Anaerobic 5-Hydroxybenzimidazole formation from aminoimidazole ribotide: an unanticipated intersection of thiamin and vitamin B_{12} biosynthesis.

Angad P. Mehta, Sameh H. Abdelwahed, Michael K. Fenwick, Amrita B. Hazra, Michiko E. Taga, Yang Zhang, Steven E. Ealick and Tadhg P. Begley

Department of Chemistry, Texas A and M University.

Supporting Information

Synthesis of (5S)-[5-$^2H_1$]- and (5R)-[5-$^2H_1$]-ribose (41 & 45)

1,2:5,6-Di-O-isopropylidene-D-allofuranose (**35**) was synthesized using the literature method[1] starting with commercially available 1,2:5,6-Di-O-isopropylidene-α-D-glucofuranose (**33**).

The syntheses of (5S)-[5-$^2H_1$]- and (5R)-[5-$^2H_1$]-ribose[2] utilized the known deuterium-labeled ribose derivative **40**.[3] The absolute stereochemistry of this structure has been previously established by conversion to the relevant thymidine derivatives followed by 2D NMR analysis.[3,4]

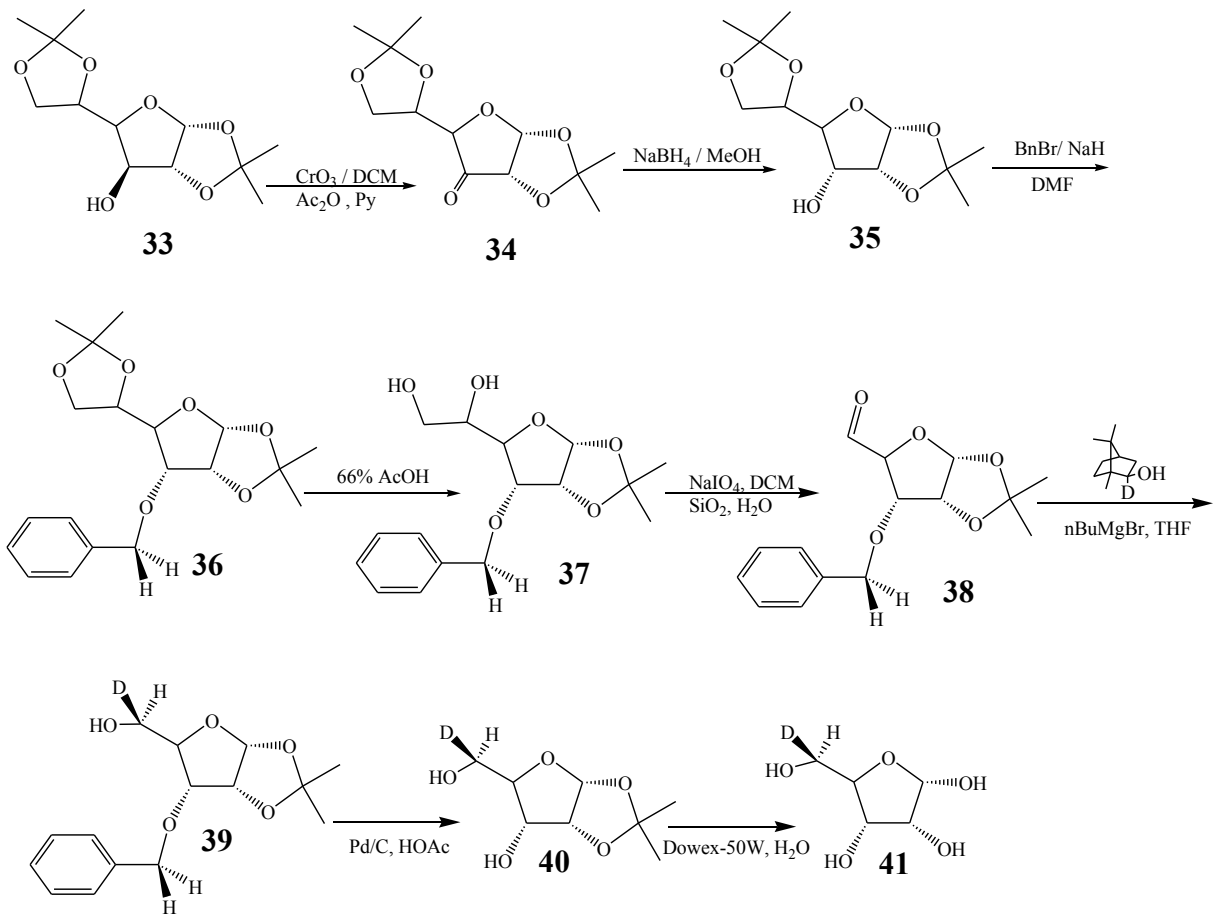


Figure S1: Scheme for the synthesis of (5S)-[5-^{2}H1]-ribose diastereomer **41.**

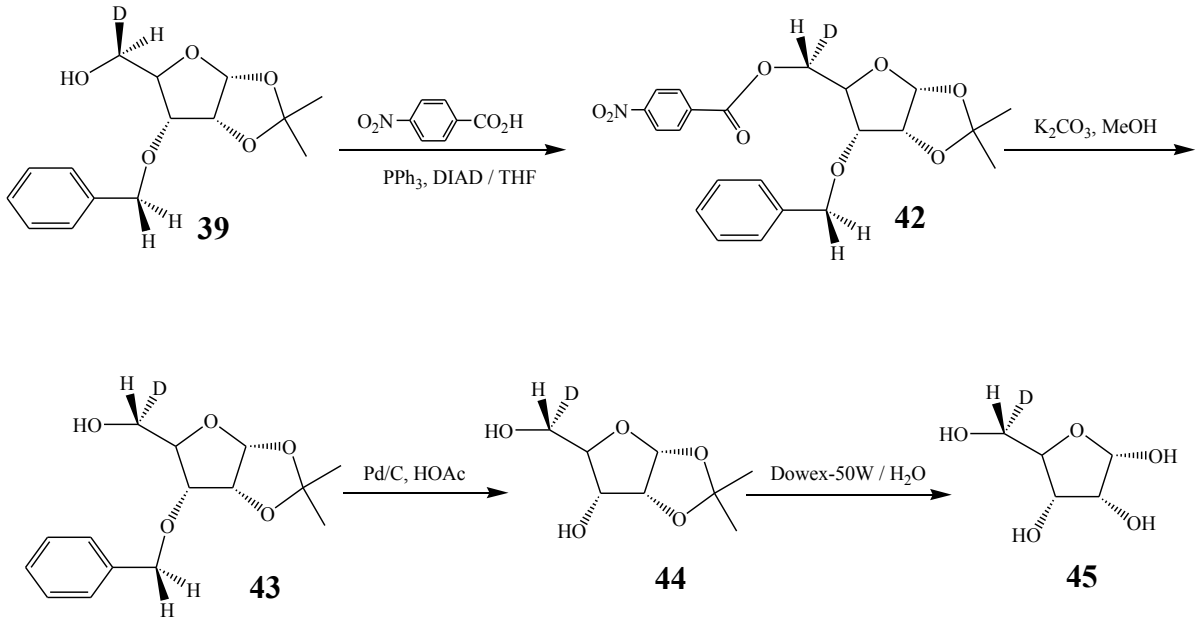
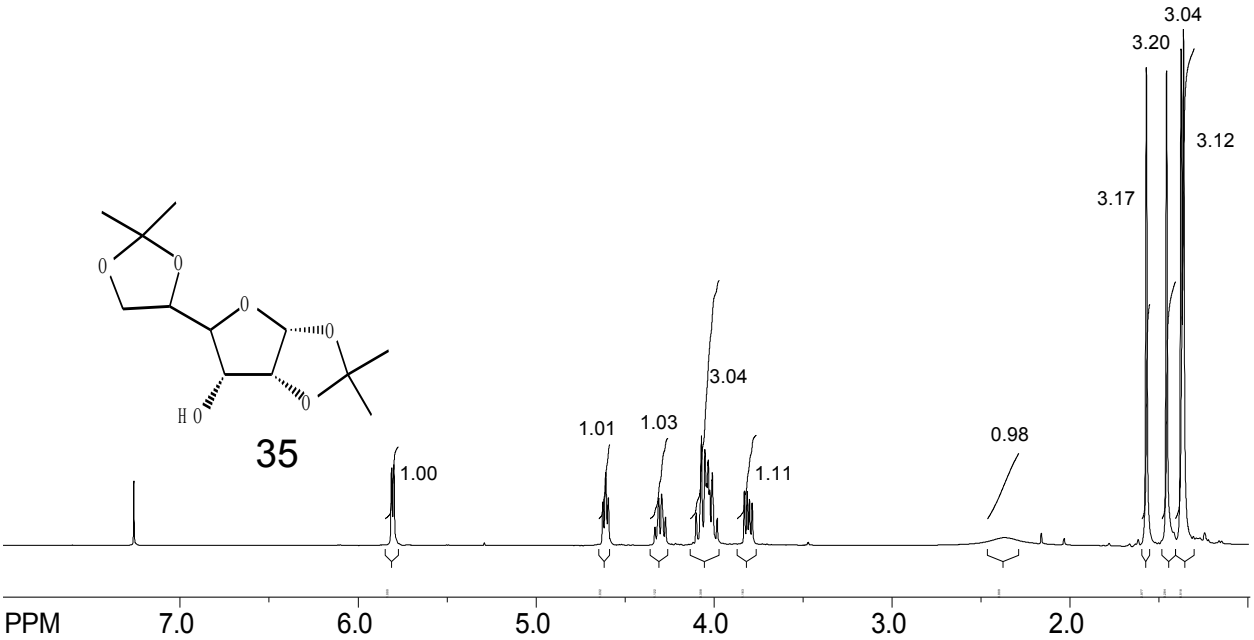


Figure S2: Scheme for the synthesis of (5R)-[5-^{2}H1]-ribose diastereomer **45.**



^{1}H NMR ($CDCl_3$) δ (ppm) 5.83 (d, 1H), 4.64 (m, 1H), 4.33 (dd, 1H), 4.07 (m, 3H), 3.85 (dd, 1H), 2.57 (br, 1H), 1.61 (s, 3H), 1.49 (s, 3H), 1.41 (s, 3H), 1.40 (s, 3H).

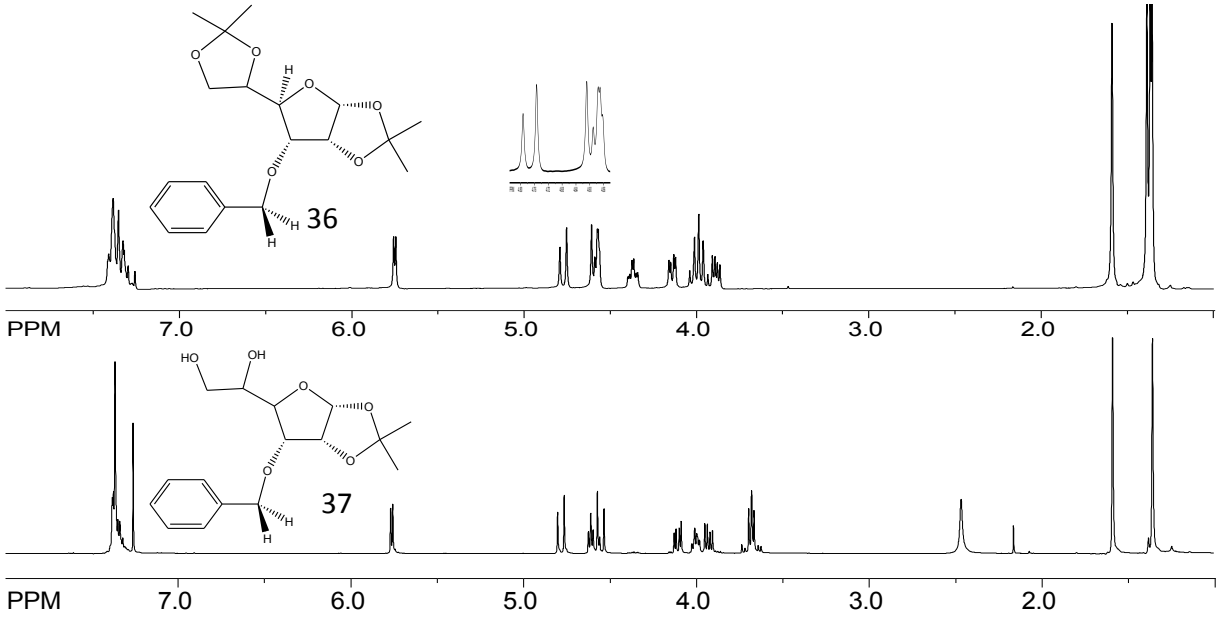


Figure S3: ^{1}H NMR ($CDCl_3$) for **35, 36 and 37**.

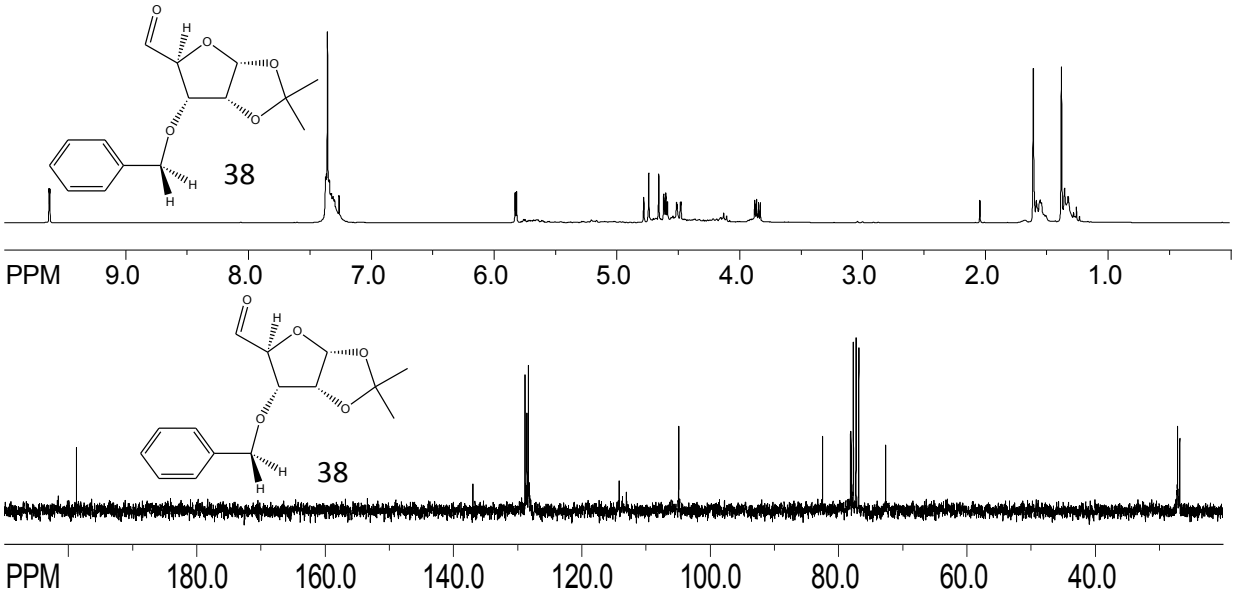


Figure S4: ^{1}H &^{13}C NMR ($CDCl_3$) for **38**.

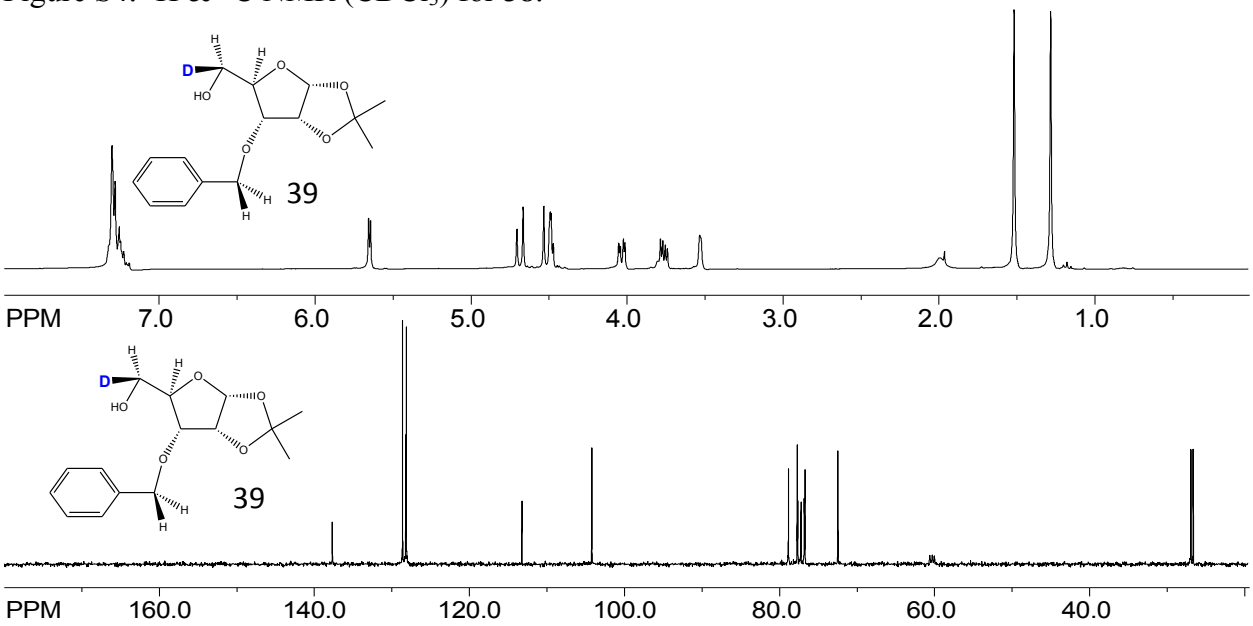
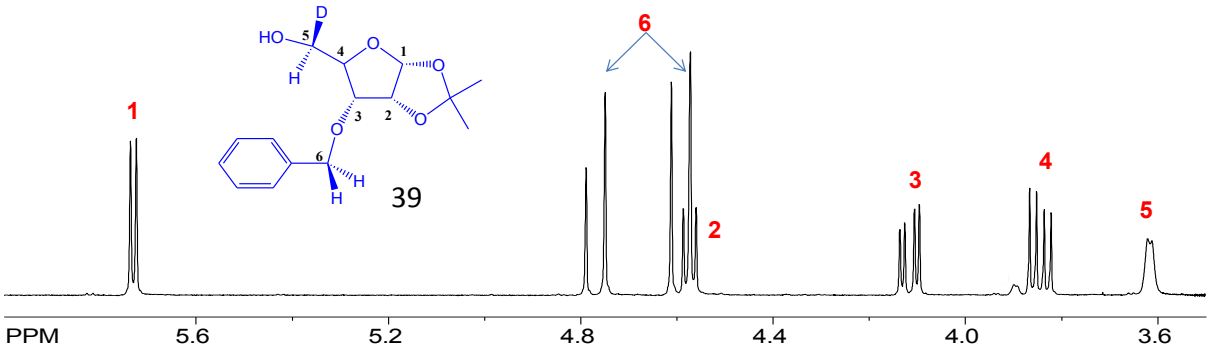
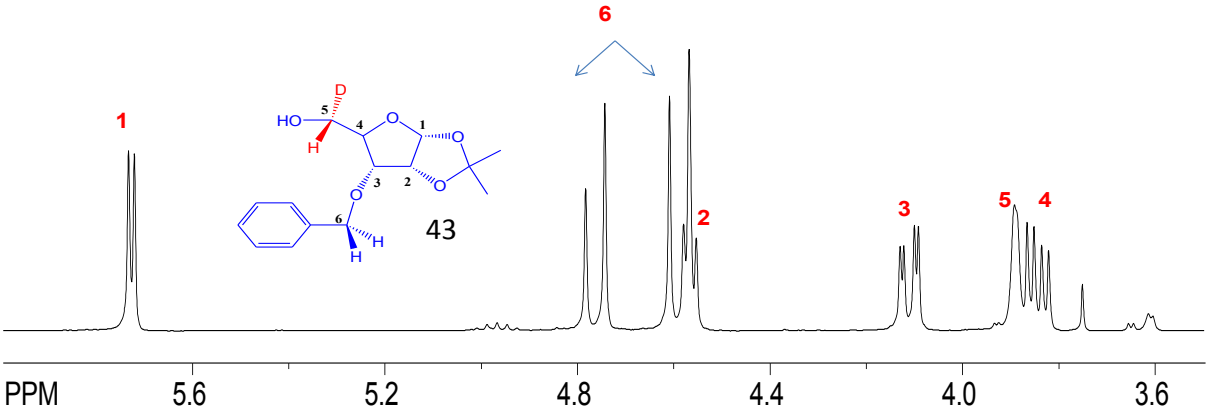


Figure S5: ^{1}H &^{13}C NMR ($CDCl_3$) for **39**.

(A)





(B)

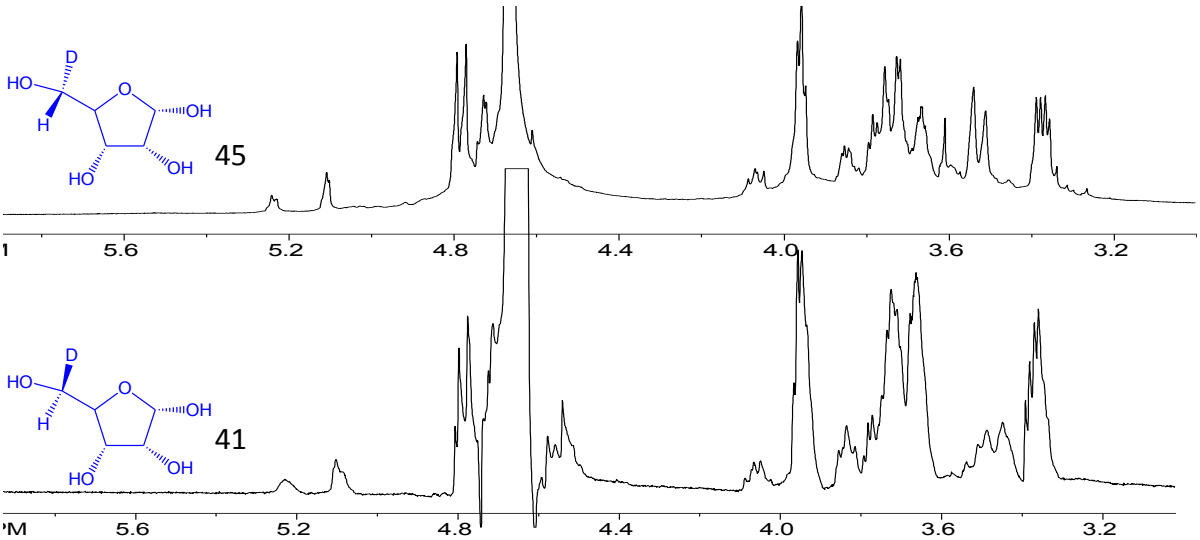


Figure S6: (A) ^{1}H NMR ($CDCl_3$) for **39 & 43** (range 3.4 – 5.8 ppm). (B) ^{1}H NMR (D_2O) for (5R)-[5-$^{2}H_1$]- and (5S)-[5-$^{2}H_1$]-ribose (45 & 41).

(5S)-[5-$^{2}H_1$]-Ribose **41**: ^{1}H NMR (D_2O) (chemical shifts for α - and β-pyranose and furanose) δ 5.38-5.20 and 4.90-4.80 (m, 1H), 4.20-4.07 (m, 1H), 3.89-3.78 (m, 2H), 3.70-3.47 (m, 1H).

(5R)-[5-$^{2}H_1$]-Ribose **45** ; ^{1}H NMR (D_2O) (chemical shifts for α - and β-pyranose and furanose) δ 5.41-5.20 and 5.00-4.80 (m, 1H), 4.24-4.09 (m, 1H), 4.00-3.75 (m, 2H), 3.73-3.45 (m, 1H).

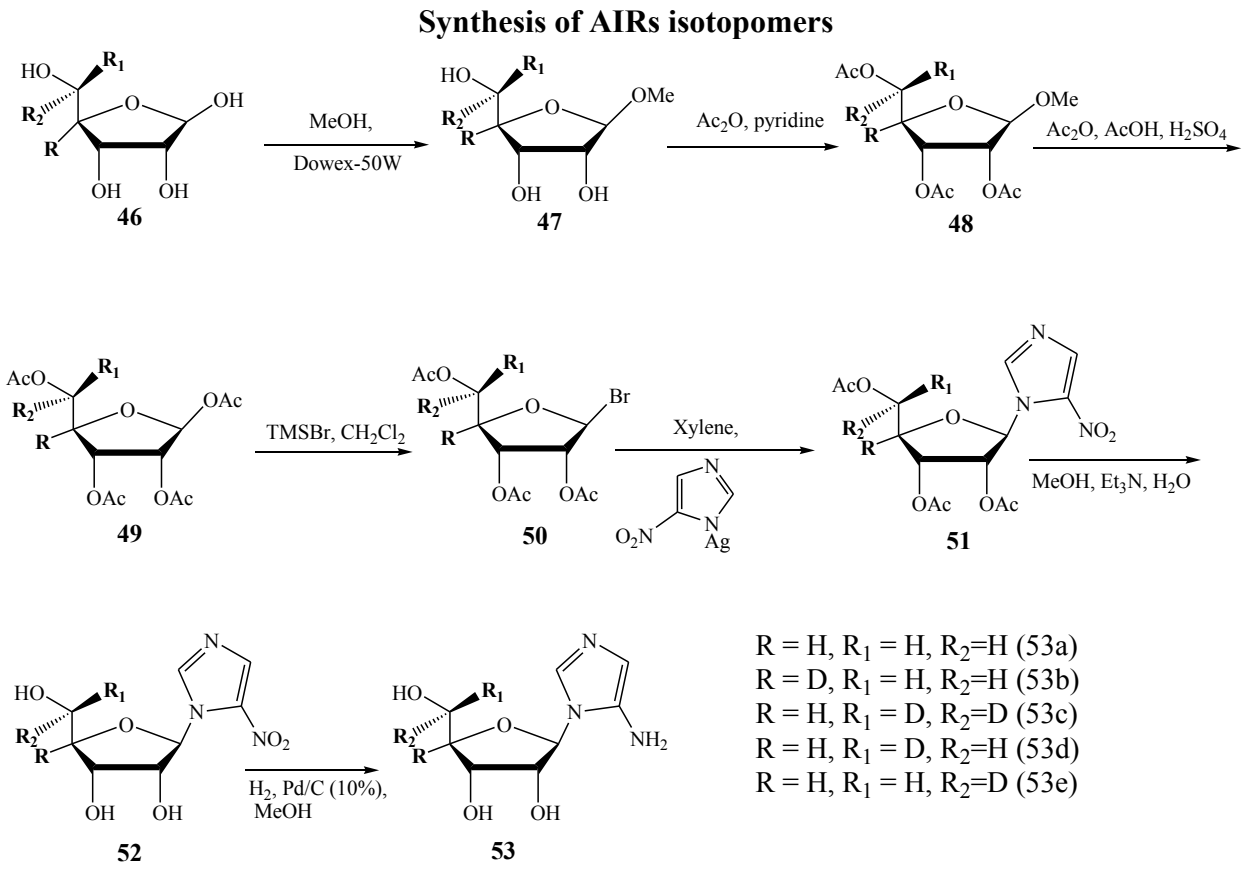


Figure S7: Synthesis of AIRs isotopomers (**52a-e**).

The syntheses and modification of AIRs isotopomers are described in our recent paper.[5,6]

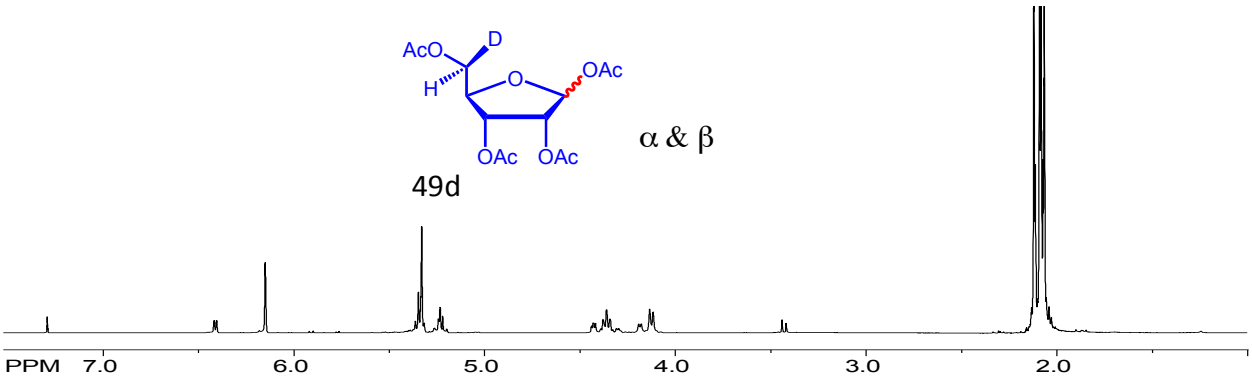


Figure S8: ^{1}H NMR ($CDCl_3$) for **49d** (R = H, R_1 = D, R_2=H).

(5S)-[5-$^{2}H_1$] Tetraacetyl ribofuranose **49d**: ^{1}H NMR ($CDCl_3$) δ 6.40 (d, 0.3, H1α), 6.13 (s, 0.7, H1β), 5.2–5.4 (m, 2, H2, H3), 4.12-4.50 (m, 1H4, 1H5), 2.10 (s, 3, CH_3CO), 2.07 (s, 3, CH_3CO), 2.06 (s, 3, CH_3CO), 2.04 (s, 3, CH_3CO).

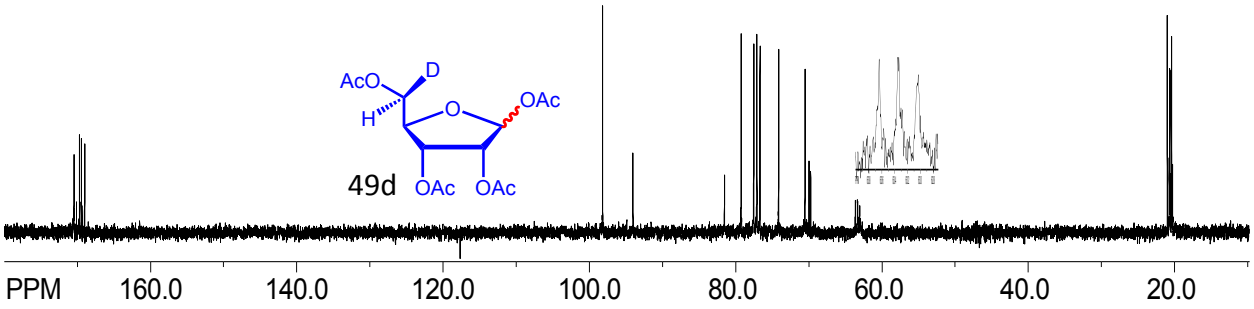


Figure S9: ^{13}C NMR ($CDCl_3$) for **49d** (R = H, R_1 = D, R_2=H).

(5S)-[5-$^{2}H_1$] Tetraacetyl ribofuranose **49d**, ^{13}C NMR ($CDCl_3$), 170.39, 169.67, 169.39, 168.96, 98.23, 93.22, 81.37, 79.38, 74.19, 70.59,69.91, 63.63 (t), 20.99, 20.67, 20.43

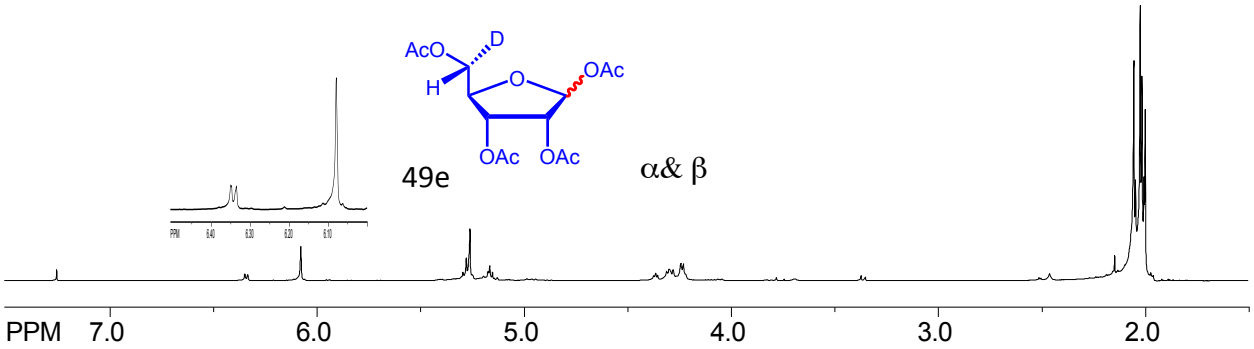


Figure S10: ^{1}H NMR ($CDCl_3$) for **49e** (R = H, R_1 = H, R_2 = D).

(5R)-[5-$^{2}H_1$] Tetraacetyl ribofuranose **49e**: ^{1}H NMR ($CDCl_3$) δ 6.40 (d, 0.3, H1′α), 6.13 (s, 0.7, H1′β), 5.2–5.4 (m, 2, H2′, H3′), 4.12-4.35 (m, 1H4′, 1H5′), 2.10 (s, 3, CH_3CO), 2.07 (s, 3, CH_3CO), 2.06 (s, 3, CH_3CO), 2.04 (s, 3, CH_3CO).

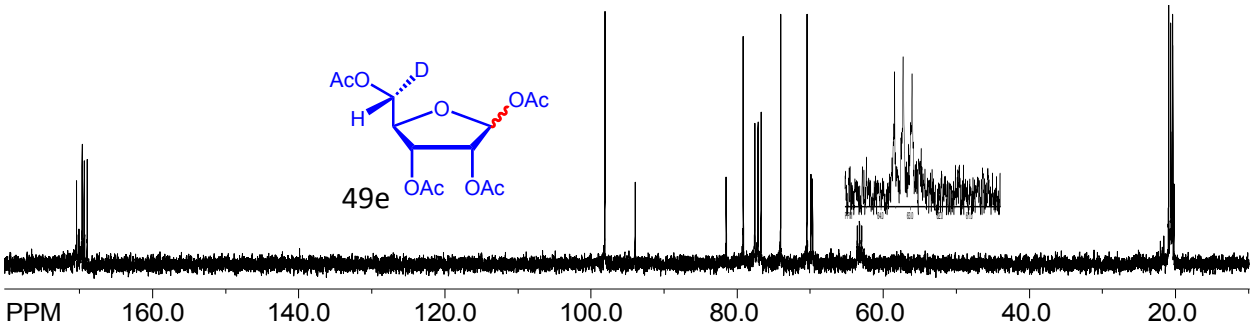


Figure S11: ^{13}C NMR ($CDCl_3$) for **49e** (R = H, R_1 = H, R_2 = D).

(5R)-[5-$^{2}H_1$] Tetraacetyl ribofuranose **49d**, ^{13}C NMR ($CDCl_3$), 170.39, 169.67, 169.39, 168.96, 98.23, 93.22, 81.37, 79.38, 74.19, 70.59,69.91, 63.63 (t), 20.99, 20.67, 20.43

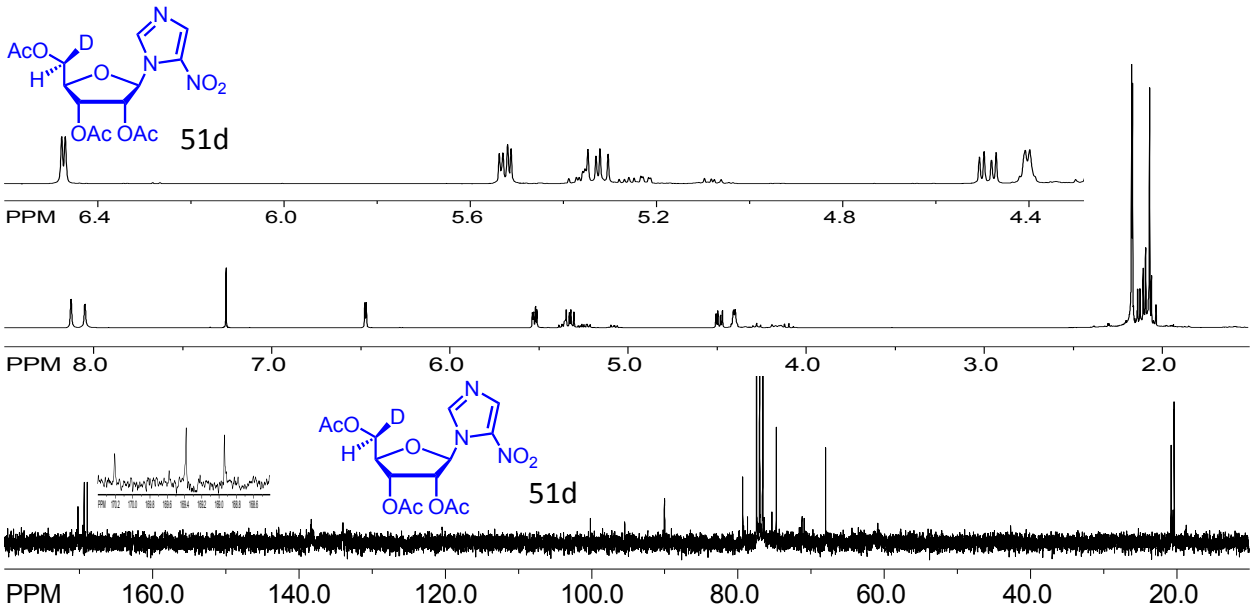


Figure S12: ^{1}H & ^{13}C NMR ($CDCl_3$) for **51d** (R = H, R_1 = D, R_2=H).

(5S)-[5-$^{2}H_1$] 5-Nitro-1-(2,3,5-tri-*O*-acetyl-β-D-ribofuranosyl)imidazole **51d**
^{1}H NMR ($CDCl_3$) δ 8.15 (s, 1H), 8.05 (s, 1H), 6.43 (d, 1H1′), 5.5 (dd, 1H2′), 5.3 (m, 1H3′), 4.5 (m, 1H4′), 4.4 (d, 1H5′) 2.16 (s, 6, 2 × CH_3CO), 2.06 (s, 3, CH_3CO).
^{13}C NMR ($CDCl_3$) :170.19, 169.57, 169.33, 138.11, 135.23, 89.43, 80.37, 79.37, 74.59, 70.03 (t), 68.91, 20.98, 20.68, 20.65

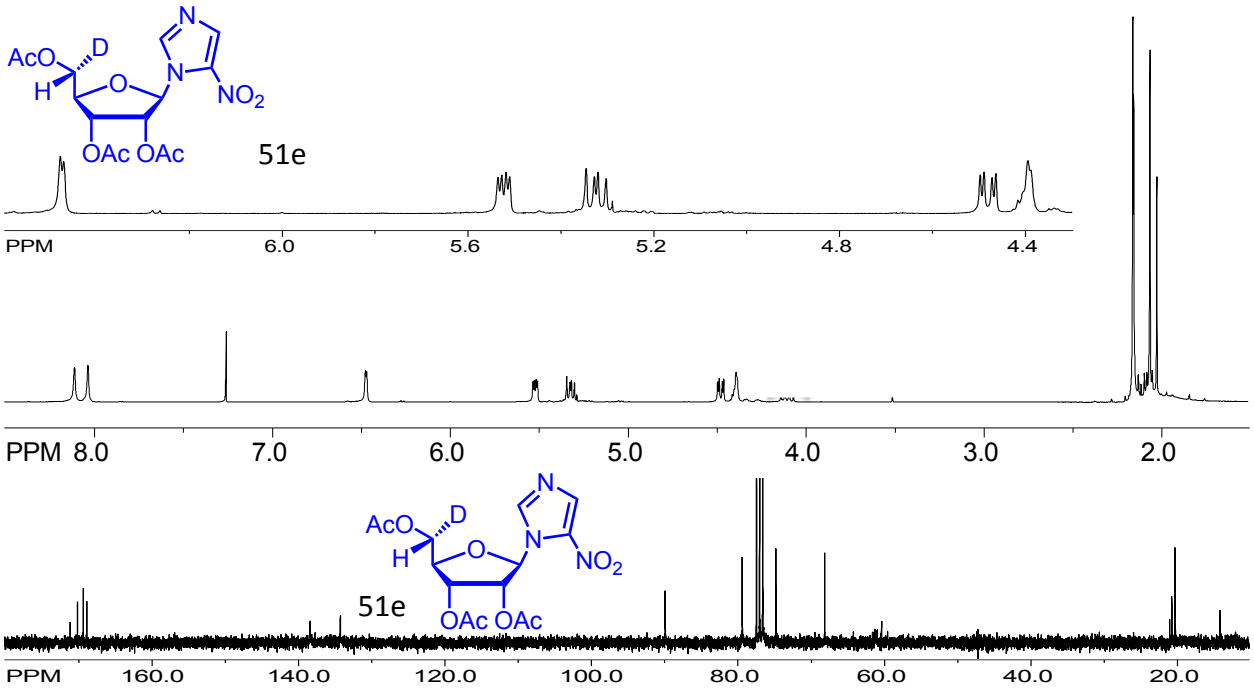


Figure S13: ^{1}H & ^{13}C NMR ($CDCl_3$) for **51e** (R = H, R_1 = H, R_2 = D).

5-Nitro-1-(2,3,5-tri-*O*-acetyl-β-D-ribofuranosyl)imidazole **51e**
^{1}H NMR ($CDCl_3$) δ 8.15 (s, 1H), 8.05 (s, 1H), 6.43 (d, 1H1′), 5.5 (d of d, 1H2′), 5.3 (m, 1H3′), 4.5 (m, 1H4′), 4.35 (d, 1H5′) 2.16 (s, 6, 2 × CH_3CO), 2.06 (s, 3, CH_3CO).
^{13}C NMR ($CDCl_3$) :170.19, 169.57, 169.33, 138.11, 135.23, 89.43, 80.37, 79.37, 74.59, 70.03 (t), 68.91, 20.98, 20.68, 20.65

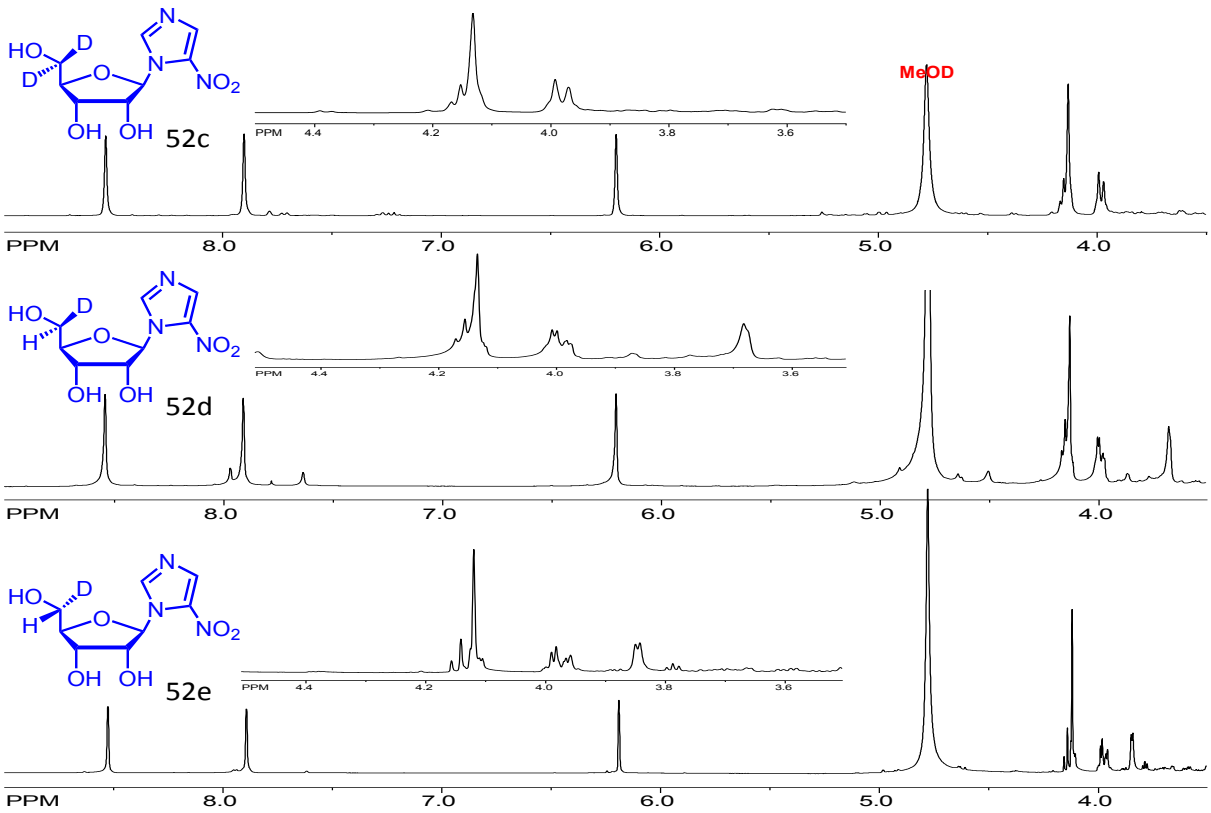


Figure S14: ^{1}H NMR (MeOD) for **52** (52c-R = H, R_1 = D, R_2 = D), (52d-R = H, R_1 = D, R_2 = H) & (52e-R = H, R_1 = H, R_2 = D) respectively.

[5′,5′-$^{2}H_2$] 5-Nitro-1-(β-D-ribofuranosyl)imidazole **52c**: ^{1}H NMR (CD_3OD) δ 8.6 (s, 1, H2), 8.0 (s, 1, H4), 6.3 (s, 1, H1′), 4.2 (m, 2, H2′, H3′), 4.0 (m, H4′).
(5S)-[5-$^{2}H_1$] 5-Nitro-1-(β-D-ribofuranosyl)imidazole **52d**: ^{1}H NMR (CD_3OD) δ 8.6 (s, 1, H2), 8.0 (s, 1, H4), 6.3 (s, 1, H1′), 4.2 (m, 2, H2′, H3′), 4.0 (m, H4′), 3.6 (d, 1, H5″).
(5R)-[5-$^{2}H_1$] 5-Nitro-1-(β-D-ribofuranosyl)imidazole **52e**: ^{1}H NMR (CD_3OD) δ 8.6 (s, 1, H2), 8.0 (s, 1, H4), 6.3 (s, 1, H1′), 4.2 (m, 2, H2′, H3′), 4.0 (m, H4′), 3.8 (d, 1, H5′).

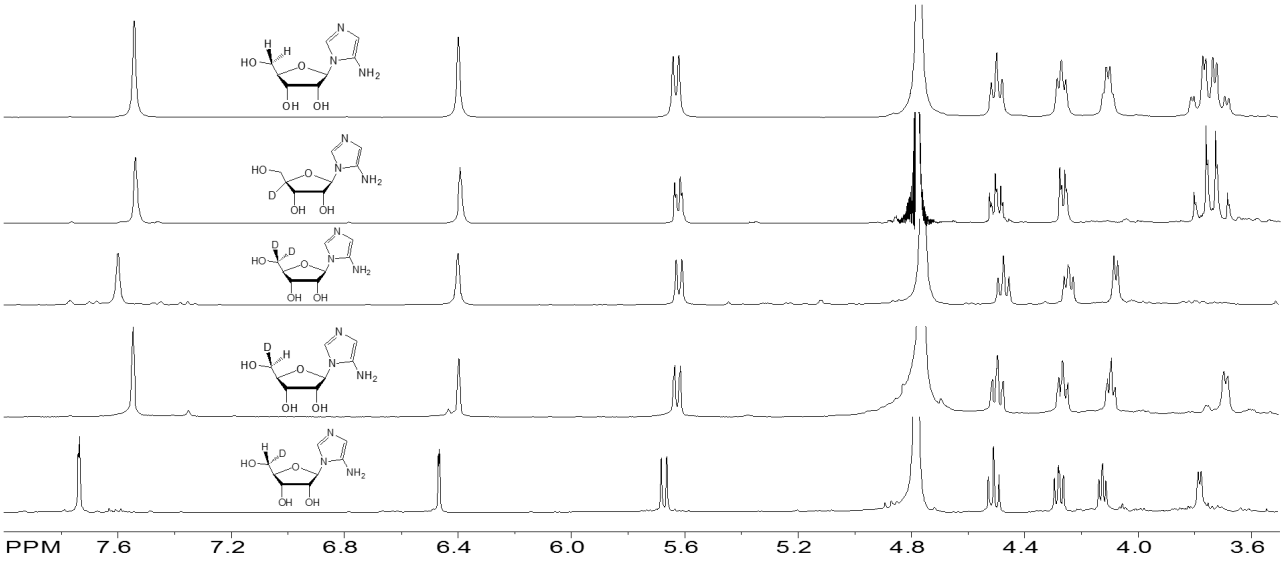


Figure S15: ^{1}H NMR (D_2O) for **53** (53a-R = H, R_1 = H, R_2 = H), (53b-R = D, R_1 = H, R_2 = H), (53c-R = H, R_1 = D, R_2 = D), (53d-R = H, R_1 = D, R_2 = H) & (53e-R = H, R_1 = H, R_2 = D) respectively.

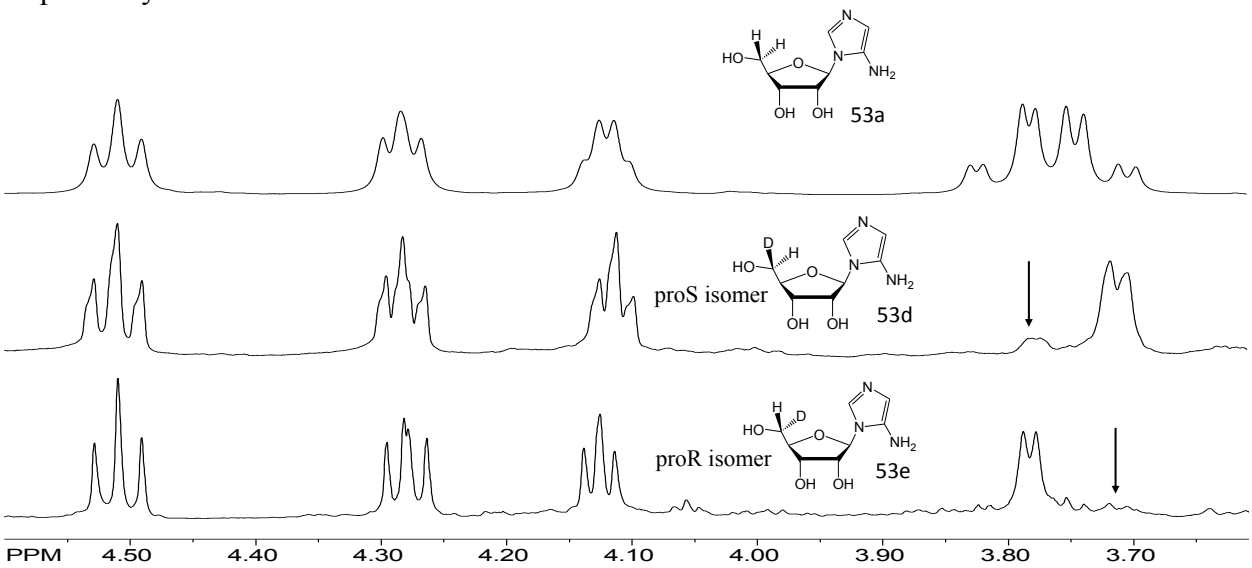


Figure S16: ^{1}H NMR (D_2O) range 3.6-4.6 ppm for **53** (53a-R_1 = H, R_2 = H), (53d-R_1 = D, R_2 = H) & (53e-R_1 = H, R_2 = D) respectively. Both isomers contain (10-15%) of the C5' epimer (indicated by arrows). This is most likely due to less than complete stereocontrol in the reduction of **38** (Figure S1).

(5S)-[5-$^{2}H_1$] 5-Amino-1-(β-D-ribofuranosyl)imidazole **53d**: ^{1}H NMR (D_2O) δ 7.6 (s, 1, H2), 6.4 (s, 1, H4), 5.6 (d, 1, H1′), 4.5 (t, 1, H2′), 4.3 (m, 1, H3′), 4.1 (d, 1, H4′), 3.7 (d, 1H5′).

(5R)-[5-$^{2}H_1$] 5-Amino-1-(β-D-ribofuranosyl)imidazole **53e**: ^{1}H NMR (D_2O) δ 7.7 (s, 1, H2), 6.4 (s, 1, H4), 5.6 (d, 1, H1′), 4.5 (t, 1, H2′), 4.3 (m, 1, H3′), 4.1 (d, 1, H4′), 3.8 (d, 1H5″).

The synthesis of 1′,2′,3′,4′,5′-^{13}C AIRs isotopomer (**53f**) was achieved as shown in Figure **S7** using UL-^{13}C-Ribose. ^{1}H NMR and ^{13}C NMR for [1′-^{13}C] 5-Amino-1-(β -D-ribofuranosyl)imidazole (**52g**), [2′-^{13}C] 5-Amino-1-(β -D-ribofuranosyl)imidazole (**53h**) and [3′-^{13}C] 5-Amino-1-(β -D-ribofuranosyl) imidazole (**53i**) were shown in our recent publications. We had previously synthesized these compounds (53f-i).[5,6] The synthesis of the labeled [Amino-^{15}N] AIRs isotopomer (**53j**) was achieved by using nitroimidazole [Nitro-N^{15}].[7]

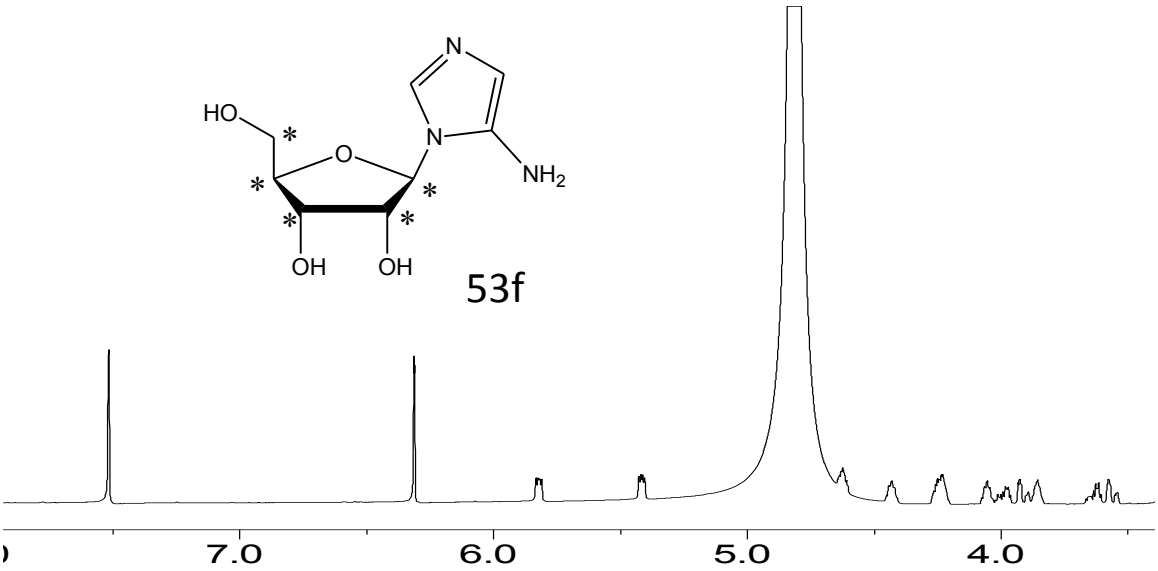


Figure S17: ^{1}H NMR (D_2O) for [UL-^{13}C] 5-Amino-1-(β-D-ribofuranosyl)imidazole (**53f**).

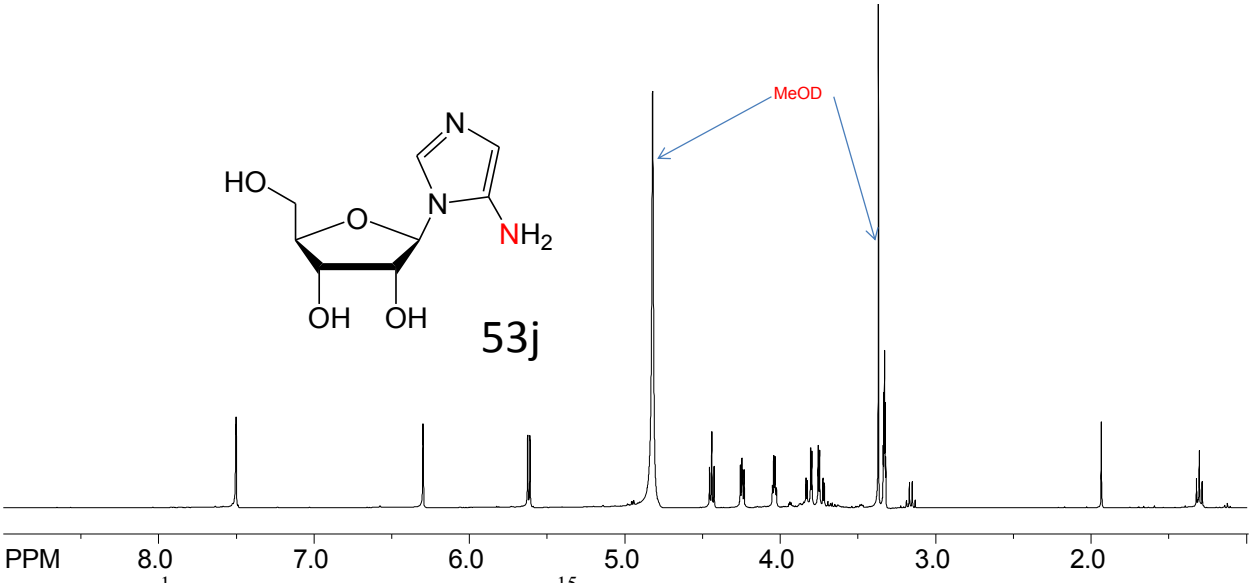


Figure S18: ^{1}H NMR (MeOD) for [Amino-^{15}N] 5-amino-1-(β-D-ribofuranosyl)imidazole (**53j**).

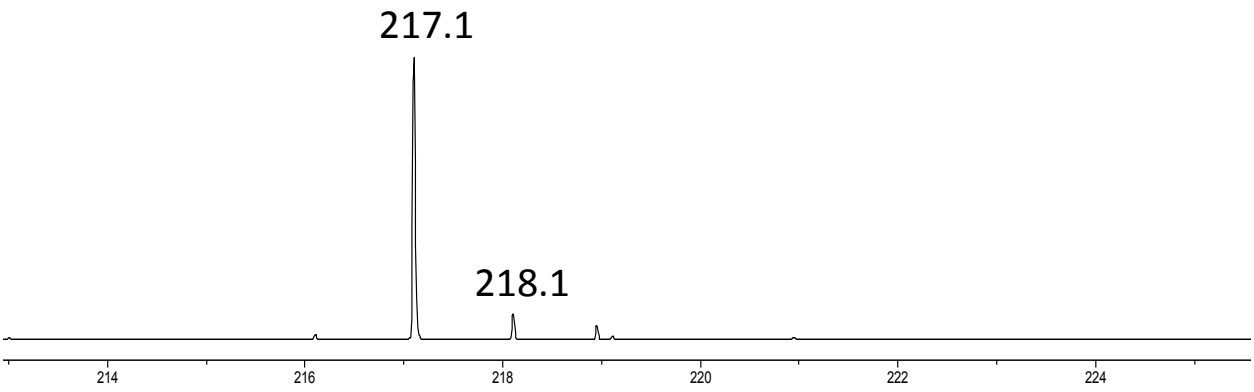


Figure S19: ESI (+ve mode) - Mass spectrum for [$^{15}NH_2$]-5-amino-1-(β-D-ribofuranosyl)imidazole (**53j**) showing complete isotopic substitution of the AIR amino group.

Over-expression and purification of HBI synthase:

The nucleotide sequence of the *Desulformonas acetoxidans* DSM 684ThiC homolog (HBI synthase) clustered with vitamin B_{12} biosynthetic genes was synthesized (see Figure **S33**) and cloned in THT vector (pET 28 vector that has a TEV protease cleavage site for removing the His tag). HBI synthase was co-expressed in the presence of a plasmid encoding the *suf* operon for *in vivo* [4Fe-4S] reconstitution in *E.coli* BL21 (DE3).[8,9] An overnight 15 mL culture was grown in LB medium in the presence of kanamycin (40 mg/L) and chloramphenicol (25 mg/L). This was then added to a 1.8 L minimal medium (M9 minimal salts 1X, 27 mL of 50% glucose, 7 mL of 1 M $MgSO_4$, 200 µl of 1 M $CaCl_2$, 72 mg kanamycin and 45 mg chloramphenicol were added to 1.8 L water). The cultures were incubated at 37 °C with shaking (180 rpm) until the OD_{600} reached 0.6 to 0.65. The cultures were then incubated at 4 °C without shaking for 3 hrs. Then 50 mg of ferrous ammonium sulfate and 50 mg of cysteine were added. This was followed by induction of the culture with 70 µM IPTG. The culture was then incubated at 15 °C with shaking (50 rpm) for 18 - 20hrs. The cultures were then incubated at 4 °C for 3 hrs without shaking. The cells were then harvested and stored in liquid nitrogen overnight before enzyme purification. For enzyme purification, the cell pellets were thawed at room temperature in an anaerobic chamber and suspended in lysis buffer (100 mM Tris-HCl, pH 7.5) in the presence of 2 mM DTT, lysozyme (0.2 mg/mL) and benzonase (100 units). This mixture was then cooled in

an ice-bath for 2 hrs. The suspension of cells was then sonicated and centrifuged to give the cell-free extract. The enzyme was then purified using standard Ni-NTA chromatography. The column was first incubated with the lysis buffer. The cell-free extract was then passed over the column, which was then washed with 8-9 column volumes of wash buffer (100 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 2mM DTT, pH 7.5). The enzyme was then eluted using 100 mM Tris-HCl, 300 mM NaCl, 250 mM imidazole, 2 mM DTT, pH 7.5. The purified enzyme was buffer exchanged into 100 mM potassium phosphate, 30% glycerol, 2 mM DTT, pH 7.5 using an Econo-Pac 10DG desalting columns (commercially available from *Bio-RAD*) and the purified enzyme was stored in liquid nitrogen. The protein as isolated contains a [4Fe-4S] cluster as indicated by the UV-Vis spectrum shown in Figure 4.

Enzymatic assays for LCMS and HPLC purification of 5-HBI:

HBI synthase (120 μM) was incubated with dithionite (10 mM), AIR (5 mM), SAM (7 mM) at room temperature and anaerobic conditions for 60 min. Controls were set up where either of AIR, SAM, dithionite or HBI synthase were absent. The protein was filtered using 10 kDa cut off filters and the small molecule pool was analyzed by HPLC and LCMS. These assays were done with unlabeled AIR, 5′,5′-$^{2}H_2$-AIR, 4-$^{2}H_2$-AIR, 1′,2′,3′,4′,5′-$^{13}C_5$-AIR, 5′-R-^{2}H-AIR, 5′-S-^{2}H-AIR, 1′-^{13}C-AIR, 2′-^{13}C-AIR, 3′-^{13}C-AIR. The same assay conditions were used for the purification of 5-HBI for NMR characterization. These assays were done with unlabeled AIR, 5′,5-$^{2}H_2$-AIR, 4-$^{2}H_2$-AIR, 2'-^{13}C-AIR.

Enzymatic assays for NMR analysis to determine the fate of 1'-C of AIR:

HBI synthase (800 μM) was incubated with dithionite (50 mM), 1′,2′,3′,4′,5′-$^{13}C_5$-AIR (20 mM) and SAM (30 mM) at room temperature for 60 min. Controls were set up in which either AIR, SAM, dithionite or HBI synthase were absent. The protein was filtered using 10 kDa cut off filters. The small molecule pool was analyzed by ^{13}C and DEPT 90 NMR experiments.

HPLC conditions for co-elution experiment and purification of 5-HBI:

HPLC-1260 series Agilent. LC-18 column - ((Supelcosil LC-18-T column (15 cm x 3 mm, 3 μm)). Following gradient was used: A: water, B: 5 mM ammonium formate, C: methanol. Following gradient was used - 0 min: 100%B; 7 min: 100%B; 25 min: 7%A, 70%B, 23%C; 27 min: 25%A, 75%C; 29 min - 25%A, 75%C; 30 min: 100%B; 45 min – 100%B).

HPLC conditions for LCMS experiments:

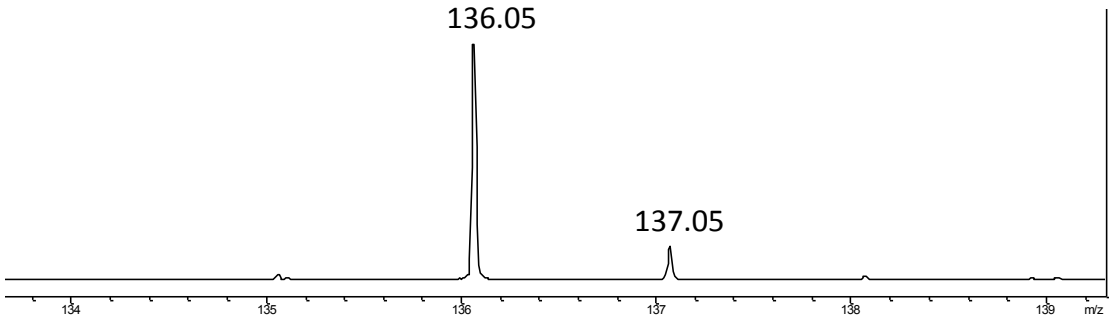
HPLC - 1260 series Agilent. LC-18 column - ((Supelcosil LC-18-T column (15 cm x 3 mm, 3 μm)). Following gradient was used - A: 5 mM ammonium formate, B: 75% methanol and 25% water. Following gradient was used - 0 min: 100% A; 7 min: 100% A; 25 min: 70% A, 30% B; 26 min: 100% B; 28 min – 100% B; 29 min: 100% A; 40 min – 100%A).

MS conditions for LCMS experiments:

Capillary, -4500 V; Capillary offset, -500 V; Nebulizer gas, 3.0 bar; Dry gas, 10.0 L/min; Dry gas temperature, 200^{o}C; Funnel 1 RF, 200.0 Vpp; Funnel 2 RF, 200.0 Vpp; ISCID, 0.0 eV; Hexapole RF, 200 Vpp; Quadrapole, Ion energy, 5.0 eV; Low mass, 200 *m/z*; Collision cell, collision energy, 10.0 eV; Collision RF, 150.0 Vpp, Transfer time, 100.0 μs; Prepulse storage, 5.0 μs. Data was processed with DataAnalysis ver. 4.0 SP4 (Bruker Daltonics, Billerica, MA).

LCMS results for HBI synthase reactions with labeled AIR isotopomers:

MS for 5-HBI:



MS for 5'-dA:

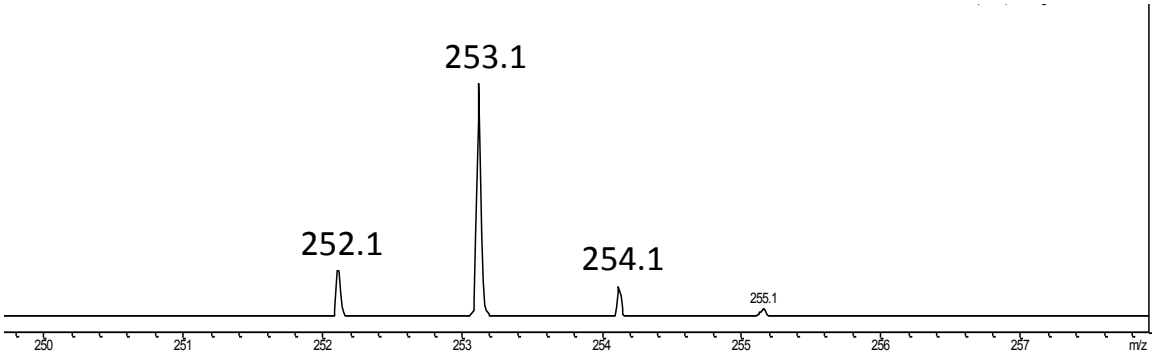
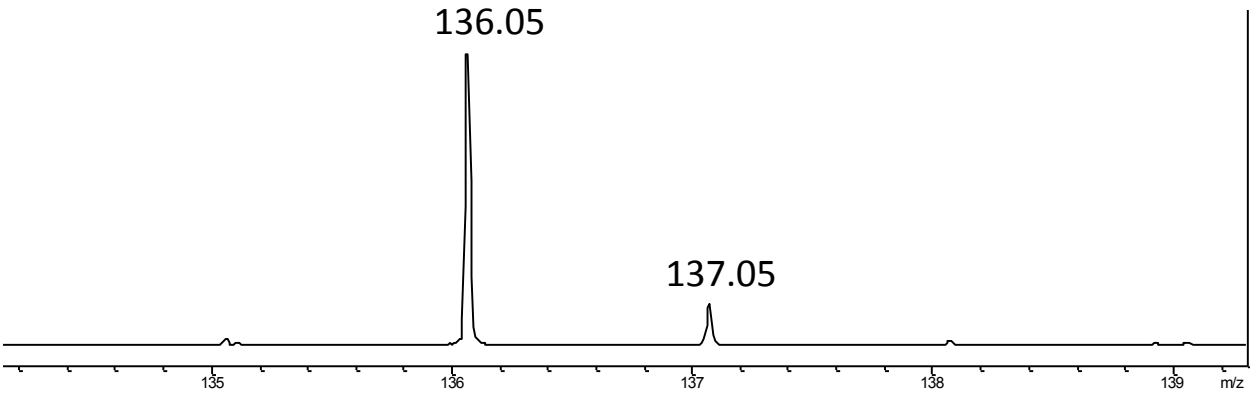


Figure S20: LCMS data for 5-HBI and 5′-dA formed in the HBI synthase reaction with 5′,5′-$^{2}H_2$-AIR.

MS for 5-HBI:



MS for 5'-dA:

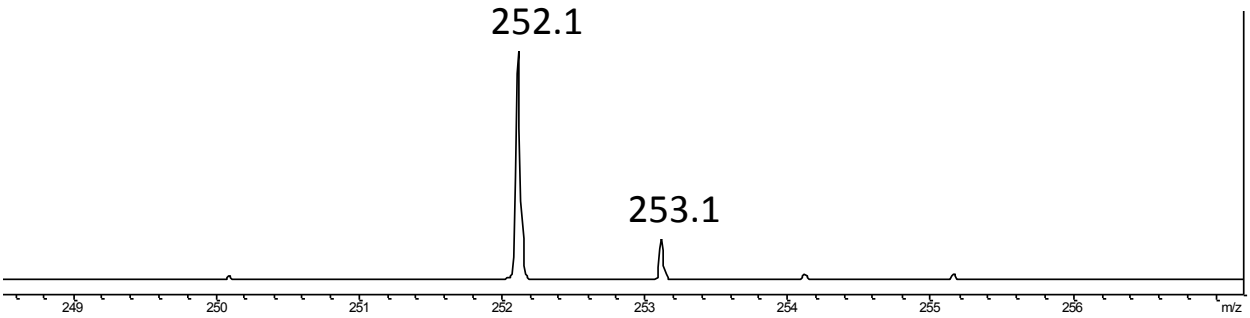
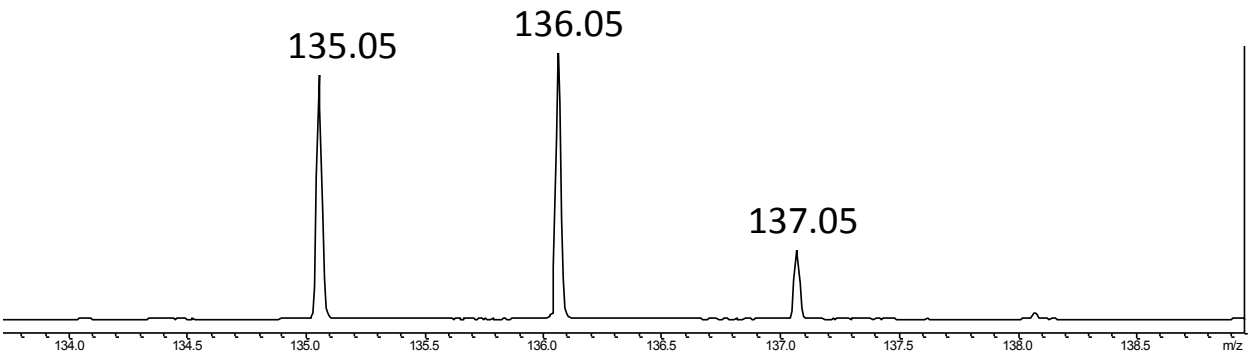


Figure S21: LCMS data for 5-HBI and 5′-dA formed in the HBI synthase reaction with 5′-R-^{2}H-AIR.

MS for 5-HBI



MS for 5'-dA

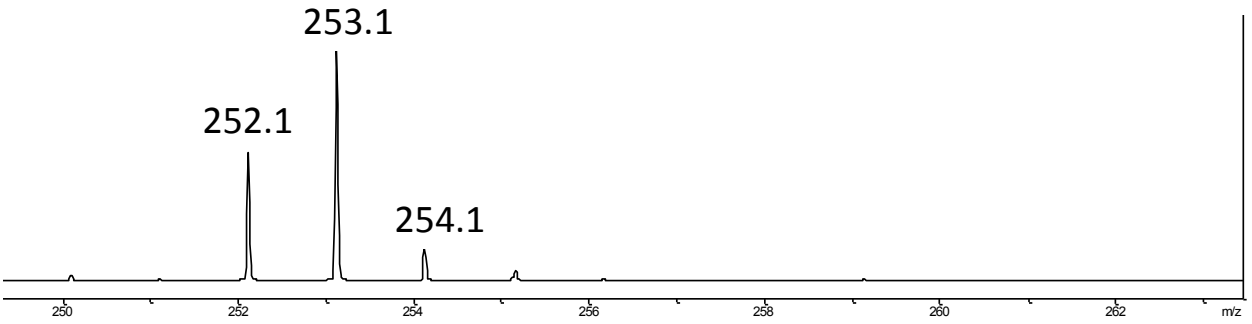
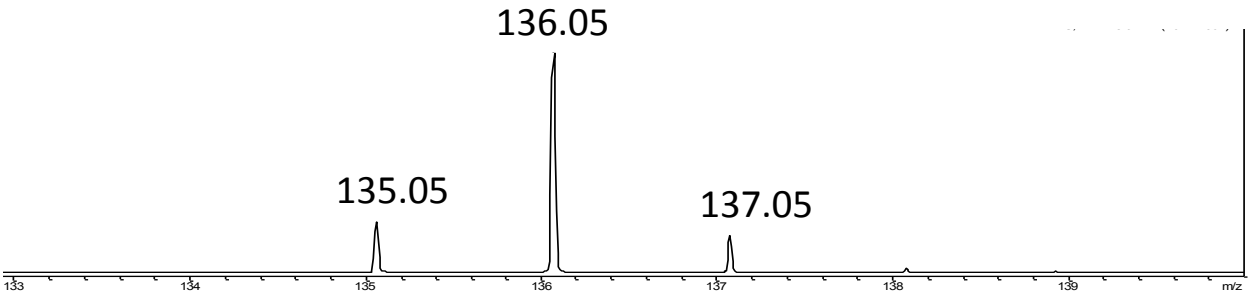


Figure S22: LCMS data for 5-HBI and 5′-dA formed in the HBI synthase reaction with 5′-S-^{2}H-AIR.

MS for 5-HBI:



MS for 5′-dA:

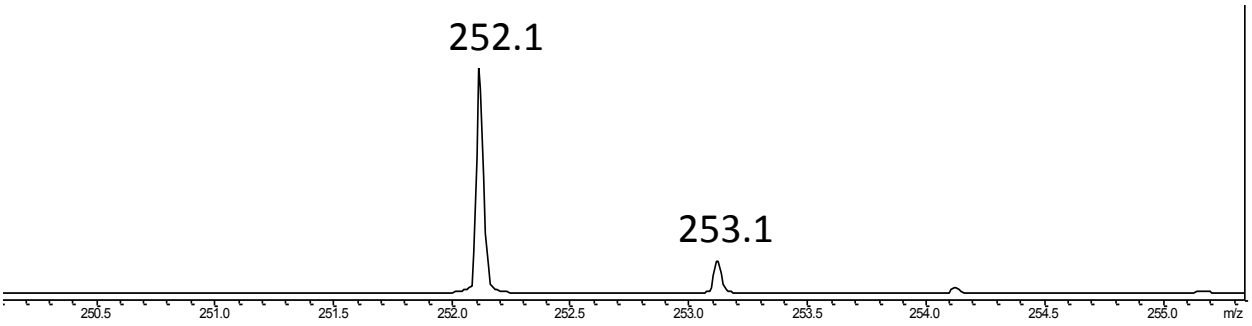


Figure S23: LCMS data for 5-HBI and 5′-dA formed in the HBI synthase reaction with 4′-^{2}H-AIR.

MS for 5-HBI:

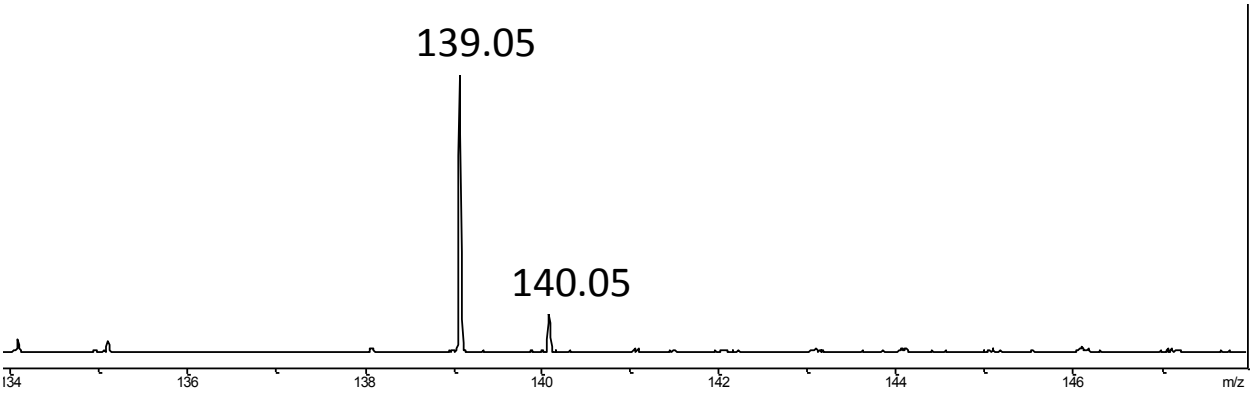


Figure S24: LCMS data for 5-HBI formed in the HBI synthase reaction with 1′,2′,3′,4′,5′-$^{13}C_5$-AIR.

MS for 5-HBI:

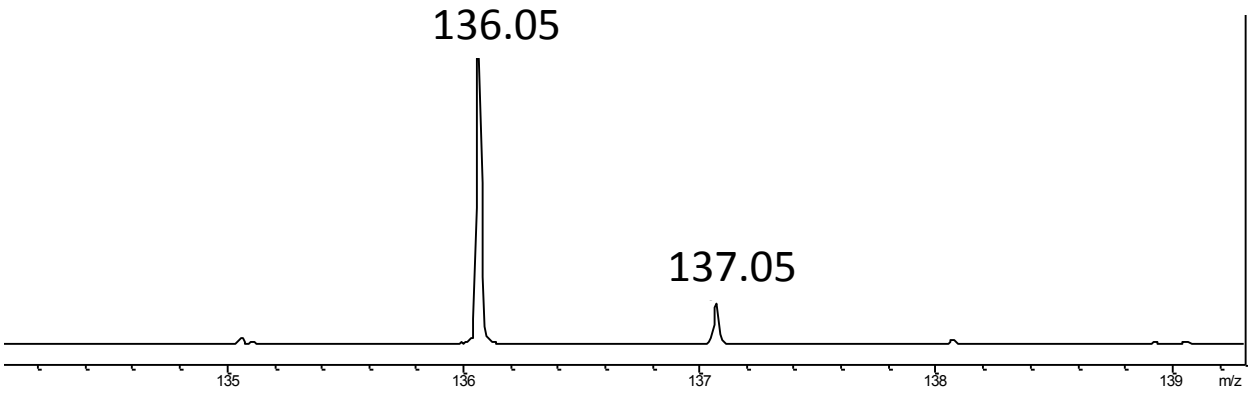


Figure S25: LCMS data for 5-HBI formed in the HBI synthase reaction with 3′-^{13}C-AIR.

MS for 5-HBI:

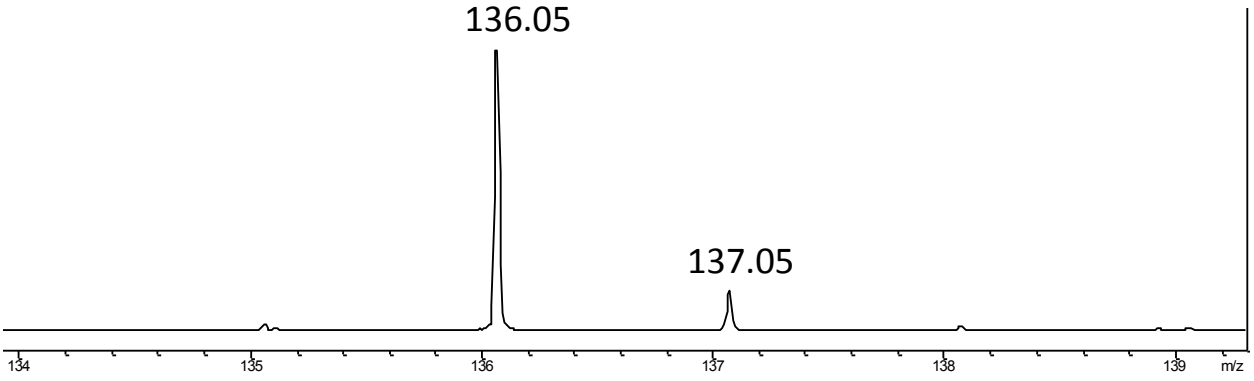


Figure S26: LCMS data for 5-HBI formed in the HBI synthase reaction with 2′-^{13}C-AIR.

MS for 5-HBI:

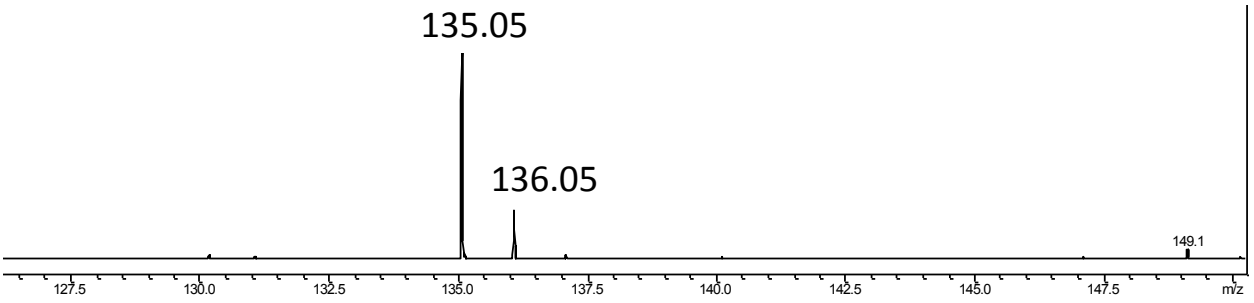


Figure S27: LCMS data for 5-HBI formed in the HBI synthase reaction with 1'-^{13}C-AIR.

MS for 5-HBI:

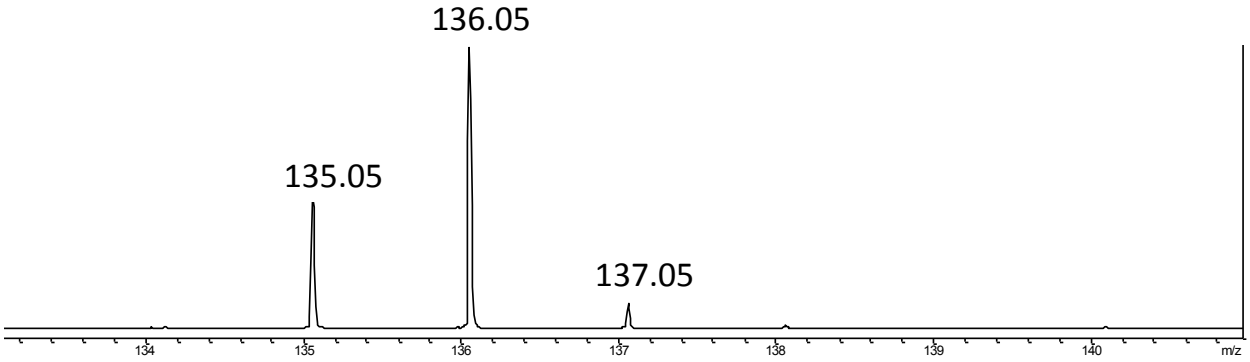


Figure S28: LCMS data for 5-HBI formed in the HBI synthase reaction with [$^{15}NH_2$]-AIR.

MS for 5'-dA:

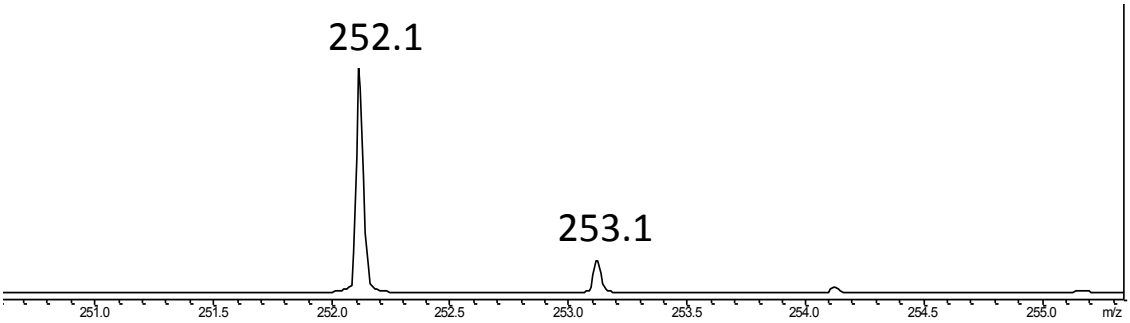


Figure S29: LCMS data for 5′-dA formed in the HBI synthase reaction with AIR.

Table S1: MS analysis of the conversion of AIR isotopomers to HBI.

	Substrate	5-HB $[M+H]^+$ Da	5'dA $[M+H]^+$ Da
1	Unlabeled AIR	135.0	252.0
2	5',5'-$^2H_2$-AIR	136.0	253.0
3	5'-R-^{2}H-AIR	136.0	252.0
4	5'-S-^{2}H-AIR*	135.0/136.0 =1	252.0/253.0 = 2.7
5	4'-^{2}H-AIR	135.0	252.0
6	1',2',3',4',5'-$^{13}C_5$-AIR	139.0	252.0
7	3'-^{13}C-AIR	136.0	252.0
8	2'-^{13}C-AIR	136.0	252.0
9	1'-^{13}C-AIR	135.0	252.0
10	^{15}N-AIR (C5-NH_2)	135.0/136.0 = 0.62	252.0

*The 5′-S-^{2}H-AIR contained approx. 10% 5′-R-^{2}H-AIR (see Figures S1 and S16). The relatively large amount of deuterium retained in HBI is most likely due to a V/K isotope effect on the hydrogen atom abstraction by the 5′-deoxyadenosyl radical.

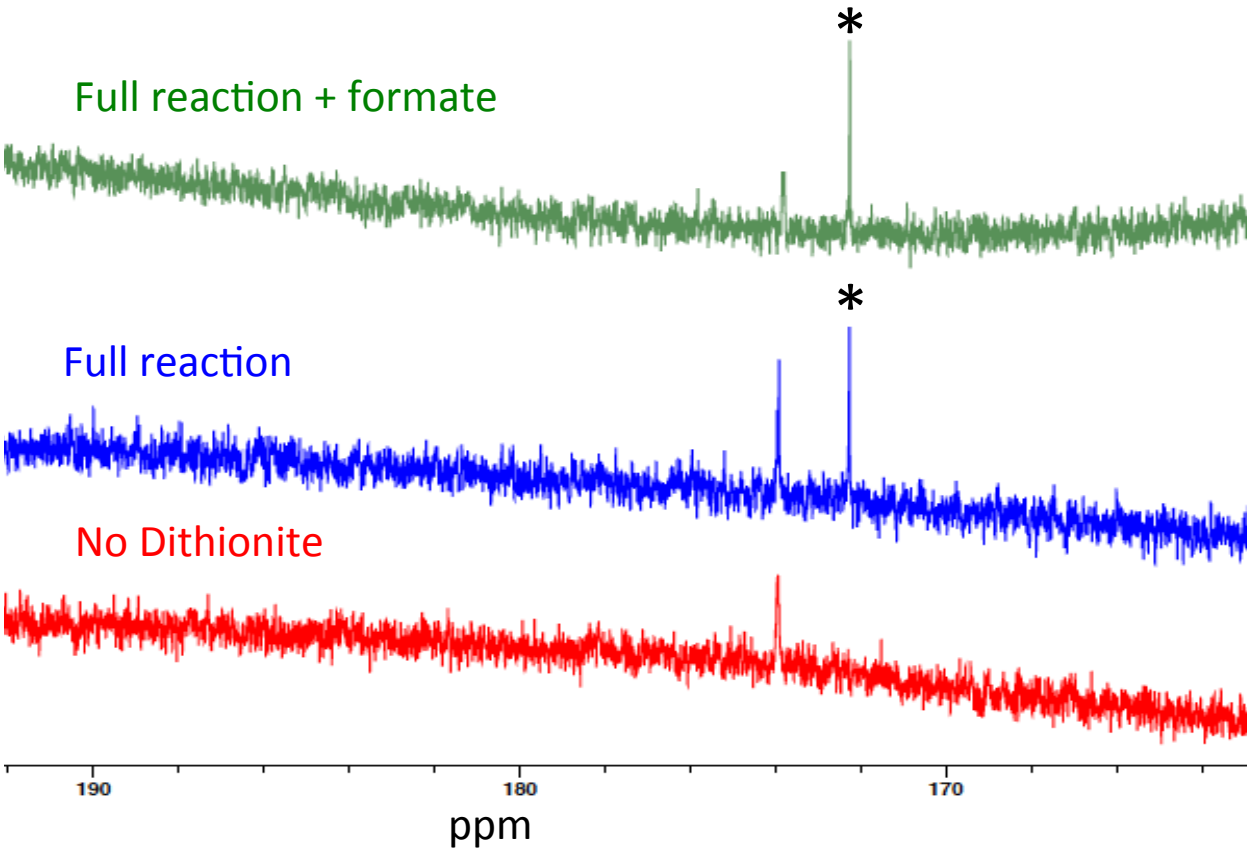


Figure S30: ^{13}C NMR of the small molecule pool from HBI synthase reactions with 1′,2′,3′,4′,5′-$^{13}C_5$-AIR. Red – No Dithionite control. Blue – Full reaction. Green – Full reaction + 50 mM formate. The asterisk indicates the formate signal.

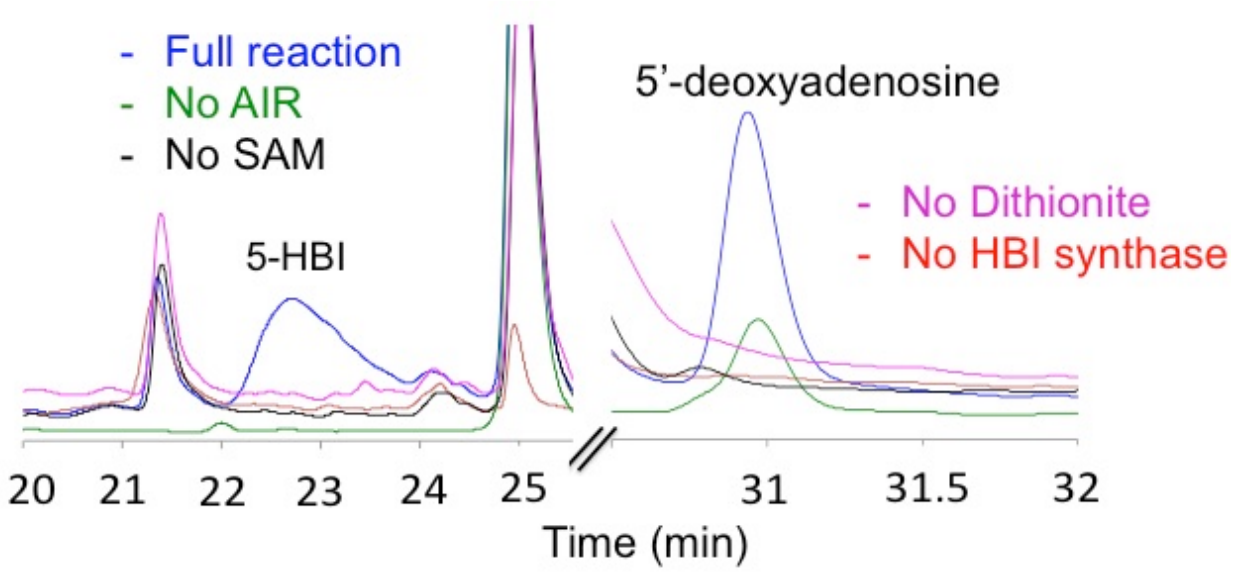


Figure S31: HPLC (290 nm) analysis of the HBI synthase reaction mixture showing the product eluting at 22.7 min (blue trace). This product was not formed in reaction mixtures lacking AIR, SAM, dithionite or HBI synthase. 5′-deoxyadenosine elutes at 31 min.

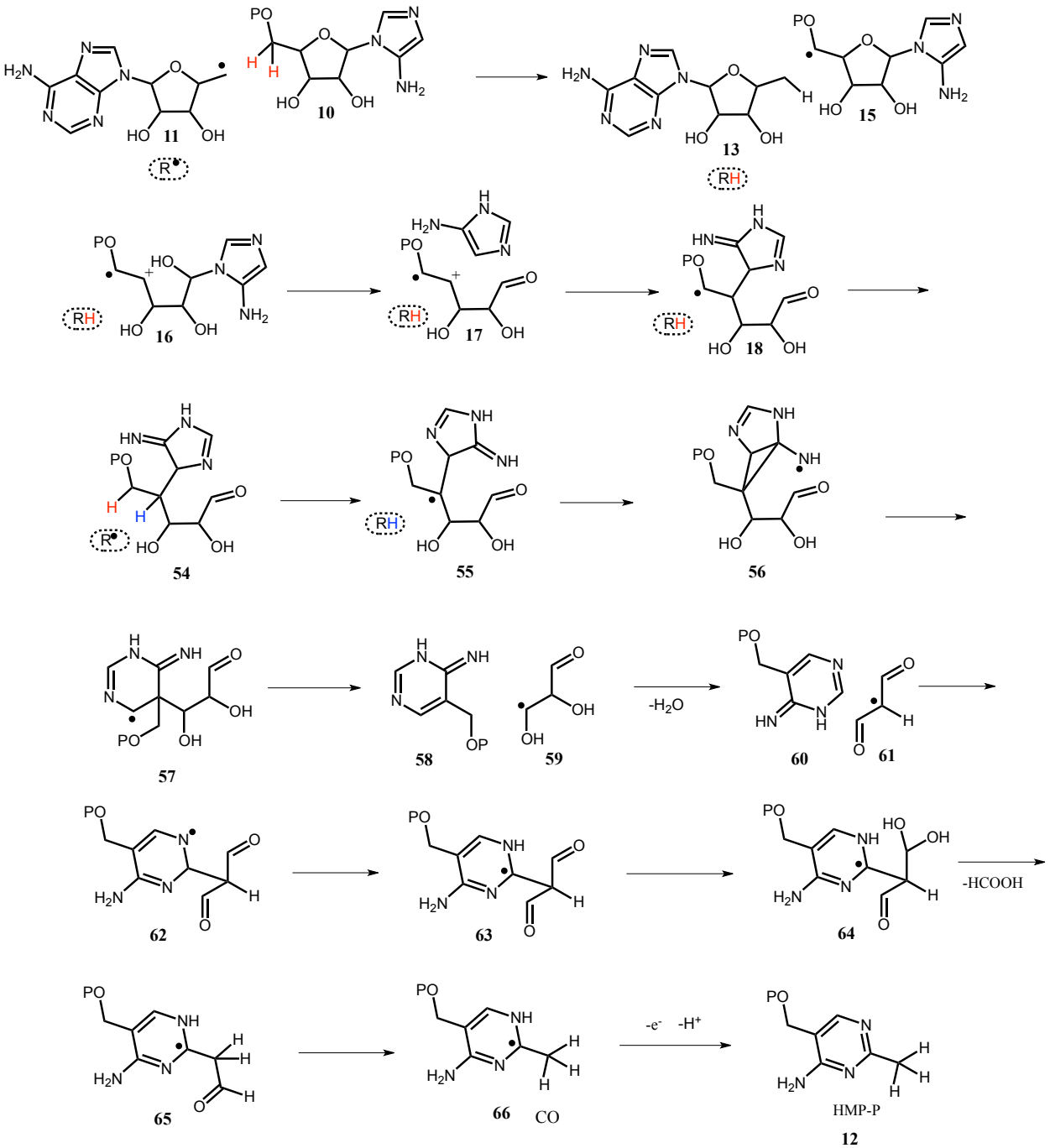


Figure S32: Proposed mechanism for the ThiC-catalyzed conversion of AIR **10** to HMP-P **12** involved in thiamin biosynthesis.[3]

(A)

http://theseed.uchicago.edu/FIG/subsys.cgi?user=master:thiamine&ssa_name=Thiamin_biosynthesis&request=show_ssa&can_alter=&check=&sort=by_variant&show_clusters=&show_minus1=

(B)

Genome ID	Organism	Variant Code	CDS(ThiI)	ThiI	Rhod(ThiI)	ThiS	ThiF	IscS	Dxs	ThiH	ThiO	ThiG	Thi4	ThiC	ThiD	Thi5	*Thi_S	ThiM
207559.3	Desulfovibrio desulfuricans G20 [B]	1					961		2156	960		959		1473	3012		962-17, 1751-17	1750
882.1	Desulfovibrio vulgaris subsp. vulgaris str. Hildenborough [B]	1				2083	2080		1344	2081		2082		1405	927		2079-17, 2350-17	2351
891.1	Desulfuromonas acetoxidans [B]	1				4763	3880, 3308		1608, 1492	3879, 3309		1426, 544		557, 2189, 704	556, 703		556-17, 545-17, 703-17, 1425-17	2051, 1944

(C)

http://theseed.uchicago.edu/FIG/seedviewer.cgi?page=Annotation&feature=fig|891.1.peg.2189&user=master:thiamine

(D)

ATGAAAACGCAAGTTGAACACGCAGTGGACGGTA
TTATTACGGAACAAATGGCGACGGTGGCGCACGACGAAGACCTGA
GCCCGGAATATATTCGCACCATGGTCGCGGAAGGTAAAATTGTGA
TCCCGAACAATTCAAACTCGACGCCGAAACCGGTCGGCATCGGTA
AAGGCCTGCGTACCAAAGTGAACGCCTCAATTGGTACGAGCTCTG
ATATCGTTAATTACCAGGCAGAAGTCCGTAAAGCACGCATTGCTG
AACAAGCAGGCGCTGACACCCTGATGGAACTGTCGGTTGGCGGTA
ATCTGGATCGTGTGCGTCGCGAAGTTCTGGCAGCAGTCAACCTGC
CGGTTGGTAATGTCCCGCTGTATCAGGCATTTTGCGATGCTACGC
GTAAATACGGCTCCGCAGATAAACTGGACCCGGAAGAACTGTTTG
ACCTGATTGAACAGCAATGCGAAGATGGTCTGGCGTTCATGGCCA
TTCATTGTGGCATCAACCGCTATACCATCGAACGTCTGCGCAAAC
AGCACTATCGTTACGGCGGTCTGGTTAGCAAAGGCGGTACGAGCA
TGGTTTCTTGGATGGAACATAACAATCGCGAAAATCCGCTGTATG
AACAATTCGACCGTGTGGTTGCCATTCTGAAAAAATACGATGTGT
GTCTGAGCCTGGGTAACGGTCTGCGTGCAGGTGCAATCCATGATT
CTCACGACCGTGCGCAGATGCAAGAACTGATTATCAATTGCGAAC
TGGCCCAGCTGGGTCGCGAAATGGGCTGTCAAATGCTGGTTGAAG
GTCCGGGCCACATGCCGCTGGATGAAGTCGAAGCAAACATTCTGA
TCCAGAAACGTATGTCAAATGAAGCTCCGTATTACATGCTGGGTC
CGATTTCGACCGATGTCGTGCCGGGCTTTGACCATATTAGTTCCG
CGATCGGTGCAGCTCAGAGTGCGCGCTATGGCGCCGATCTGATTT
GCTACATCACCCCGGCAGAACACCTGGCTCTGCCGAACGAAGATG
ACGTGCGCTCCGGTGTTGAAGCGGCCCGTGTTGCCACGTATATCG

GTGATATGAATAAATACCCGGACAAAGGCCGTCAGCGCGATAAAG
CGATGAGCAAAGCCCGTCGCGATCTGCAGTGGGACAAACAATTTG
AACTGGCACTGATGCCGGAACAGGCTCGTCAAGTGCGCGATAGTC
GTCTGCCGGAAGAAGAACATTCCTGCACCATGTGTGGCAACTTCT
GTGCTGCTAACGGTTCTAAAACGCTGTTCGACGGCGACCTGCAAG
GCGATAAATGCTGACTCGAG

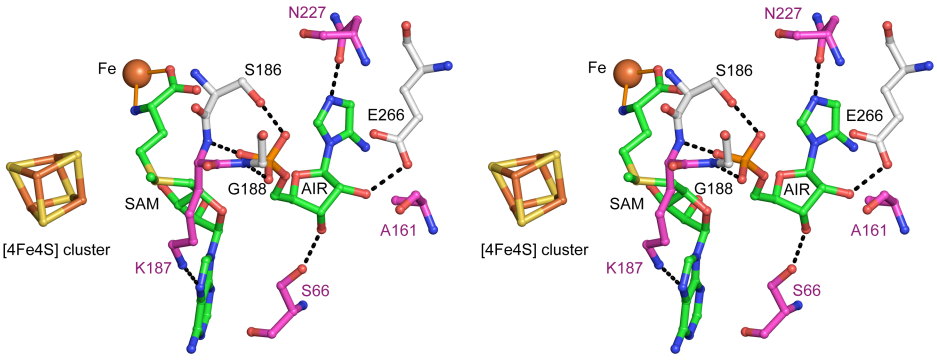
Figure S33: (A) Seed database link to identify organisms with multiple copies of ThiC. (B) The *Desulfuromonas acetoxidans* [B] *bzaF* (annotated as 2189) was gene synthesized and codon optimized. (C) Seed database link showing the gene neighborhood for *D.acetoxidans* [B] *bzaF* (annotated as 2189). (D) Codon optimized sequence of *D.acetoxidans* [B] *bzaF*.

This gene was cloned in the THT vector (a modified pET 28b vector with a TEV cleavage site[10]).

Homology modeling procedure

The homology model was generated using the SWISS-MODEL online server,[11] providing the crystal structure of the AtThiC/AIR/SAH/Fe complex as a template. The model was then inspected and manually adjusted as needed using COOT.[12]

A



B

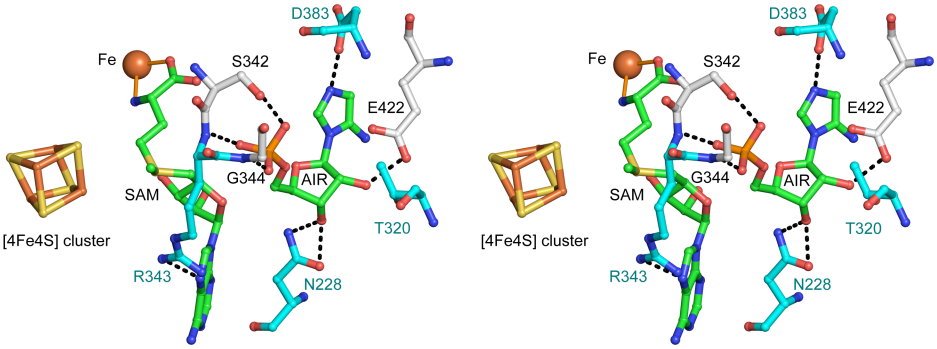


Figure S34: (A) Homology model for HBI synthase. (B) X-ray structure of the ThiC active site.

References:

(1) Loiseleur, O.; Ritson, D.; Nina, M.; Crowley, P.; Wagner, T.; Hanessian, S. *J. Org. Chem.* **2007**, *72*, 6353.

(2) Hanes, J. W.; Burns, K. E.; Hilmey, D. G.; Chatterjee, A.; Dorrestein, P. C.; Begley, T. P. *J. Am. Chem. Soc.* **2008**, *130*, 3043.

(3) Ono, A.; Ono, A.; Kainosho, M. *Tetrahedron Lett.* **1997**, *38*, 395.

(4) Tate, S.-i.; Kubo, Y.; Ono, A.; Kainosho, M. *J. Am. Chem. Soc.* **1995**, *117*, 7277.

(5) Chatterjee, A.; Hazra, A. B.; Abdelwahed, S.; Hilmey, D. G.; Begley, T. P. *Angew. Chem., Int. Ed.* **2010**, *49*, 8653.

(6) Lawhorn, B. G.; Mehl, R. A.; Begley, T. P. *Org. Biomol. Chem.* **2004**, *2*, 2538.

(7) Saikia, A.; Sivabalan, R.; Gore, G. M.; Sikder, A. K. *J. Sci. Ind. Res.* **2014**, *73*, 485.

(8) Hänzelmann, P.; Hernández, H. L.; Menzel, C.; García-Serres, R.; Huynh, B. H.; Johnson, M. K.; Mendel, R. R.; Schindelin, H. *Journal of Biological Chemistry* **2004**, *279*, 34721.

(9) Mehta, A. P.; Abdelwahed, S. H.; Begley, T. P. *J. Am. Chem. Soc.* **2013**, *135*, 10883.

(10) Fenwick, M. K.; Mehta, A. P.; Zhang, Y.; Abdelwahed, S. H.; Begley, T. P.; Ealick, S. E. *Nat Commun* **2015**, *6,* 6480.

(11) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. *Nucleic Acids Res* **2003**, *31*, 3381.

(12) Emsley, P.; Cowtan, K. *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2126.