Supporting Information

Direct Observation of Multiple Tautomers of Oxythiamine and their Recognition by the Thiamine Pyrophosphate Riboswitch

Vipender Singh,[†,‡,¶] Chunte Sam Peng,[†] Deyu Li,[†,‡,¶] Koyel Mitra,[§] Katherine J. Silvestre,[†] Andrei Tokmakoff,[†,╫] and John M. Essigmann[*,†,‡,¶]

[†]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

[‡]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

[¶]Center for Environmental Health Sciences, Massachusetts Institute of Technology Cambridge, MA 02139

[╫]Present address: University of Chicago, 929 E. 57th St., Chicago, IL 60637.

[§]Current Address: Dept. of Medicine, University of California San Diego, 9500 Gilman Drive, San Diego, CA 92122

[*]To whom correspondence should be addressed.

Mailing Address: Room 56-669; 77 Massachusetts Avenue, Cambridge, MA 02139

Telephone number: 617-253-6227. E-mail: jessig@mit.edu

Notes

Note 1. From the variable temperature NMR spectra of oxythiamine in DMF, we can tell the first appearance of three distinguishable peaks of the tautomeric proton is -50 °C. By using the equation: exchange rate = $\sqrt{2}/(\pi\Delta\nu) = 1/(2.22\Delta\nu)$, we can calculate the exchange rate at this temperature between peak 16.6 and 15.5 ppm equals $1/(2.22\Delta\nu) = 1/(2.22 * 1.1 \text{ ppm} * 500 \text{ Hz/ppm}) = 1/(2.22 * 550 \text{ Hz}) = 8.2 * 10^{-4}$ S = 0.82 ms. The other two exchange rates can also be calculated: the exchange rate between 15.5 and 13.3 is 0.41 ms and the exchange rate between 16.6 and 13.3 is 0.27 ms.

Note 2. Generally, the chemical shift of aromatic amide proton is around 10-14ppm, e.g. the amide proton is 10.8 ppm in guanosine and 11.3 ppm in thymidine; the chemical shift of aza-phenol is around 9.5 ppm, e.g. the phenol proton in 6-methyl-3-pyridinol is 9.6 ppm; the chemical shift of endocyclic aromatic amino proton is around 8.0 ppm, e.g. the amino proton of in pyrrole is 8.0 ppm. To make these data comparable, all ^{1}H NMR chemical shift data are taken from Spectral Database for Organic Compounds, SDBS (http://riodb01.ibase.aist.go.jp/sdbs/cgi-bin/cre_index.cgi?lang=eng.)

Note 3. According to the spectroscopic data from SDBS, Spectral Database of Organic Compounds, the peak at 8.6 ppm was assigned to the proton at the C6' position, the 9.7 ppm peak corresponds to two 4'-NH_2 amino protons. The peak at 10.3 ppm was assigned to proton at the 2-position of the thiazole ring. The peak at 15.7 ppm was assigned to the HCl proton. The remaining peaks, below 8.5 ppm can be assigned using the thiamine hydrochloride spectrum deposited at the SDBS database. Following is the link to the SDBS database: http://riodb01.ibase.aist.go.jp/sdbs/cgi-bin/direct_frame_top.cgi

Table S1. Showing the pH of 20 mg/ml of oxythiamine in various concentrations of HEPES buffer pH 7.5 and in water.

20 mg/ml Oxythiamine	**pH**
H_2O	1.4
0.1 M HEPES pH 7.50	3.48
0.25 M HEPES pH 7.50	6.17
0.5 M HEPES pH 7.50	6.91
1.0 M HEPES pH 7.50	7.23

Figure S1. Vibrational normal modes of 4'-keto-N1'H-OxyT tautomer with two D_2O molecules from DFT calculations performed using QChem software. (*S8*)

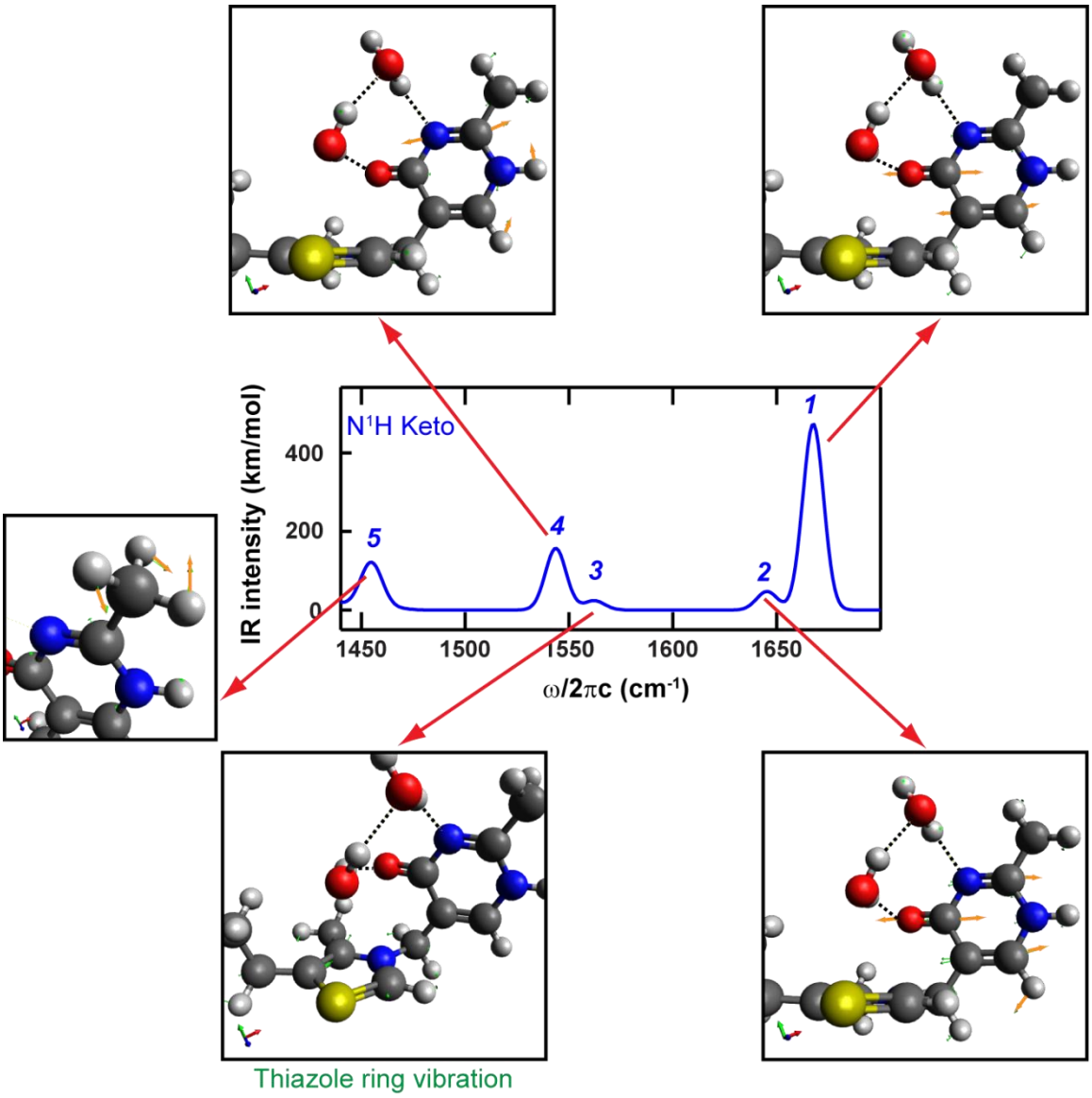


Figure S2. Vibrational normal modes of 4'-keto-N3'H-OxyT tautomer with two D_2O molecules from DFT calculations.

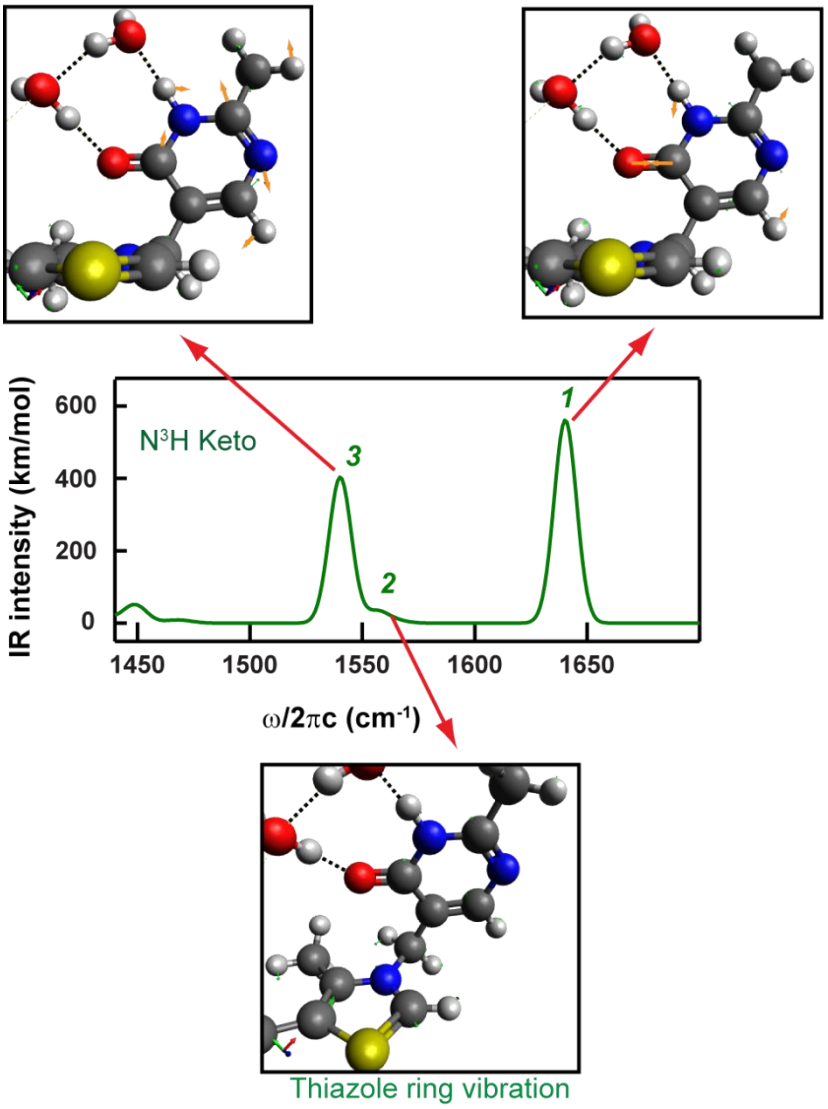
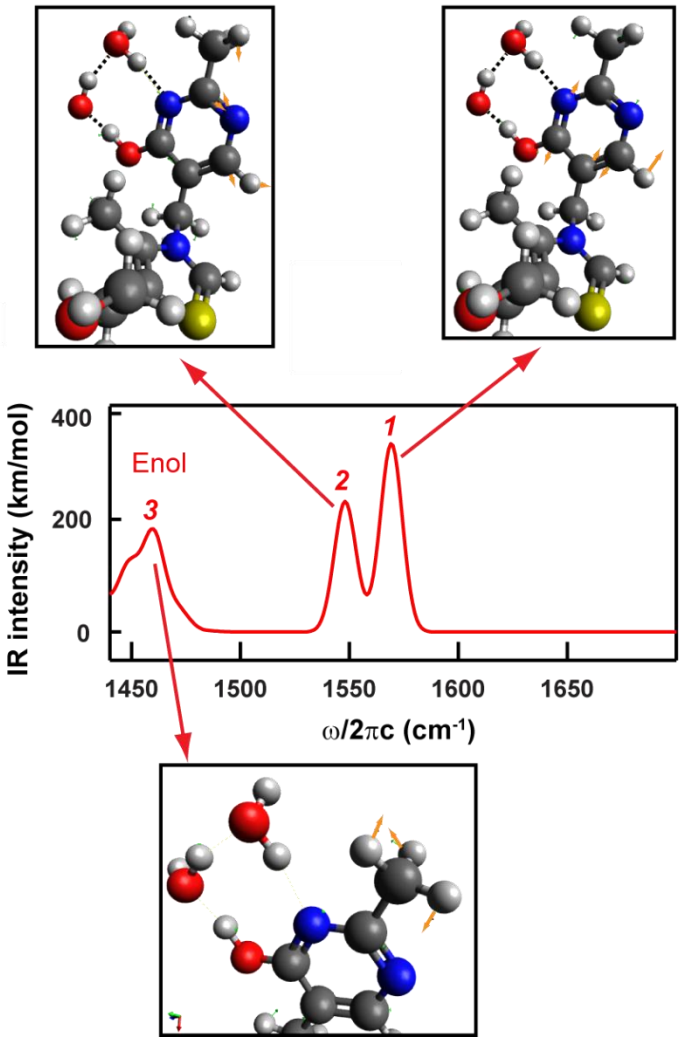


Figure S3. Vibrational normal modes of 4'-enol-OxyT tautomer with two D_2O molecules from DFT calculations.



Method M1. Fitting routine for FTIR and 2D IR spectra.

To calculate and fit the experimental FTIR and 2D IR spectra, a nonlinear response function was used and described in detail previously.(S*1*, S*2*) Since the three tautomers of OxyT have distinct vibrations in the 1580 – 1680 cm^{-1}, only the experimental spectra in this frequency range were fit. Before fitting the FTIR spectra, we applied a linear correction from 1580 to 1700 cm^{-1} to the temperature dependent OxyT FTIR spectra after the temperature dependent buffer spectra have been subtracted. Three uncoupled anharmonic oscillator with Kubo line-shape were used to fit both the FTIR and 2D IR spectra. The Kubo line-shape is calculated from an energy-gap correlation function which contains a single exponential decay $\Delta exp(-|t|/\tau)$, and a $\delta(t)/T_2$ term to produce the homogeneous linewidth in the 2D IR spectra.(S*3*) A simultaneous fit of both the FTIR and 2D IR spectra (Figure S4) well reproduced the experimental data (Figure S5) with the fitting parameters listed in Table S2.

In this model since the three anharmonic oscillators are uncoupled, no cross-peaks should be present in the calculated spectrum. The off-diagonal features seen in the experimental spectrum are from the tails of the Lorentzian peaks extending to the X1/X3 cross-peaks (see Figure 2A) which do not exist in the calculation. In addition, the DFT calculations show that both the keto tautomers may have contributions to the X2 mode as vibrations corresponding to the thiazole ring vibrations (Figures S1 and S2) are observed at 1560 cm^{-1}. We have neglected this contribution in this model due to their very weak intensities. The errors in determining the tautomer populations from neglecting these contributions are small and within the error bars reported in this study.

Figure S4. (a) Experimental FTIR of OxyT in TPP buffer taken at room temperature (blue curve), and calculated FTIR (red curve) using the fitting parameters listed in Table S2 (b) Experimental (left) and fitted (right) 2D IR spectra.

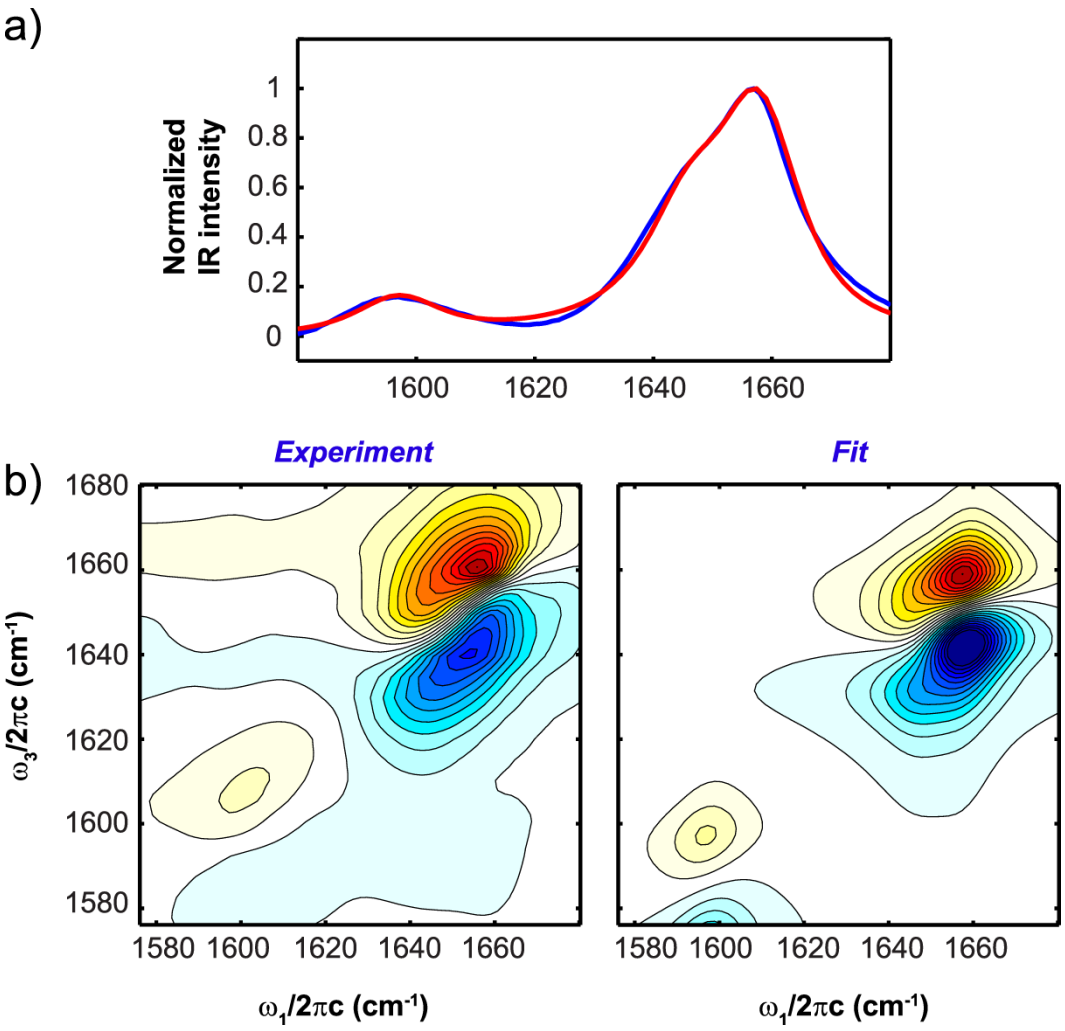


Table S2. Fitting parameters from simultaneous fitting of experimental FTIR and 2D IR at room temperature. The model includes three independent oscillators representing the three tautomers seen in OxyT in TPP buffer. The ω symbol is the transition frequency; $\Delta\omega$ is the diagonal anharmonicity; μ_{10} is the transition dipole strength for the zero- to one- quantum states; μ_{21} is the transition dipole strength for the one- to two- quantum states; T_2, Δ, and τ are the lineshape parameters. The Kubo line-shape function was used: $g(t) = t/T_2 + \Delta^2\tau^2\left[exp(-t/\tau) - 1 + t/\tau\right]$.

	4'-keto-N1'H-OxyT	4'-keto-N3'H-OxyT	4'-enol-OxyT
N	0.5	0.42	0.08
ω	1658	1646	1597
$\Delta\omega$	8.5	12.9	19
μ_{10}	0.7	0.68	0.74
μ_{21}	1.12	0.93	1.29
T_2	1515	600	2066
Δ	45.6	40.1	52.8
τ	617	80.7	593.6

In order to determine the error bars on the tautomer populations, we tested the robustness of the two keto populations in reproducing the FTIR spectrum. We fit the FTIR spectrum by fixing μ_{10} from the simultaneous fit and either one of the keto populations, and then allowed the other parameters (another keto population and line-shape parameters) to float. We stepped the population of 4'-keto-N1'H-OxyT from 25 % to 75 %, and calculated the residual relative to the experimental spectrum. We then defined the acceptance for a good fit to be 20% deviation of the relative residual from the best fit. We found that the experiment can be well reproduced with the population of 4'-keto-N1'H-OxyT varying between 50 % to 60 %. Similar procedure was done for the population of 4'-keto-N3'H-OxyT and it was found to be between 30 % to 44 %. The population of the enol tautomer was found to vary between 7 % to 9 %. The result shows about 14 % error in determining the keto populations as opposed to the 2 % error in enol population. The overlap between the two keto peaks leads to the larger uncertainty. Other sources of error include baseline uncertainty in the FTIR spectra, lineshapes, and the fact that we may be neglecting keto contributions to the X2 mode. However, self-consistent fitting the 2D IR spectrum in addition to the FTIR enhances our confidence limits.

Figure S5. Variable temperature FTIR spectra from 5 °C to 90 °C: (a) experiment; (b) fit. The fitting procedure is the same as described for fitting the room temperature spectrum. The transition dipole strengths were assumed to be independent of temperature. Transition frequencies were allowed to float within ± 3 cm^{-1} from the values determined at room temperature to account for the temperature induced frequency shifts.

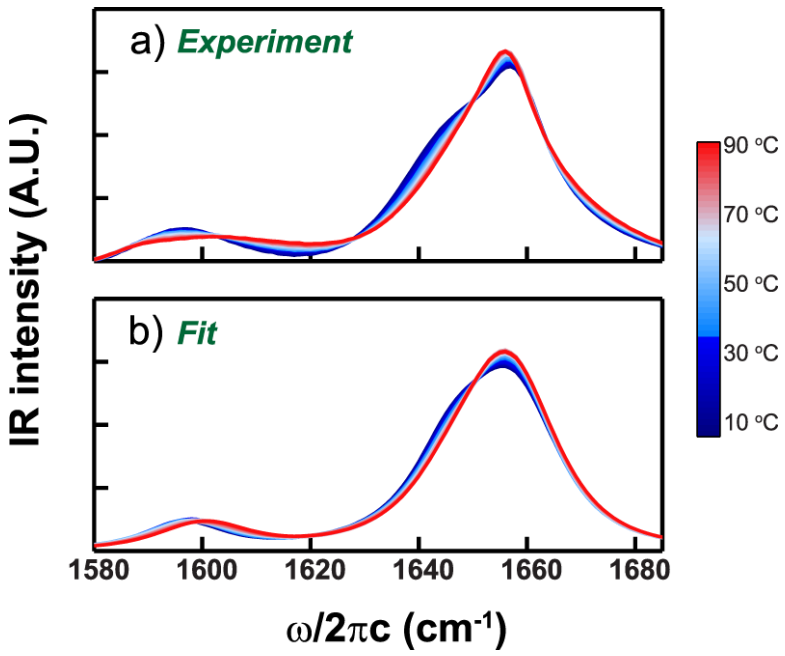
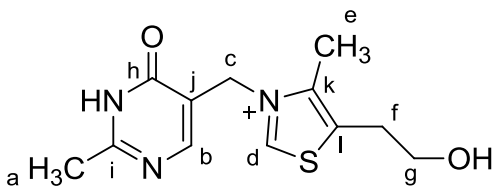


Table S3. ^{1}H and ^{13}C NMR signal assignments of OxyT in DMF-d_7 at 20 °C.



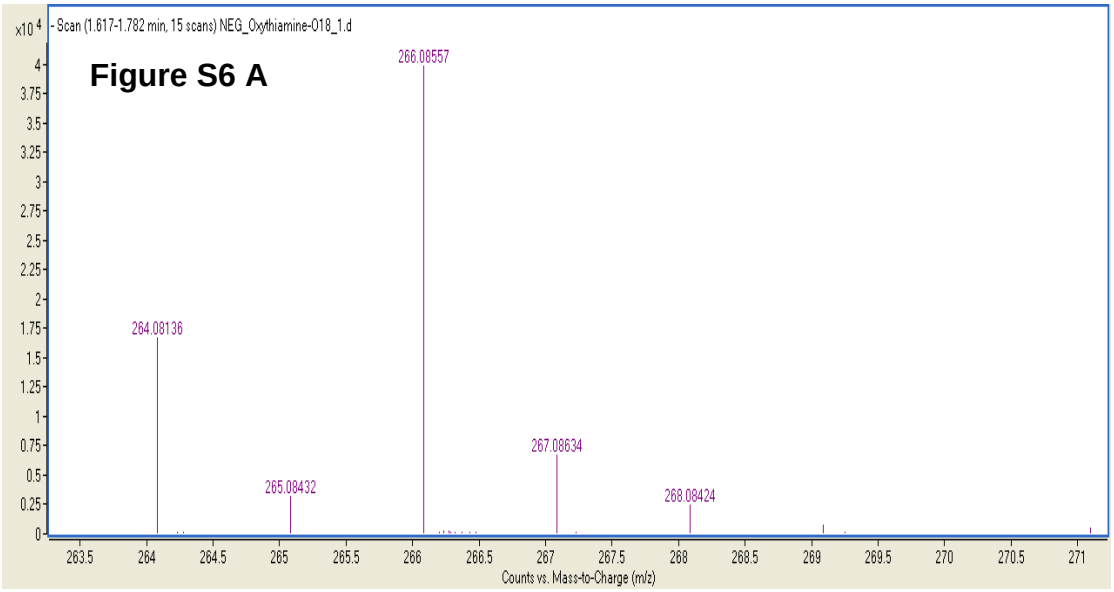
^{1}H position	^{1}H chemical shift in DMF-d_7	^{1}H peak pattern and coupling constant in DMF-d_7	^{13}C chemical shift in DMF-d_7
a	2.68	Singlet	21.7
b	8.35	Singlet	*
c	5.66	Singlet	51.5
d	10.30	Singlet	159.0
e	2.50	Singlet	12.4
f	3.17	Triplet J=5.5	30.1
g	3.79	Triplet J=5.5	61.2
h			*
i			*
j			100.8
k			143.4
l			136.6
HOD	3.74	Singlet (br)	
DMF aldehyde	8.03	Singlet	163.2
DMF methyl a	2.92	Quintet	34.9
DMF methyl b	2.75	Quintet	29.8

* Due to the tautomerization of the pyrimidine ring, these carbon signals did not display in the ^{13}C NMR spectrum.

Figure S6. A. Quadrupole time-of-flight (QTOF) mass spectrometry data for OxyT in the negative mode. **B.** Extracted intensities and retention times for each m/z of OxyT for the purpose of calculating ^{18}O enrichment. **C**. Mass spectrometry data of OxyT in the positive mode

Expected masses for ^{16}O and ^{18}O OxyT in the negative mode:

OxyT – [^{16}O] – 264.10 ; OxyT – [^{18}O] – 266.10



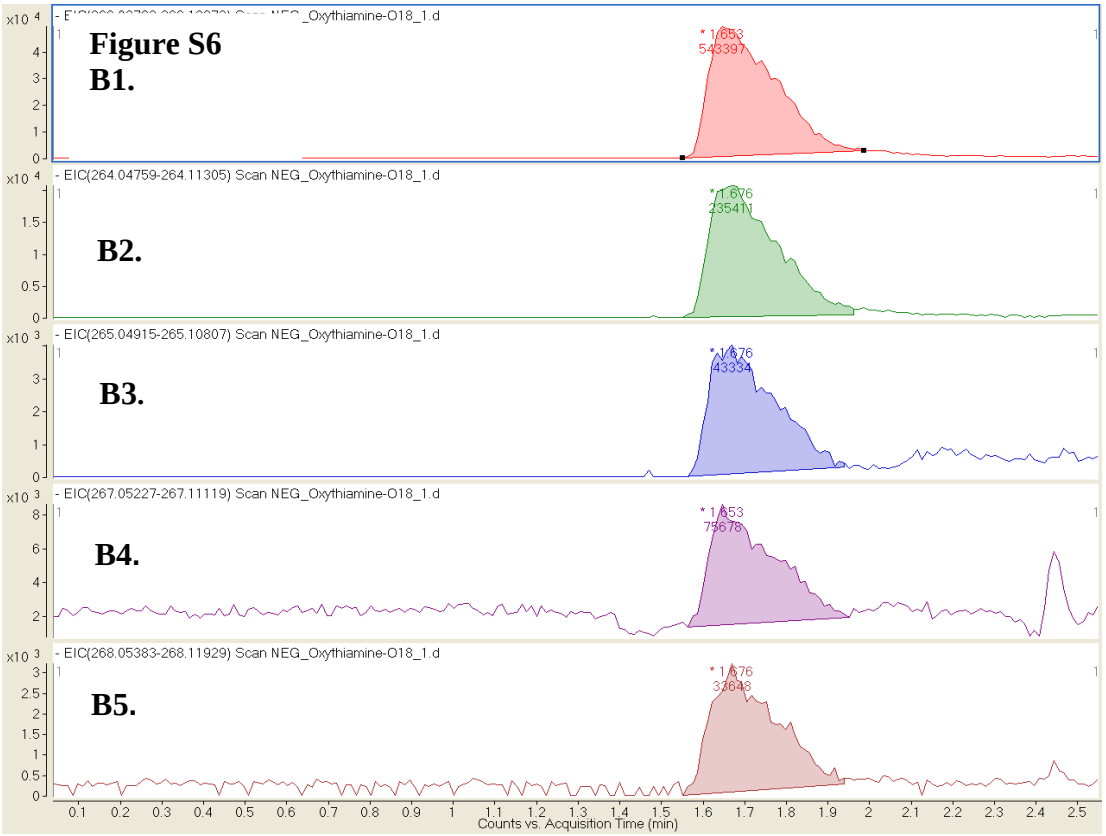
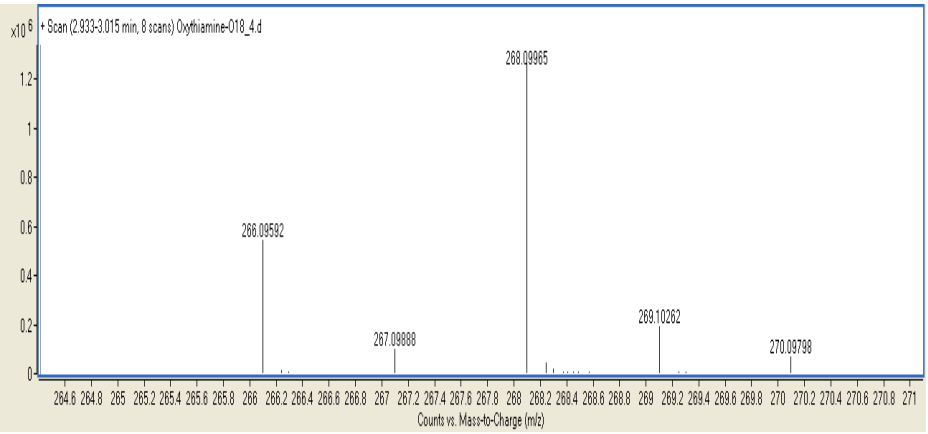


Figure S6. C. QTOF mass spectrometry data for OxyT in the positive mode.

OxyT – [^{16}O] – 264.10 (266.10 in the positive mode)

OxyT – [^{18}O] – 266.10 (266.10 in the positive mode)



Calculation of ^{18}O Enrichment Factor

The enrichment factor for ^{18}O OxyT was obtained by dividing the sum of the area from all the peaks from ^{18}O OxyT (B1 + B4 + B5) by the total area of peaks for both ^{16}O and ^{18}O OxyTs (B1+B2+B3+B4+B5) as shown in figure S6 B.

$$^{18}O - Enrichment = \frac{Area(B1 + B4 + B5)}{Area(B1 + B2 + B3 + B4 + B5)} = 0.725$$

Methodology for determination of ^{18}O enrichment for OxyT- Identification and quantification of ^{18}O enrichment of synthesized OxyT was accomplished by coupling the HPLC to an Agilent 6510 quadrupole time-of-flight (QTOF) mass spectrometer equipped with an electrospray ionization (ESI) source operated in both negative and positive ion mode (supporting information). Operating parameters were as follows: ESI capillary voltage, 43500 V; gas temperature, 350 $^{\circ}$C; drying gas flow, 10 L/min; nebulizer pressure, 30 psi; fragmentor voltage, 175 V; and *m*/*z* scan range, 100-1000. HPLC. oxythiamine (4'-^{16}O and 4-^{18}O) were resolved with a Dionex Acclaim PolarAdvantage (C16 2.1× 150 mm, 3 µm particle size) eluted with 0.1% acetic acid in water at a flow rate of 200 microliter/min at 27 C: 0–10min, 0% B; 10–40 min, 80%; 40–41min,100,41-51, 100% and 52-53 min, 0% and post run of 15 minutes 2–7%; 28–30 min, 7%; 30–31 min, 7–100%; 31–41 min, 100%. Results are shown in Figure S8A, S8B and S8C.

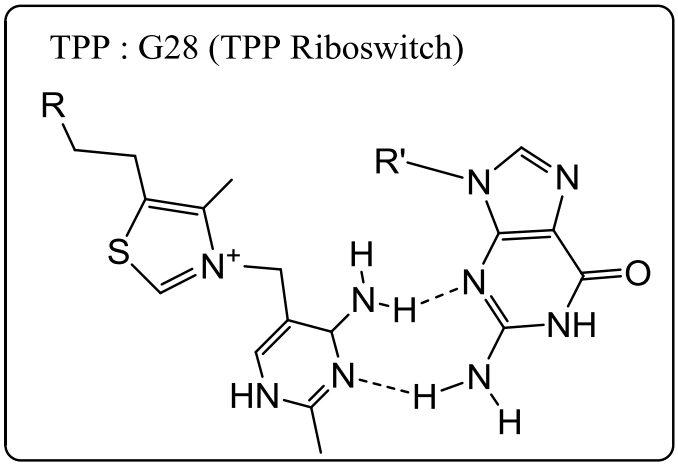
Table S4. Titration experiment showing the effect of addition of unlabeled TPP on O-18 BIE measured using [P-33, O-18] and [P-32, O-16] OxyTPP (K_d ~ 1,67 M)). At saturating concentrations of unlabeled TPP (K_d ~ 50 nM), the labeled OxyTPP is mostly present in the unbound form and a BIE similar to control with no riboswitch was obtained.

TPP concentration	O-18 Binding Isotope Effect
0.1	**0.971**
0.25	**0.982**
0.5	**0.992**
1.0	**1.009**
3.0	**1.005**

Table S5. The 4'-^{18}O BIEs for the desolvation of tautomers of OxyT, as mimicked by removal of three explicit water molecules. BIEs were calculated using the B3LYP functional and 6-31G (d, p) basis set implemented in Gaussian09. (S9)

	4'-^{18}O BIEs
Desolvation of 4'-keto-N1'H-OxyT	0.997
Desolvation of 4'-keto-N3'H-OxyT	0.997
Desolvation of 4'-enol-OxyT	1.004
Full deprotonation the N1' proton of 4'-keto-N1'H-OxyT tautomer	1.0008
Full deprotonation the N3' proton of 4'-keto-N3'H-OxyT tautomer	1.0007

Figure S7. A comparison of hydrogen bonding interactions between TPP and G28 (guanine at the 28th position) of the TPP riboswitch.



R= pyrophosphate.

R'=riboswitch backbone.

Figure S8. Showing that OxyTPP specifically binds to the TPP riboswitch, no non-specific binding with unrelated RNA (thrombin RNA aptamer) was observed.

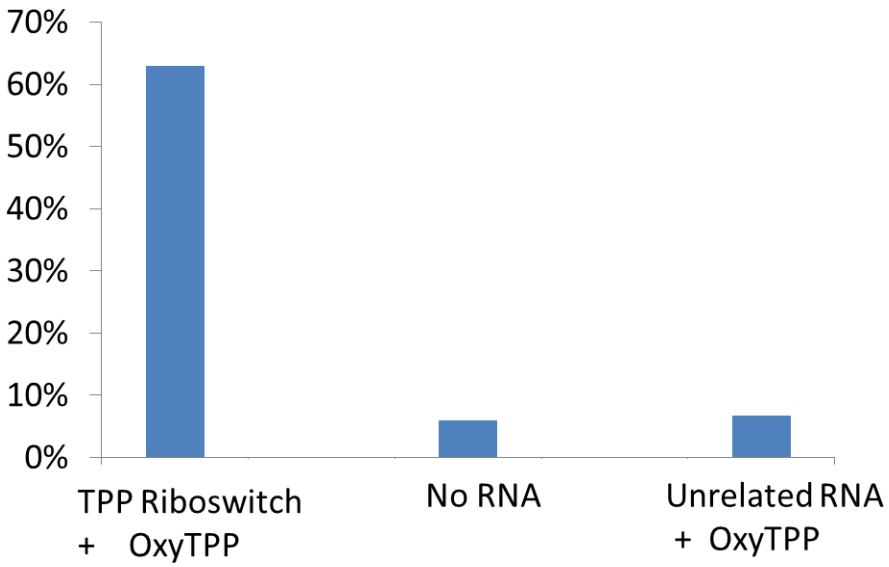
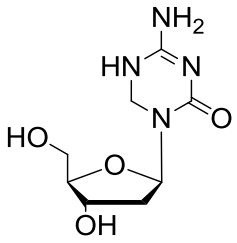
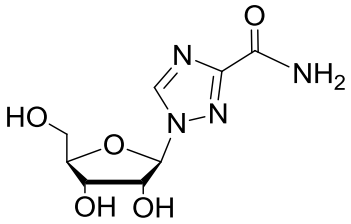


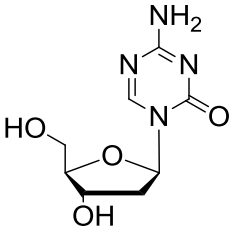
Figure S9. Tautomerizable/Rotamerizable nucleoside analogues with antiviral and anticancer properties: 5, 6-dihydro-5-aza-2'-deoxycytidine (KP1212),(S*4*) drug ribavirin (1-D-ribofuranosyl-1, 2, 4-triazole-3-carboxamide),(S*5*) 5-aza-2'-deoxycytidine,(S*6*) 5-hydroxy-2'-deoxycytidine.(S*7*)



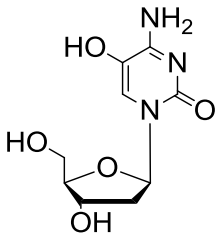
KP1212
(5,6-dihydro-5-aza-2'-deoxycytidine)



Ribavirin
(1-D-ribofuranosyl-1,2,4-triazole-3-carboxamide)



5-aza-2'-deoxycytidine



5-hydroxy-2'dexocycytidine

Method M2. Synthesis of double labeled OxyTPP**.**

The ^{18}O labeled OxyT was synthesized by deamination of thiamine by heating it at high temperature in the presence of sulphuric acid in ^{18}O water. About 50 mg of thiamine was added to a round bottom flask that has 1.82 ml of ^{18}O (> 97%) water, and then slowly 0.68 ml of H_2SO_4, to about 10 N concentration, was added to the reaction in a drop wise manner. The reaction mixture was heated to 80-90 $^{\circ}C$ for about 12-14 hours under argon. Following which it was cooled to room temperature. The reaction was quenched by neutralizing the pH by adding about 0.5 grams of NaOH, additional ^{18}O water was added to dissolve the NaOH pellets. The solvent was evaporated under vacuum using Rotavap. The reaction was then washed twice first with 12.5 ml ethanol and then with equal volume of methanol. It was filtered remove the undissolved salt. The filtrate was concentrated under vacuum and purified using a silica column using a solvent containing 45% water, 40% isopropanol and 15% acetate buffer. The OxyT obtained was again purified using HPLC first using a reverse phase C-18 column and then on a C-16 column using 10 mM ammonium acetate buffer pH 5.0. The purified oxythiamine was examined by LC-MS Q-TOF (Liquid Chromatography Quadrapole-Time-of-Flight) mass spectrometry to check for its purity and ^{18}O enrichment.

The ^{18}O labeled OxyT was double labeled with ^{32}P or ^{33}P at the terminal phosphate position using ^{18}O OxyT and ^{32}P or ^{33}P labeled ATP using thiamine pyrophosphokinase (TPK) enzyme. The histidine-tagged TPK enzyme was expressed and purified using *E. coli* for this purpose using affinity Nickel column chromatography. The doubled labeled OxyTPP was purified using a 20% denaturing PAGE gel, bands were excised and the labeled OxyTPP was

eluted with water by soaking the crushed gel overnight in water at 4 °C. The labeled compound was eluted from the gel fragments by centrifugation using Millipore spin columns with 10,000 KDa cutoffs filters. The eluted compound was concentrated and purified again using a 20% non-denaturing PAGE gel. The double labeled OxyTPPs were subsequently purified on HPLC with a C-16 column using 10 mM ammonium acetate buffer pH 5.0. Following purification the radioactivity was quantitated using scintillating counting. The ^{32}P and ^{33}P were mixed in approximately 1:4 ratios. The mixtures were finally purified again using 10 mM ammonium acetate pH 5.0, evaporated to dryness and dissolved in water and stored at -20 °C for BIE experiments.

Method M3. Synthesis of TPP Riboswitch.

The TPP riboswitch containing the T7 polymerase binding site was cloned into pUC19 vector and was amplified using *E. coli* cells. The RNA was generated by *in vitro* transcription using T7 polymerase. It was purified using 8% denaturing PAGE gel. The RNA was eluted from gel and exchanged into water by washing it 5 times with 50 ml of autoclaved deionized water using 50 ml Millipore stirred-cell using 10 KDa membrane filters. It was concentrated and stored at minus 20 °C.

References

S1. Sung, J., and Silbey, R. J. (2001) Four wave mixing spectroscopy for a multilevel system, *J. Chem. Phys. 115*, 9266–9287.

S2. Khalil, M., Demirdöven, N., and Tokmakoff, A. (2003) Coherent 2D IR Spectroscopy: Molecular Structure and Dynamics in Solution, *J. Phys. Chem. A 107*, 5258–5279.

S3. Schmidt, J. ., Sundlass, N., and Skinner, J. . (2003) Line shapes and photon echoes within a generalized Kubo model, *Chem. Phys. Lett. 378*, 559–566.

S4. Mullins, J. I., Heath, L., Hughes, J. P., Kicha, J., Styrchak, S., Wong, K. G., Rao, U., Hansen, A., Harris, K. S., Laurent, J.-P., Li, D., Simpson, J. H., Essigmann, J. M., Loeb, L. A., and Parkins, J. (2011) Mutation of HIV-1 Genomes in a Clinical Population Treated with the Mutagenic Nucleoside KP1461, *Plos One 6*, e15135.

S5. Crotty, S., Cameron, C., and Andino, R. (2002) Ribavirin's antiviral mechanism of action: lethal mutagenesis?, *J. Mol. Med. 80*, 86–95.

S6. Dapp, M. J., Clouser, C. L., Patterson, S., and Mansky, L. M. (2009) 5-Azacytidine Can Induce Lethal Mutagenesis in Human Immunodeficiency Virus Type 1, *J. Virol. 83*, 11950–11958.

S7. Loeb, L. A., Essigmann, J. M., Kazazi, F., Zhang, J., Rose, K. D., and Mullins, J. I. (1999) Lethal mutagenesis of HIV with mutagenic nucleoside analogs, *Proc. Natl. Acad. Sci. 96*, 1492–1497.

S8 QChem, Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; Jr, R. A. D.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Voorhis, T. V.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Iii, H. L. W.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Iii, H. F. S.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

S9 Gaussian 09, Revision A.1, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford, CT, 2009.Frisch, M. *Gaussian09*; Gaussian Inc. Wallingford CT 2009.