Supporting information for:

Biosynthesis of the Thiamin-Thiazole in Eukaryotes: Identification of a Thiazole Tautomer Intermediate

*Abhishek Chatterjee[1], Frank C. Schroeder[2], Christopher T. Jurgenson[1], Steven E. Ealick[1] and Tadhg P. Begley[1]**

[1]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853

[2] Boyce Thompson Institute for Plant Research, Cornell University, Ithaca, NY 14853

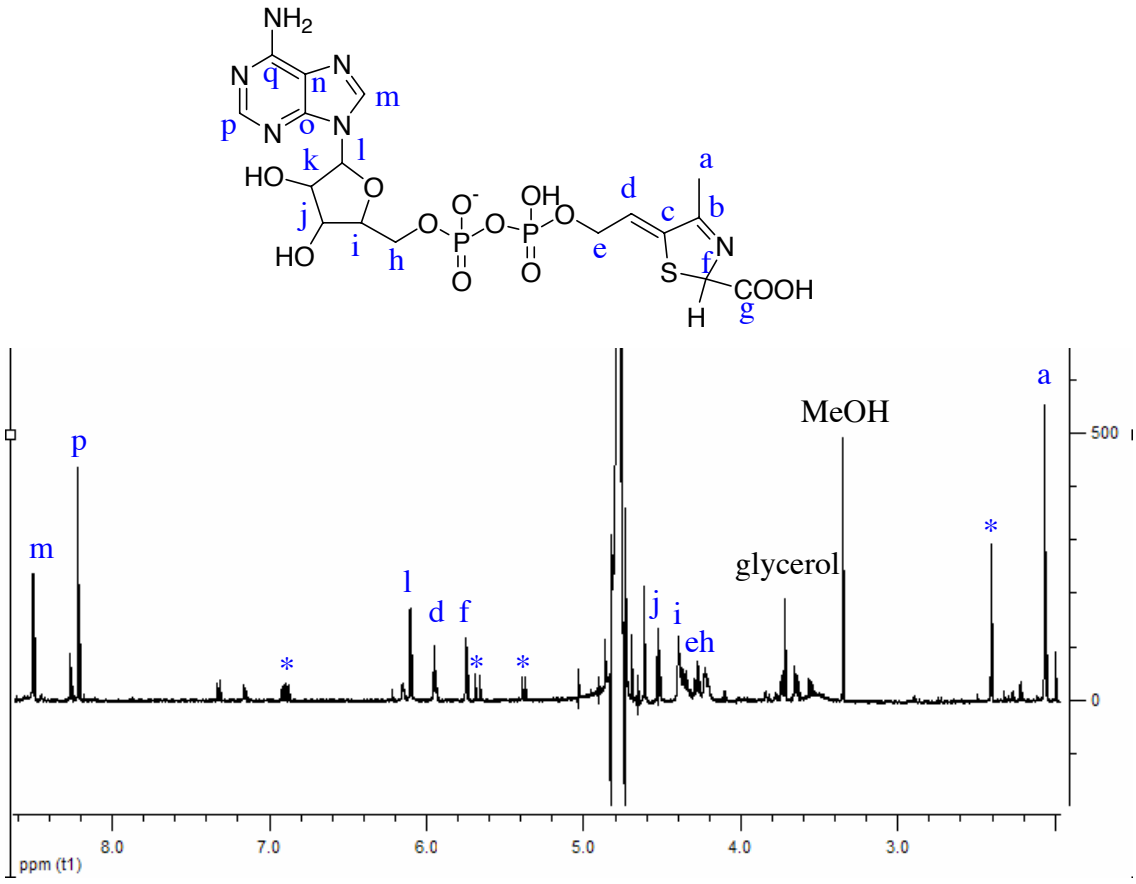
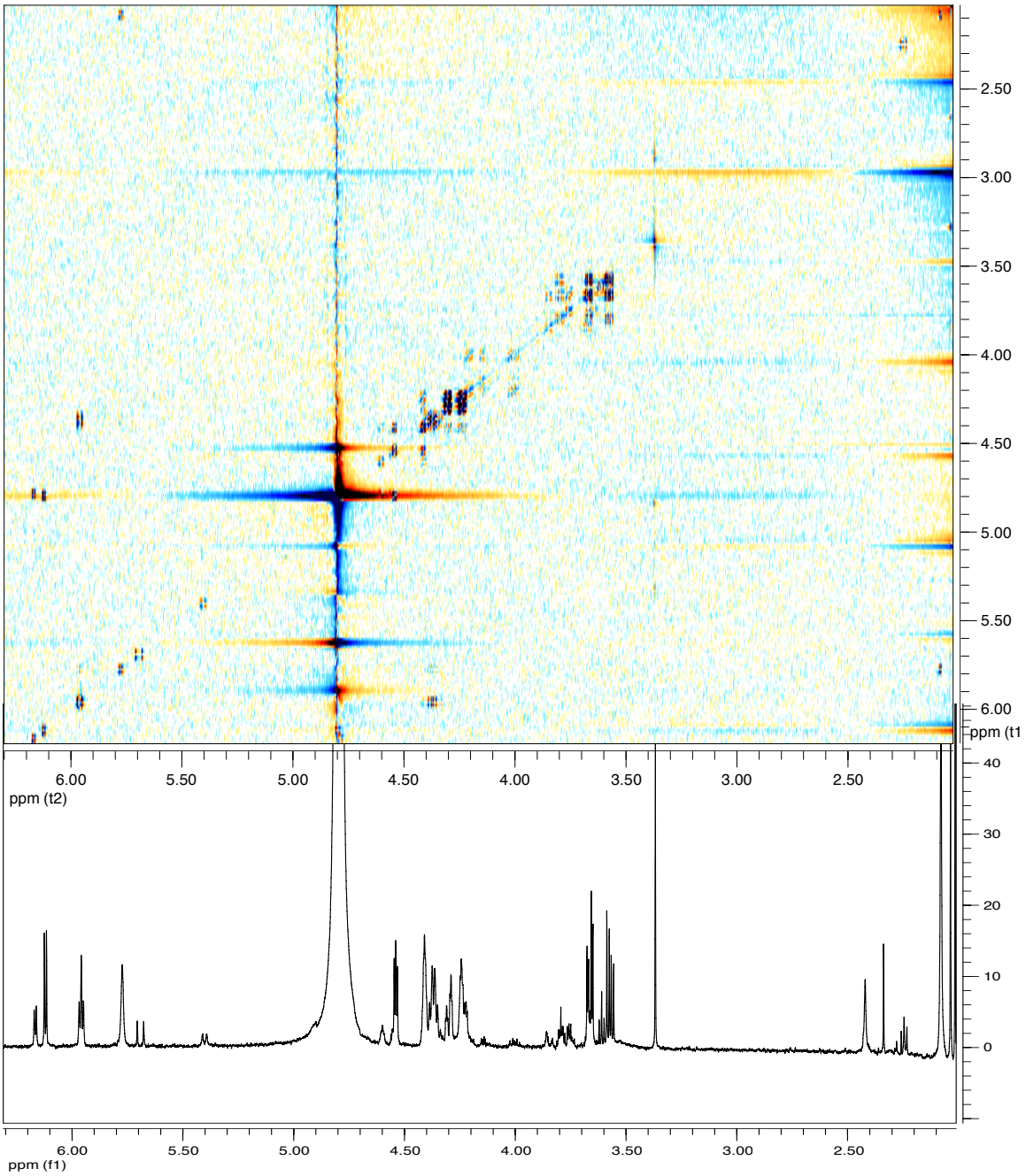


Figure SI 1: A proton NMR of the purified **13**. The assignments were done based on COSY (vide infra). The major impurities included methanol, glycerol, and degradation products (ADP and **16**). The proton signals correlating to **16** are marked with (*).


2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ppm (t1
6.00
5.50
5.00
4.50
4.00
3.50
3.00
2.50
ppm (t2)
40
30
20
10
0
6.00
5.50
5.00
4.50
4.00
3.50
3.00
2.50
ppm (f1)

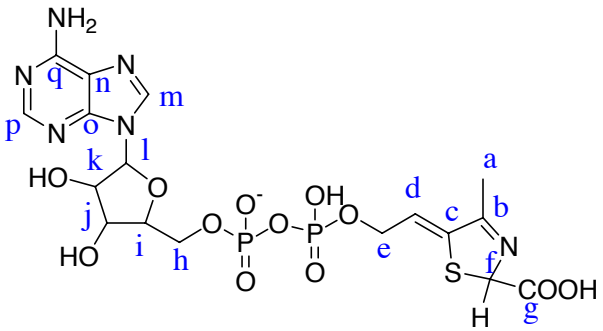

NH2
N
q
N
n
m
p
N
o
N
k
l
HO
O
j
HO
i
h
O
O-
P
O
O
OH
P
O
O
e
d
c
a
b
N
S
f
COOH
H
g

Figure SI 2: The COSY correlation diagram overlapped with the proton spectrum for direct comparison. The spin system corresponding to an ADP- moiety in the metabolite is easily identified. It reveals the proton **k** , which is hidden under the water peak. Also a minor ribose-spin system, very similar to the ADP-moiety of the intermediate, was found. Comparing it with standards and also from HPLC evidences this was found to be due to ADP, one of the degradation product of the purified intermediate. The methyl unit {**a**} (2.08 ppm) and the CH unit {**f**} at 5.78 ppm was found to couple to each other. Also a mutually strongly coupled CH {**d**} (5.95 ppm) and CH_2 {**e**} (4.37 ppm) units were observed. The multiplicity of the CH_2 {**e**} proton signal suggested that it is connected to the phosphate.

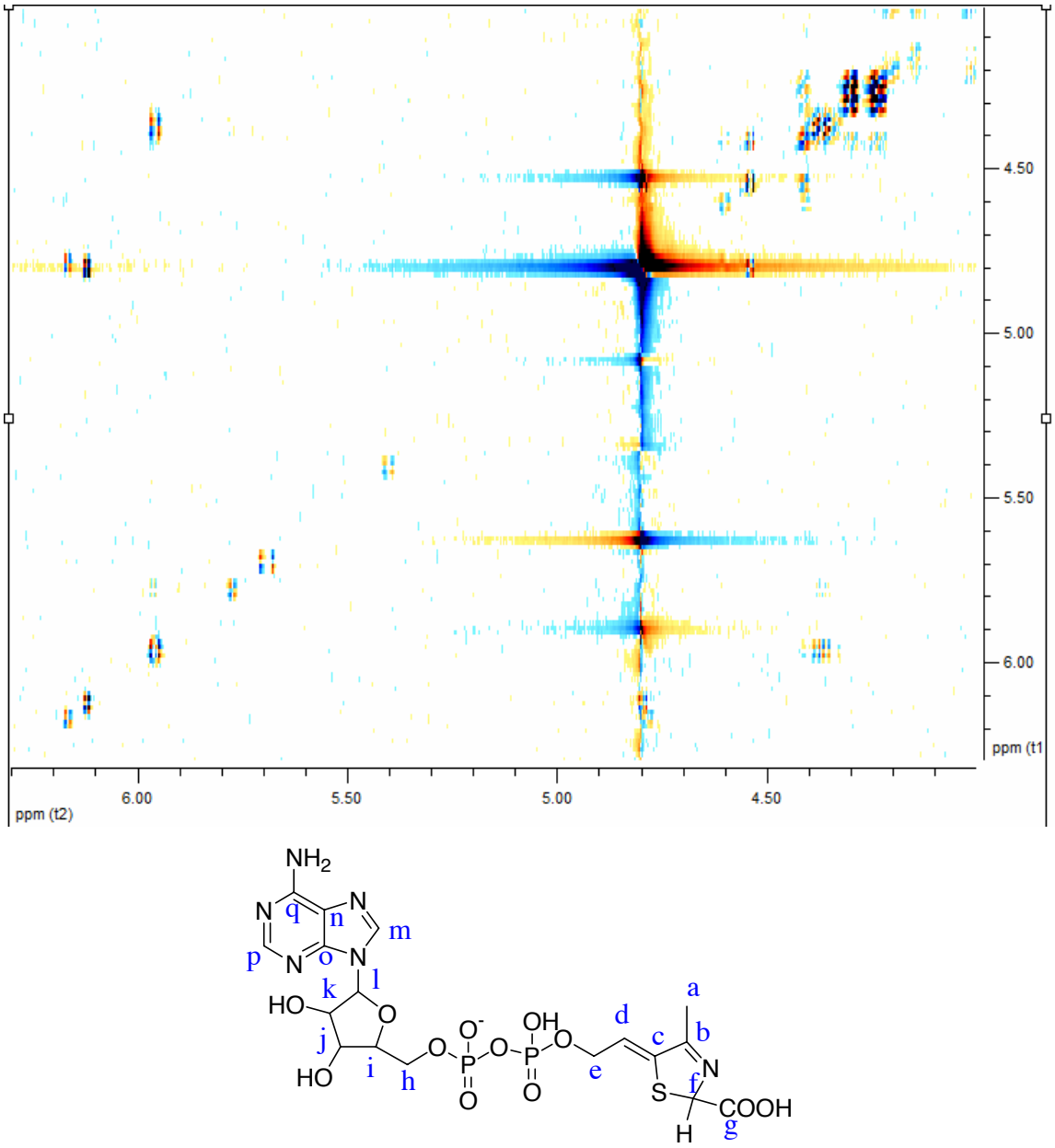


Figure SI 3: Magnified region of the dqfCOSY correlation diagram. The CH unit {**f**} at 5.78 ppm shows very weak correlation with both the CH_2 {**e**} and the CH {**d**} protons of the ADP-CH_2-CH unit.

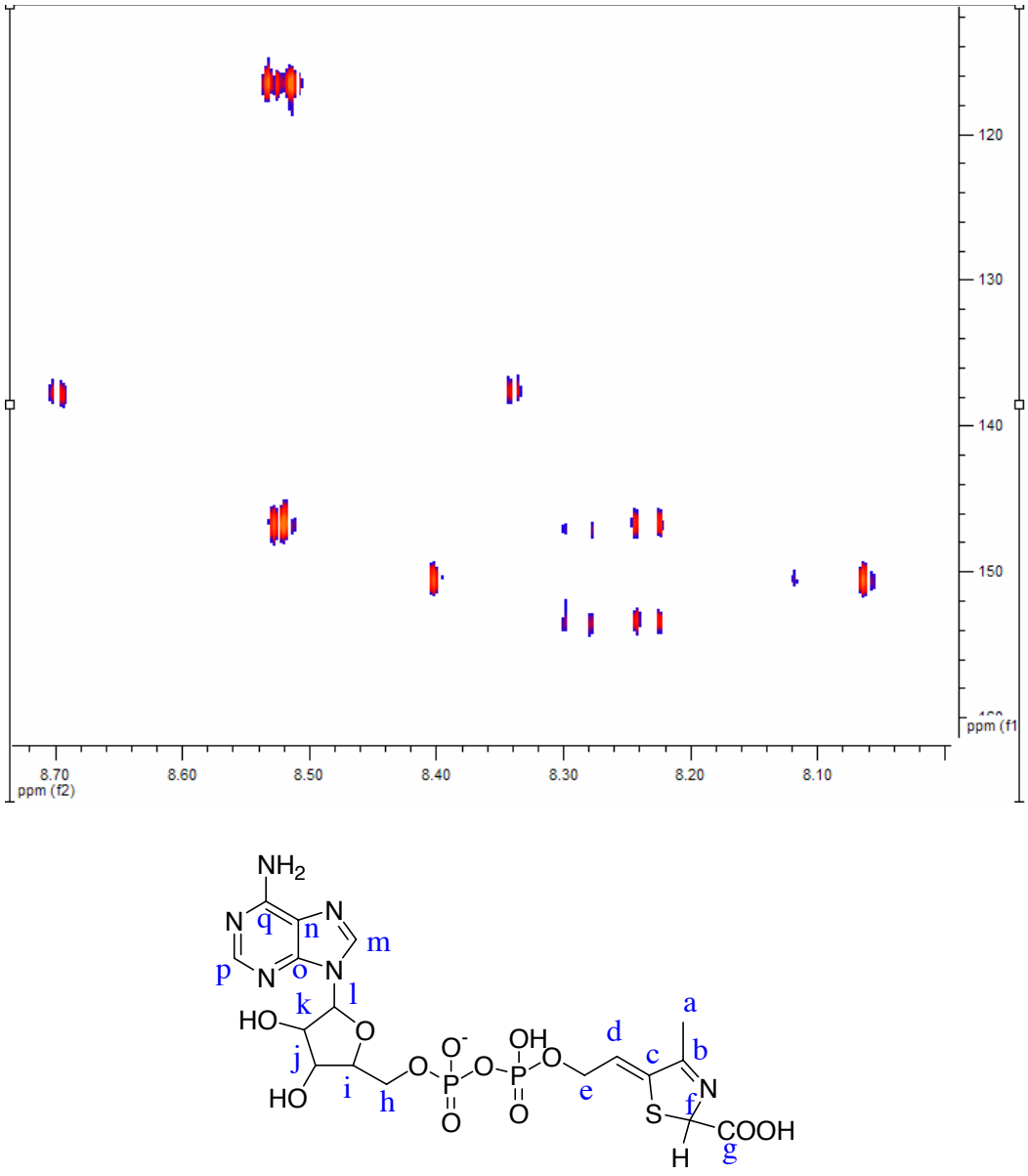


Figure SI 4: The magnified aromatic region of the HMBC spectra. Two adenine protons (**m** 8.52 and **p** 8.23 ppm) couple to two carbon atoms each (**m** coupling to **o** 146.6 ppm and **n** 116.4 ppm; **p** coupling to **q** 153.3 and **o** 146.6 ppm). Also both the protons couple to the carbon atom it is attached to (137.7 ppm for **m**; 150.4 ppm for **p**). These show up as these correlations have unusually high coupling constants and are not filtered out by the normal parameters used. These correlations appear as doublets with large coupling constants. Also, the minor degradation product ADP which has a proton peak at 8.29 ppm shows two carbon cross peaks very similar to the proton **p**. This probably comes from the adenine moiety of ADP.

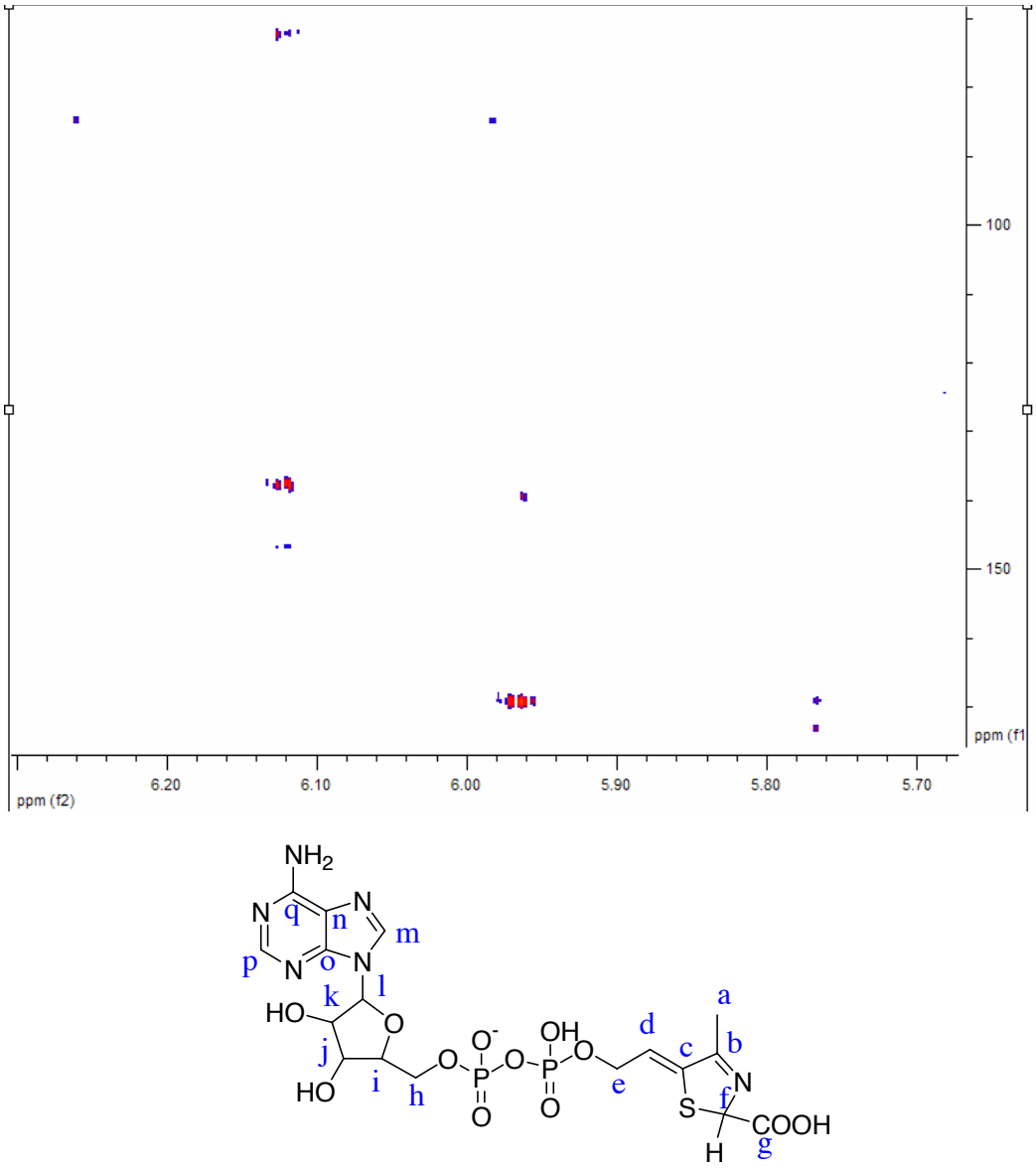


Figure SI 5: Magnified region of HMBC spectra. The ribose C-1 proton {**l**} shows three long range correlations: with **k** 71.9 ppm, **m** 137.7 ppm and **o** 146.6 ppm. It also shows a weak one bond correlation to the carbon it is attached to (84.6 ppm signal appearing as a doublet with a large coupling constant). The proton **d** (5.96 ppm) shows two long range correlations: with **c** 139.3 ppm and **b** 169.2 ppm. The proton **f** (5.76 ppm) shows two weak correlations: with **g** 173.2 ppm and with **b** 169.2 ppm.

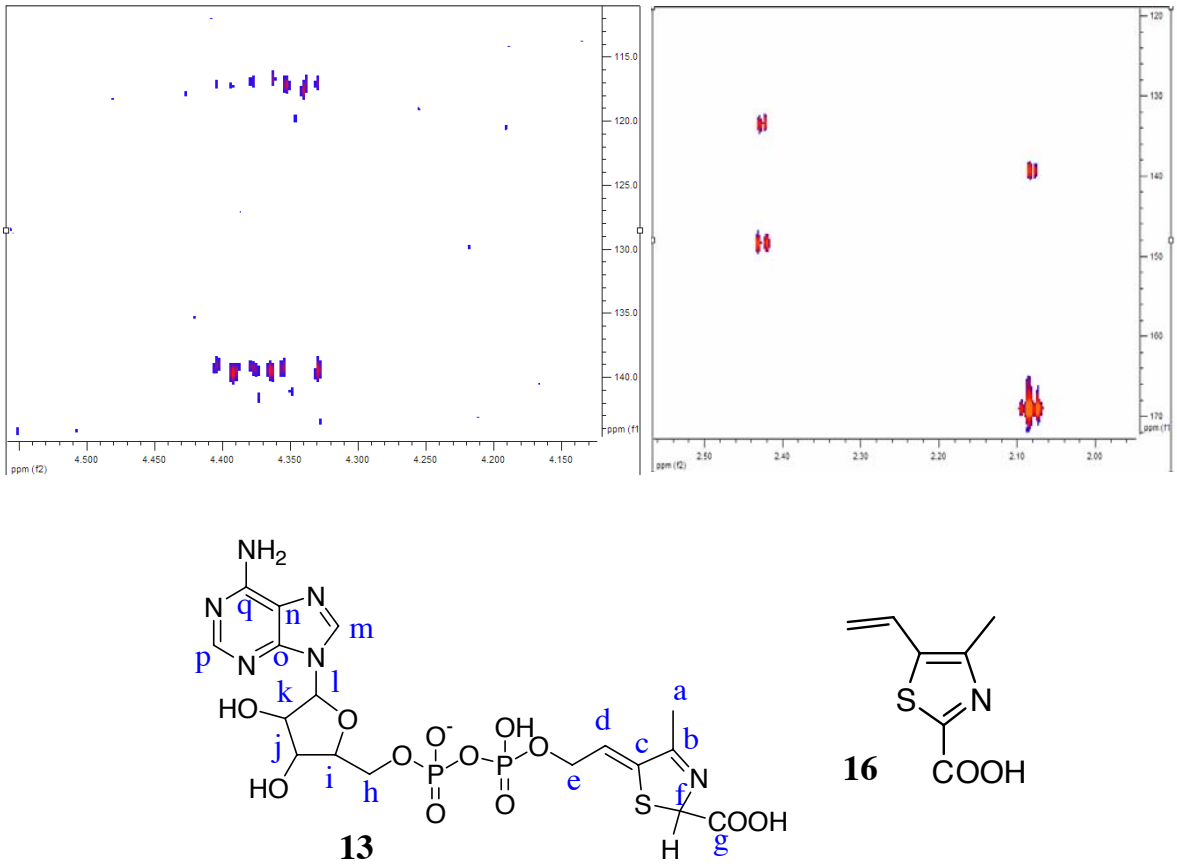


Figure SI 6: Magnified region of HMBC spectra. Left: **e** (4.37 ppm) protons show two long range correlation with **c** 139.3 ppm and **d** 117.05 ppm. Right: **a** prtons (2.08 ppm) show two strong correlations: with **b** 169.2 ppm and **c** 139.3 ppm. The methyl unit of **16**, the degradation product of **13**, gives rise to a proton peak at 2.43 ppm, which couples with two aromatic carbon atoms of the thiazole ring at 133.4 ppm and 148.3 ppm.

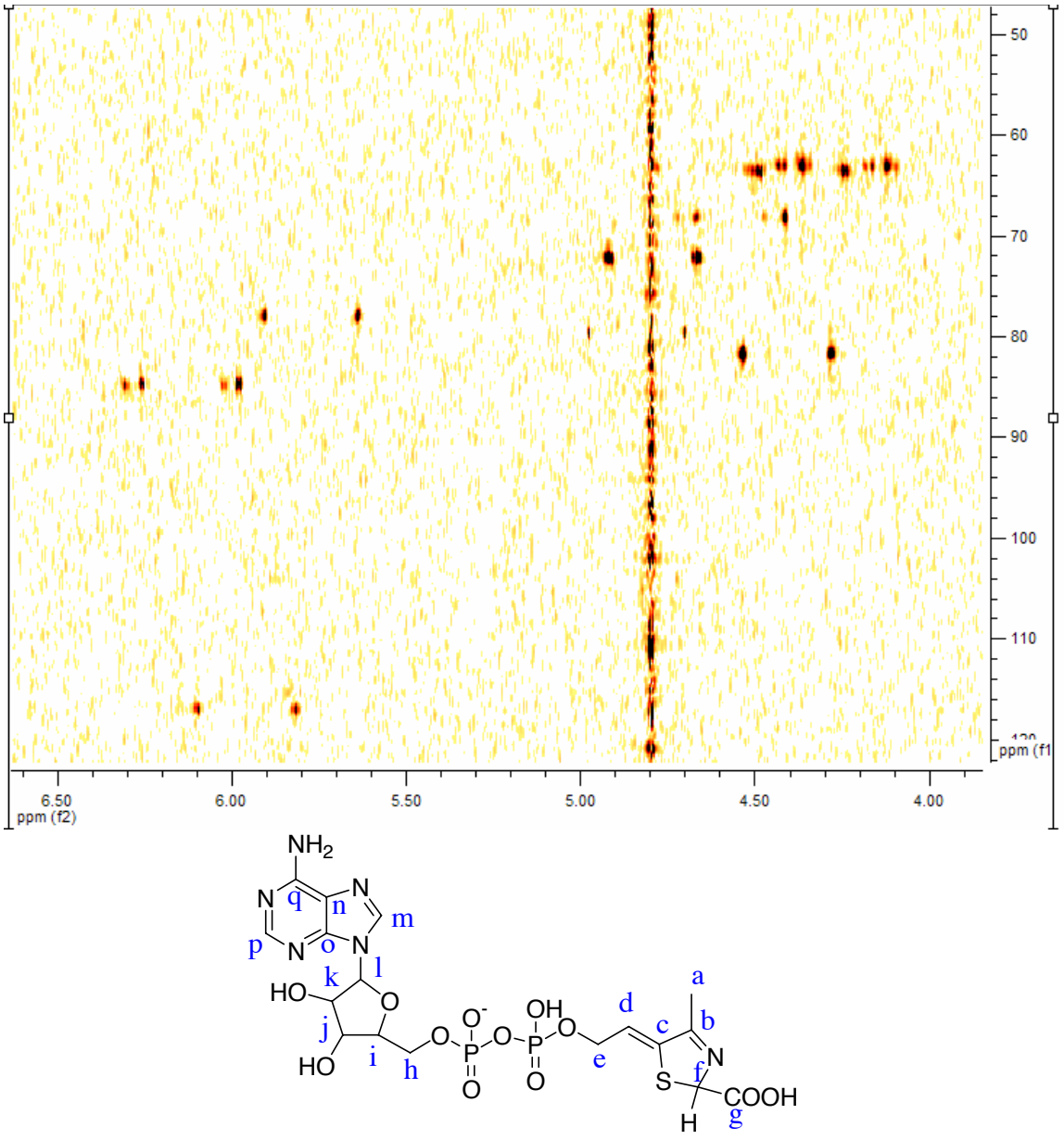
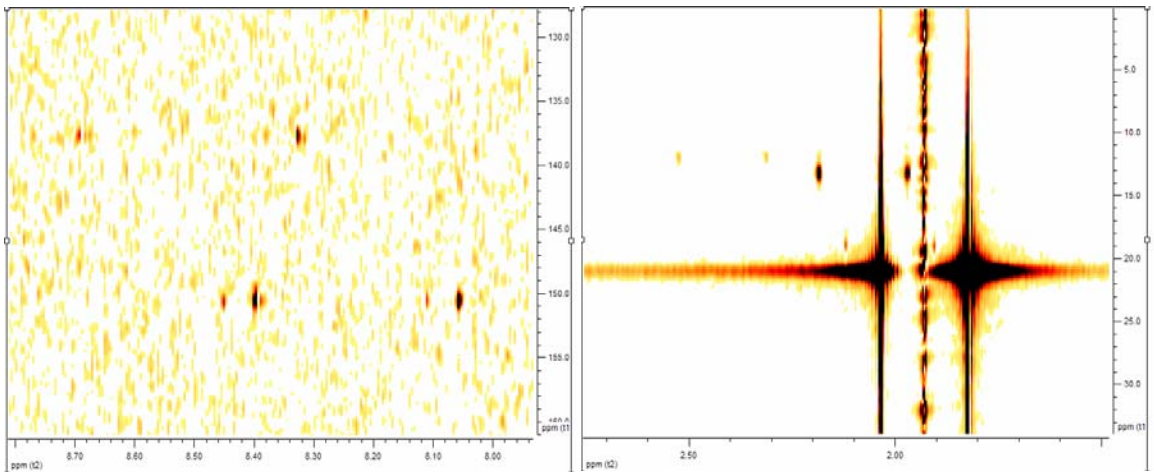


Figure SI 7: Magnified region of the HMQC (one bond ^{1}H-^{13}C) correlation. In this particular diagram, chemical shifts for the following carbon atoms are obtained: **d** (117 ppm) attached to proton 5.96 ppm; **e** (63.5 ppm) attached to proton 4.37 ppm; **f** (77.8 ppm) attached to proton 5.76 ppm; **h** (62.9 ppm) attached to proton 4.24 ppm; **i** (81.7 ppm) attached to proton 4.41 ppm; **j** (68.1 ppm) attached to proton 4.54 ppm; **k** (72.1 ppm) attached to proton 4.8 ppm; **l** (84.6 ppm) attached to proton 6.12 ppm. This correlation diagram also reveals the presence of another ribose unit (from ADP, the degradation product of **13**).



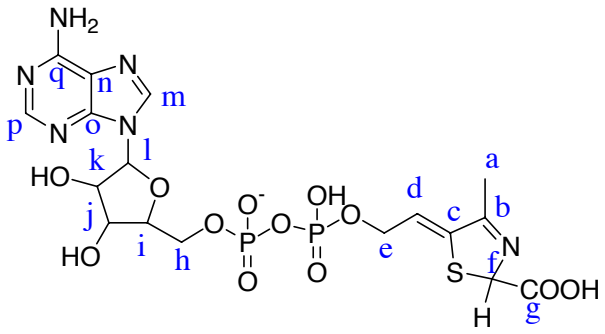


Figure SI 8: Magnified images from the HMQC correlation diagram. Left: Adenine carbon **m** (137.7 ppm) attached to proton 8.52 ppm and **p** (150.4 ppm) attached to proton 8.23 ppm are observed. Right: The methyl carbon **a** (13.2 ppm) attached to proton 2.08 ppm is visible. Also, the carbon atom of the methyl unit of the degradation product 16 is visible at 11.9 ppm attached to the corresponding protons resonating at 2.41 ppm. Residual acetate gives rise to the strong background signal.

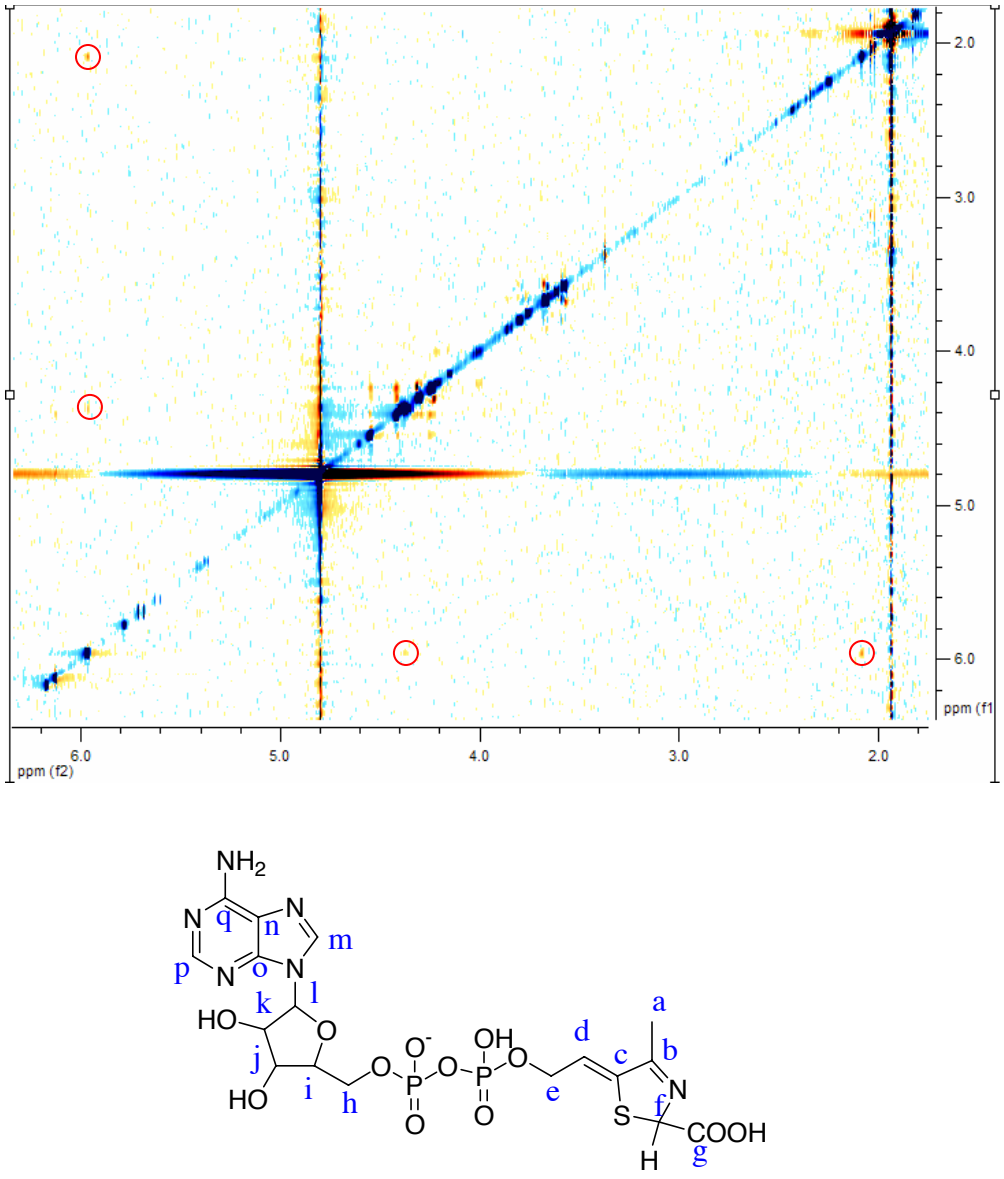
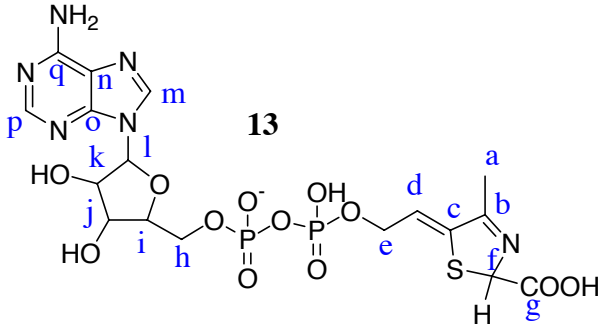


Figure SI 9: A ROESY correlation diagram showing NOE correlation between methyl proton **a** (2.08 ppm) and the proton **d** (5.96 ppm). No correlation could be observed between **a** and **e** protons. Proton **d** show weak correlation with proton **e** (4.37 ppm).

Summary of all the chemical shifts and the correlations deduced from the NMR spectroscopic analysis of compound 13:



13

Position	$\delta^{13}C$ [ppm]	$\delta^{1}H$ [ppm]	Relevant HMBC Correlations	Relevant COSY Correlations	Relevant ROESY Correlations
a	13.2	2.08	C-b, C-c	H-f	H-d
b	169.2	-	-	-	-
c	139.3	-	-	-	-
d	117.05	5.96	C-b, C-c	H-e, H-f	H-a, H-e
e	63.5	4.37	C-c, C-d	H-d, H-f	H-d
f	77.8	5.76	C-b, C-g	H-a, H-d, H-e	-
g	173.2	-	-	-	-
h	62.9	4.24	-	H-i	-
i	81.7	4.41	-	H-h, H-j	-
j	68.1	4.54	-	H-i, H-k	-
k	72.1	4.8	-	H-j, H-l	-
l	84.6	6.12	C-k, C-m, C-o	H-k	-
m	137.7	8.52	C-n, C-o	-	-
n	116.4	-	-	-	-

o	146.6	-	-	-	-
p	150.4	8.23	C-o, C-q	-	-
q	153.3	-	-	-	-

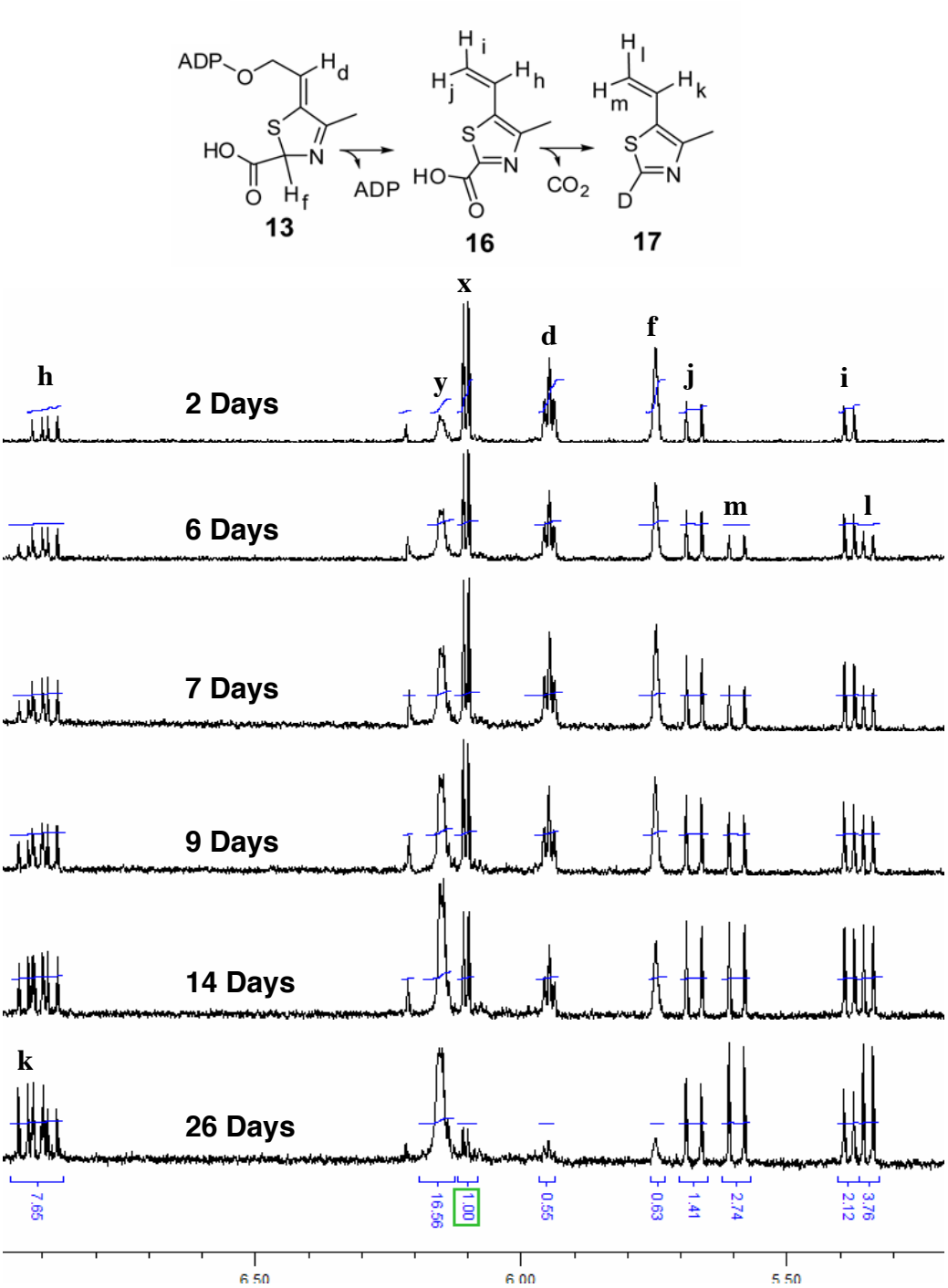


Figure SI 10: Extended ^{1}H NMR spectroscopic analysis of the degradation of **13** in D_2O. Resonances x and y represent the anomeric protons on the ribose of **13** and ADP respectively.

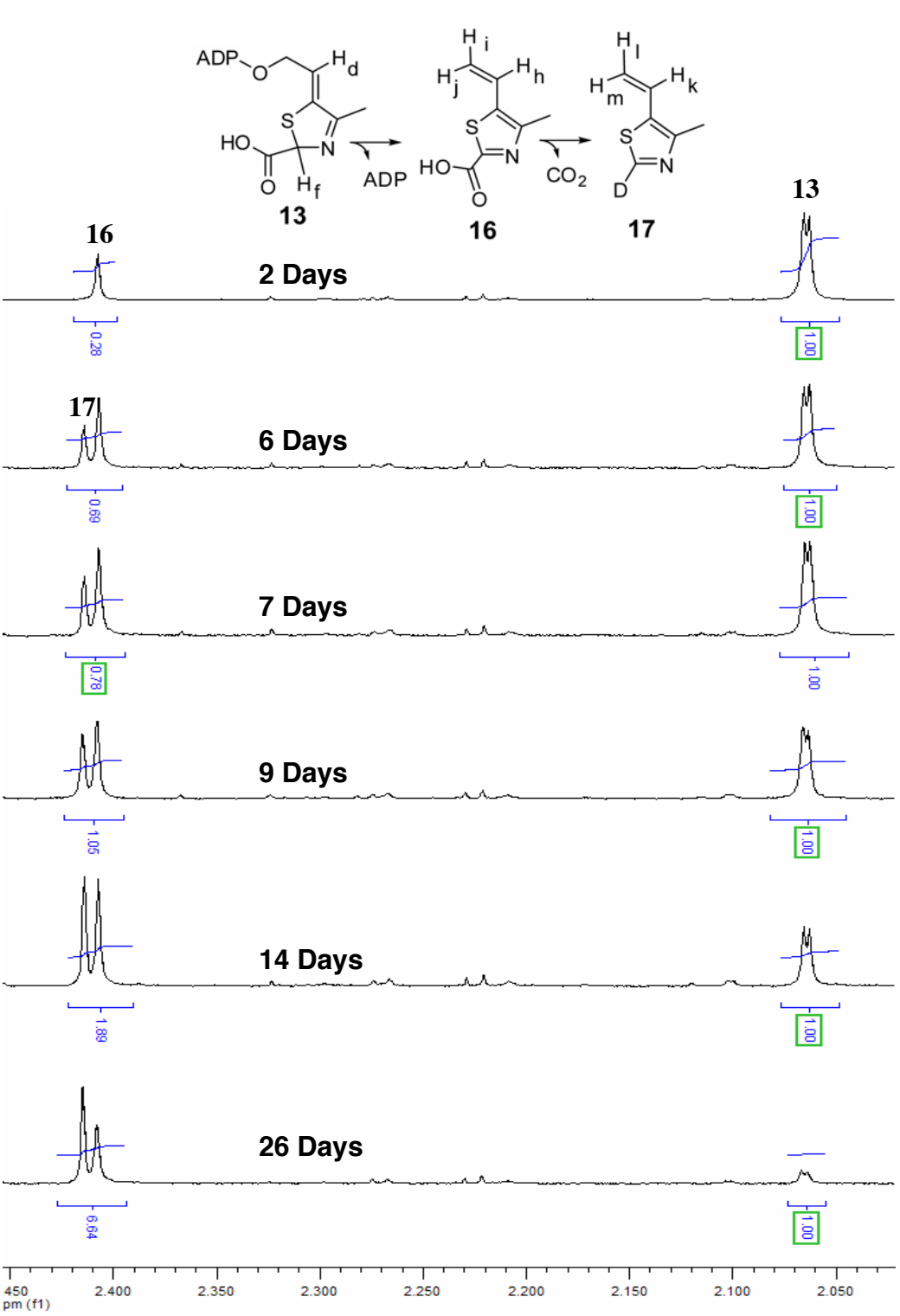


Figure SI 11: Extended ^{1}H NMR spectroscopic analysis of the degradation of **13** in D_2O (magnified region showing the methyl peaks of **13**, **16** and **17**). The integration values of the methyl peaks were used for the kinetic analysis of the degradation reaction.

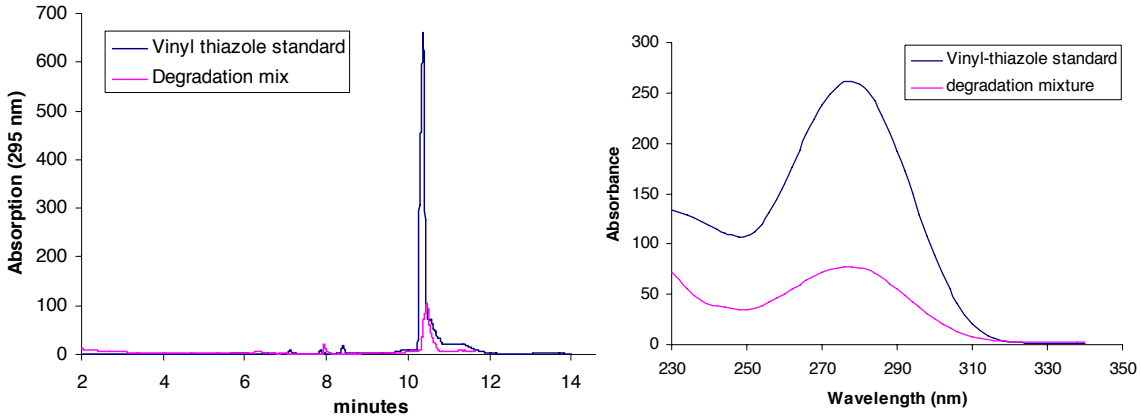


Figure SI 12: Identification of 4-methyl-5-vinylthiazole (**17**) as the final degradation product of **13**. Left: HPLC chromatogram demonstrating that the degradation product from a sample of **13** (purple trace), incubated for 40 days at room temperature, comigrates with an authentic standard of **17** (blue trace). Right: UV absorption spectra of an authentic standard of **17** (blue) and the degradation product of **13** (purple) are identical.