR spectroscopy. NMR spectroscopy samples were taken directly from the reaction mixture and diluted with D_2O (9:1 H_2O/D_2O), then taken for analysis. All spectra shown here (Figures S3–S6) were run at room temperature with water suppression and were referenced to the DMSO (2.50 ppm) which was present in the reaction mixture. As shown in Figure 3, the diagnostic peaks for the three main compounds were the following: **1d** (d at 7.93 ppm), **1b** (dd at 7.24 ppm), and **1a** (d at 9.03 ppm). These peaks were integrated in order to determine relative percent composition of the reaction mixtures. However, as in entry 1, complex mixtures of products prevented accurate integration of the diagnostic peaks, thus product compositions could not be reported.

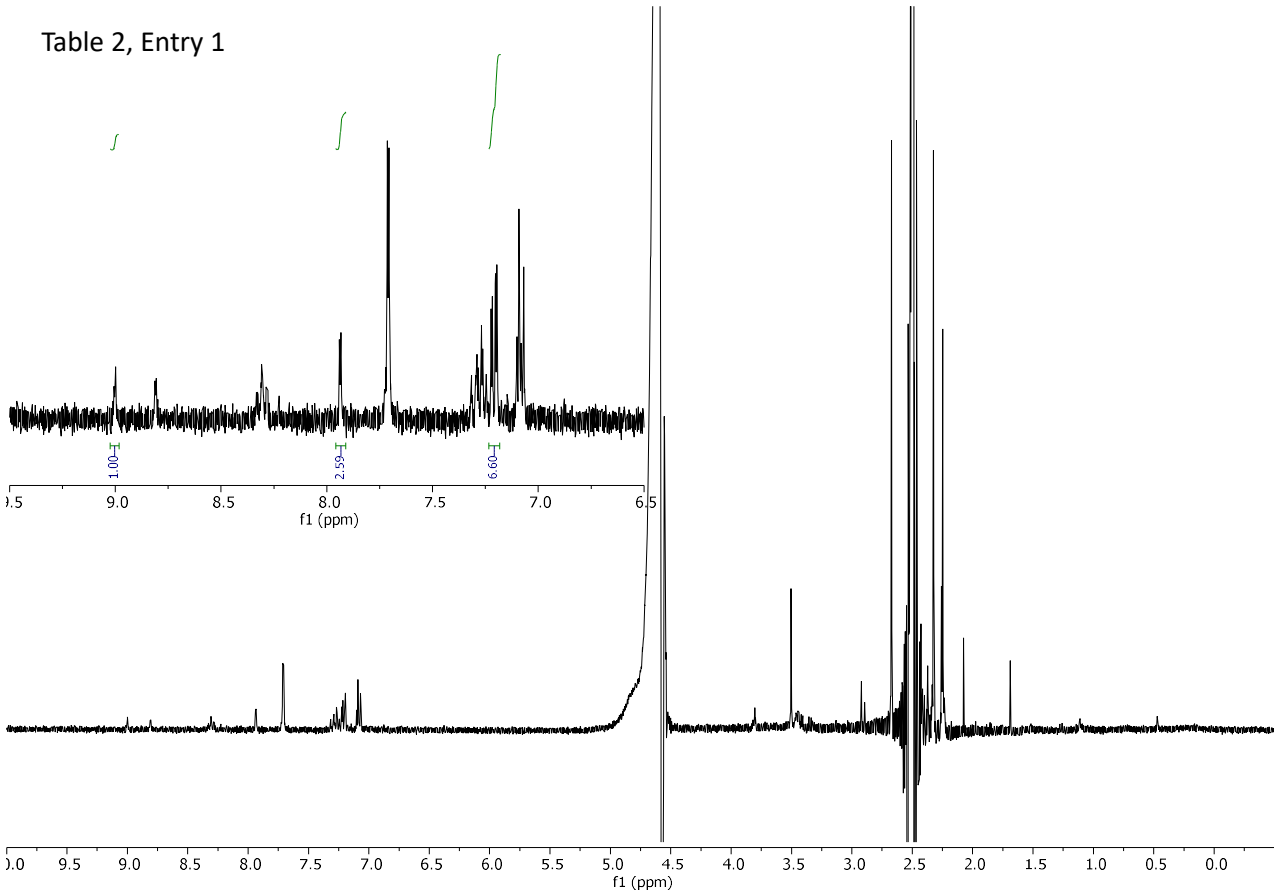


Figure S4. ^{1}H NMR spectrum (9:1 H_2O/D_2O) showing product composition of the control reaction run in the absence of V_2O_5 after 20 h (entry 1, Table 2). The presence of **1a**, **1b** and **1d** is confirmed, however the overlapping peaks (particularly 7.24 ppm) does not allow accurate integration.

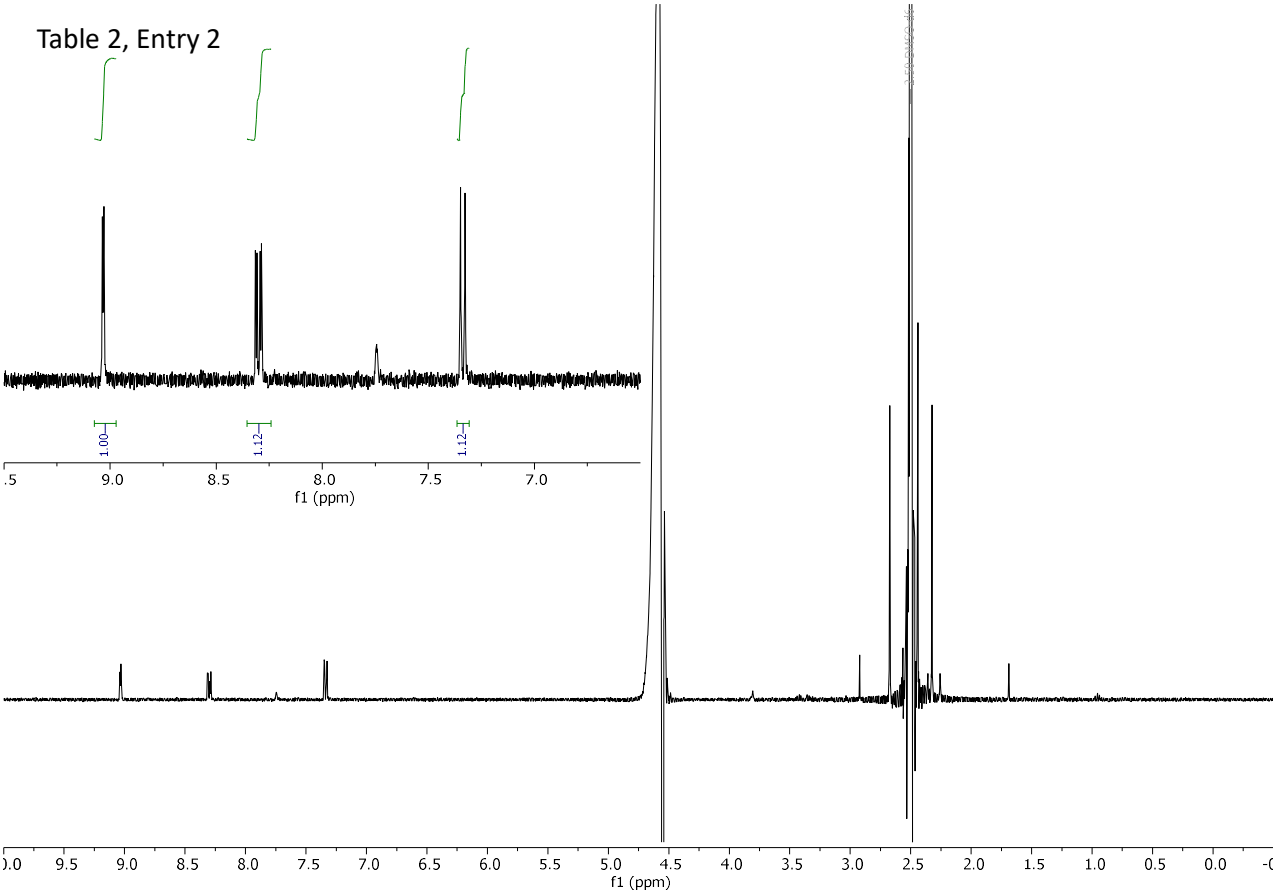


Figure S5. ^{1}H NMR spectrum (9:1 H_2O/D_2O) showing product composition of the control reaction run in the absence of Hyd1 after 20 h (entry 2, Table 2). Only **1a** is detected in the aromatic region (inset). The peak at 7.75 ppm is from the FMN cofactor (Müller, 2014).

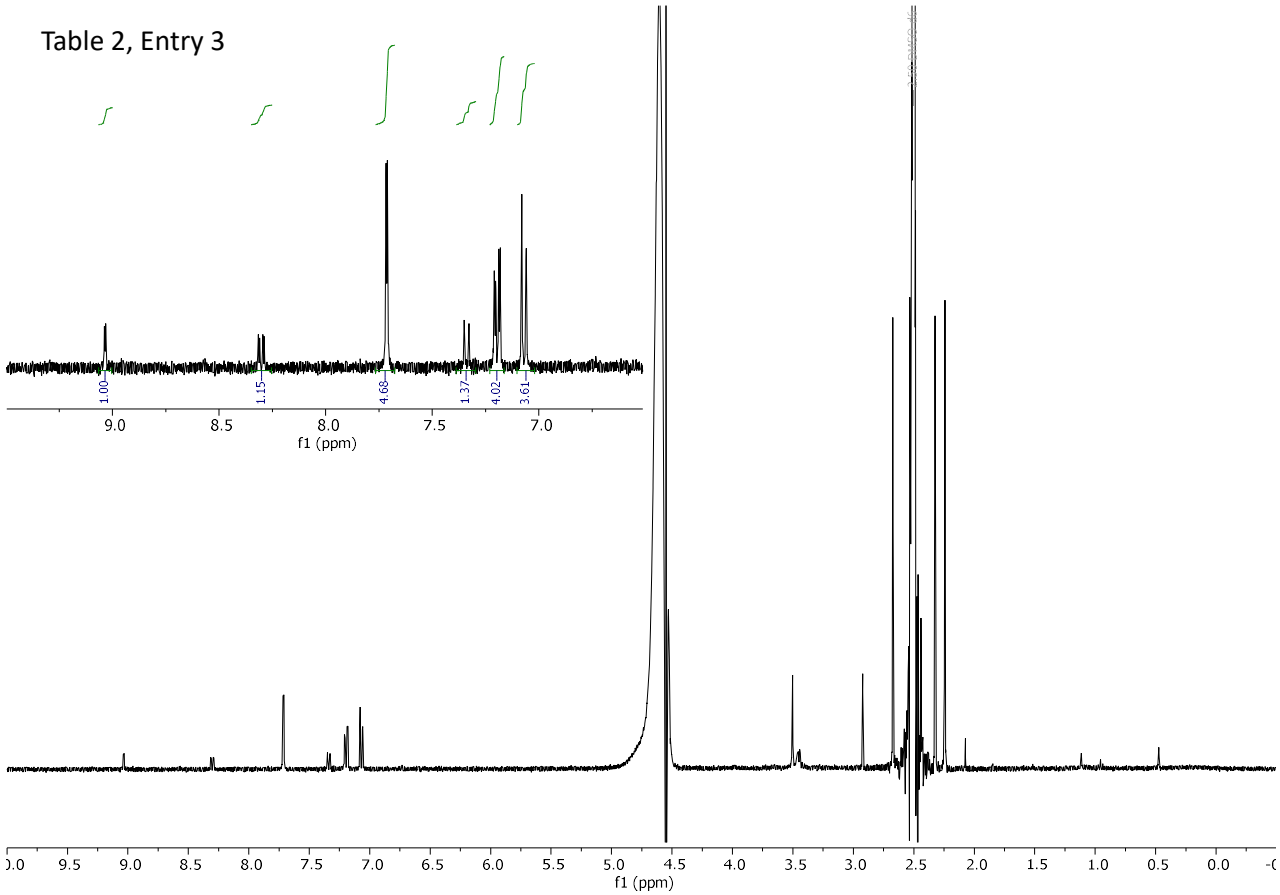


Figure S6. ^{1}H NMR spectrum (9:1 H_2O/D_2O) showing product composition of the control reaction run in the absence of NR after 20 h (entry 3, Table 2). A 4 : 1 ratio of **1b** to **1a** is noted, with no detection of **1d** (see aromatic region, inset).

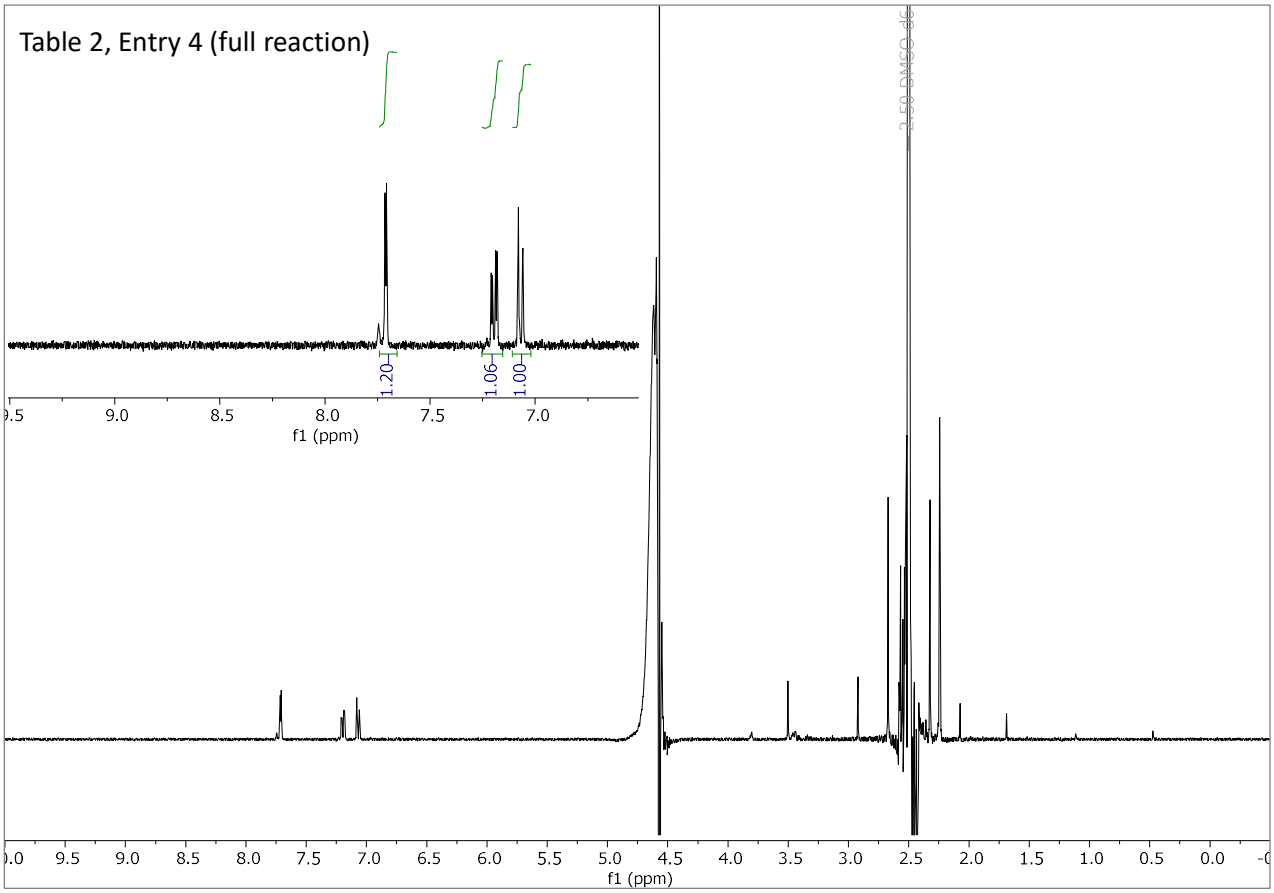


Figure S7. ^{1}H NMR spectrum (9:1 H_2O/D_2O) showing product composition of the full reaction (run with all of the reagents) after 20 h (entry 4, Table 2). Only **1b** is detected in the aromatic region (inset). The small peak at 7.75 ppm is from the FMN cofactor (Müller, 2014).

Table S3. Reproducibility for 'full reaction' conditions (entry 4, Table 2) run in duplicate

Entry	Conversion to 1b (%)[a]	
	3 h	6 h
1	8	49
2	6	46
Standard deviation	1.4	2.1

Reaction conditions: 300 µL scale, 10 mM **1a**, 1 mM FMN, 0.18 mg/mL Hyd1 (3.95 mU), 1 mg/ml NR-17, 2 mM V_2O_5 in potassium phosphate buffer (100 mM, pH 7.0) with 5 vol% DMSO at 35°C in a pressure vessel (1 bar H_2) for 20 h. [a]Determined using ^{1}H NMR spectroscopy.

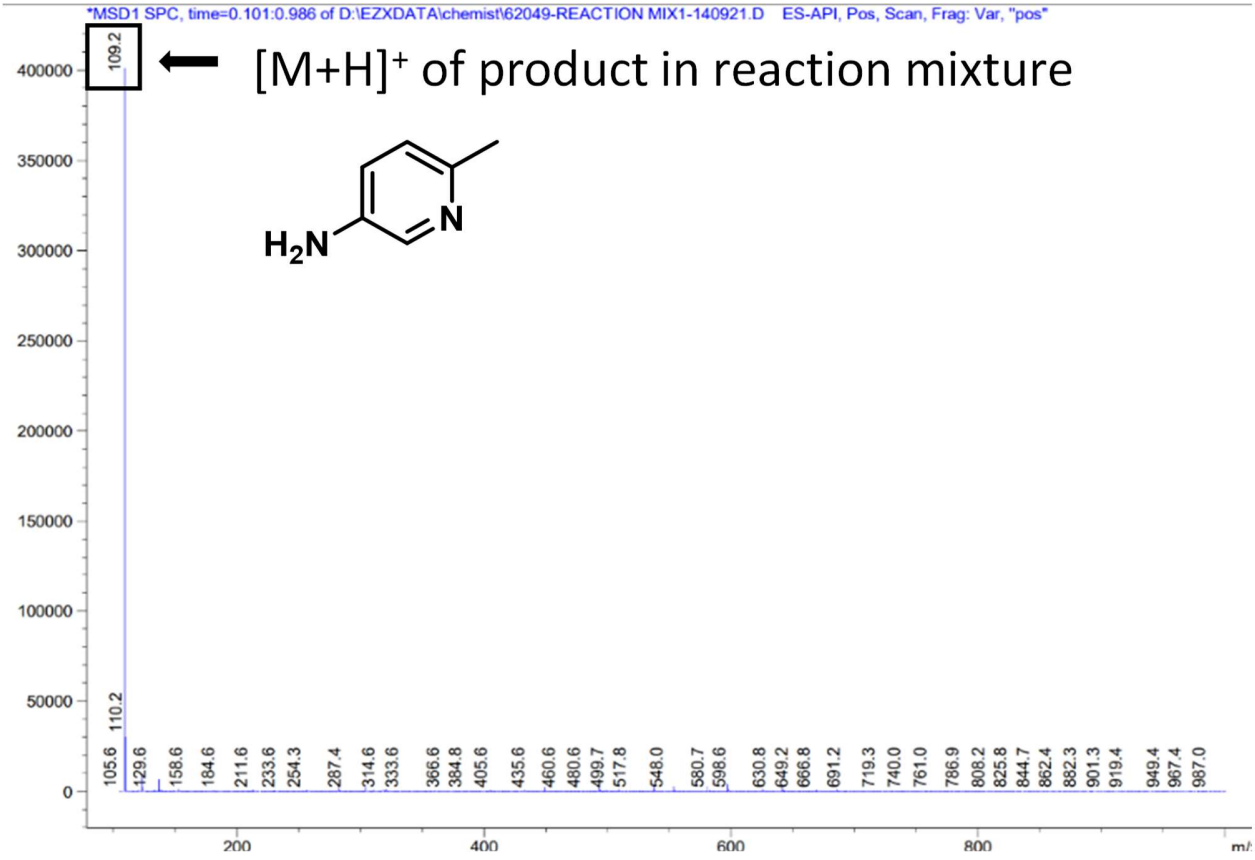


Figure S8. Mass spectrum obtained from the reaction run for 19 h in entry 5, Table 2.

3 References

Dubini, A. *et al.* (2002) 'How bacteria get energy from hydrogen: a genetic analysis of periplasmic hydrogen oxidation in Escherichia coli', *Int. J. Hydrog. Energy*. Pergamon, 27, pp. 1413–1420. doi: 10.1016/S0360-3199(02)00112-X.

Feng, P. *et al.* (2016) 'Access to a new class of synthetic building blocks via trifluoromethoxylation of pyridines and pyrimidines', *Chem. Sci.* The Royal Society of Chemistry, 7, pp. 424–429. doi: 10.1039/C5SC02983J.

Gibson, D. G. *et al.* (2009) 'Enzymatic assembly of DNA molecules up to several hundred kilobases', *Nat. Methods*. Nature Publishing Group, 6, pp. 343–345. doi: 10.1038/nmeth.1318.

Müller, F. (2014) 'NMR Spectroscopy on Flavins and Flavoproteins', in Weber, S. and Schleicher, E. (eds) *Methods Mol. Biol.* New York: Springer, pp. 229–306. doi: 10.1007/978-1-4939-0452-5_11.

Redwood, M. D. *et al.* (2008) 'Dissecting the roles of Escherichia coli hydrogenases in biohydrogen production', *FEMS Microbiol. Lett.*, 278, pp. 48–55. doi: 10.1111/j.1574-6968.2007.00966.x.

Rowbotham, J. S. *et al.* (2020) 'Bringing biocatalytic deuteration into the toolbox of asymmetric isotopic labelling techniques', *Nat. Commun.* Nature Publishing Group, 11, p. 1454. doi: 10.1038/s41467-020-15310-z.

Sargent, F. (2016) 'The Model [NiFe]-Hydrogenases of Escherichia coli', *Adv. Microb. Physiol.* Academic Press, 68, pp. 433–507. doi: 10.1016/BS.AMPBS.2016.02.008.

Skibinski, D. A. G. *et al.* (2002) 'Regulation of the hydrogenase-4 operon of Escherichia coli by the σ54-dependent transcriptional activators FhlA and HyfR', *J. Bacteriol.*, 184, pp. 6642–6653. doi: 10.1128/JB.184.23.6642-6653.2002/FORMAT/EPUB.

Weyman, P. D. *et al.* (2011) 'Heterologous expression of Alteromonas macleodii and Thiocapsa roseopersicina [NiFe] hydrogenases in Escherichia coli', *Microbiology*, 157, pp. 1363–1374. doi: 10.1099/mic.0.044834-0.