Electronic Supplementary Information for:

Formic Acid Catalyzed Isomerization of Protonated Cytosine: A Lower Barrier Reaction for Tautomer Production of Potential Biological Importance

Lingxia Jin*,[a] Mengdan Lv[a] Mengting Zhao[a] Rui Wang[a] Caibin Zhao[a] Jiufu Lu[a] Ling Wang[a]

Wenliang Wang*,[b] Yawen Wei[c]

[a]*Shaanxi Key Laboratory of Catalysis, School of Chemical & Environment Science, Shaanxi Sci-Tech University, Hanzhong 723001*

[b]*Key Laboratory for Macromolecular Science of Shaanxi Province, School of Chemistry and Chemical Engineering, Shaanxi Normal University, Xi'an 710062*

[c]*Institute of publication Science，Chang'an University, Xi'an 710064*

Corresponding authors. Tel: +86-916-2641660; e-mail: jinlx@snut.edu.cn (L.X. Jin); Tel: +86-29-81530815 Fax: +86-29-81530727; e-mail: wlwang@snnu.edu.cn (W. L. Wang).

Table contents:

Section S1 Computational Methods

All the calculations were performed using the Gaussian 09 package.[1] The more reliable energies were conducted at the CBS-QB3 level of theory.[2] This highly accurate method is a composite approach, which is a five-step method starting with a CBSB7 geometry and frequency calculation, followed by CCSD(T), MP4SDQ, and MP2 single-point calculations as well as a CBS extrapolation. Moreover, many previous investigations have proposed that the CBS-QB3 method can provides adequately accurate energies, with a standard deviation of about 1.5 kcal•mol^{-1}.[3-5]

All the stationary points including reactant complexes, transition states, and products have been optimized using CBSB7 (B3LYP/6-311G (2d,d,p)) method. Vibrational frequency calculations have been conducted at the level of theory used for optimization to characterize the nature of each stationary point as a minimum (real frequencies) or transition state (only one imaginary frequency). Intrinsic reaction coordinate (IRC)[6] calculations have been carried out at the same level of theory from each transition state to ensure that the obtained transition states connected the appropriate reactants and products. Besides, to further investigate the effect of solvation, the obtained stationary points based on the gas phase geometries were further optimized using polarized continuum model (PCM)[7] with dielectric constant 78.39 at CBS-QB3 method.

1.1 IRC and reaction force *F*(ξ) analysis

The minimum energy paths (MEP) have been obtained by Intrinsic Reaction Coordinate (IRC) in mass-weighted Cartesian coordinates. In order to understand the mechanisms of tautomerisation reaction, the reaction force $F(\xi)$ has been introduced by the negative of the derivative of the potential energy $E(\xi)$ with respect to the reaction coordinate ξ: [8]

$$F(\xi) = -\frac{\mathrm{d}E}{\mathrm{d}\xi} \tag{1}$$

As the potential energy, the $F(\xi)$ presents a universal form along the reaction coordinate. For an elementary step it exhibits a minimum between the reactant and the transition state and a maximum between the transition state and the product. These particular points define the transition state region and correspond to transitions between different stages of the elementary step.[9-12]

1.2 Rate constants calculation and proton tunneling correction

The theoretical rate constants were calculated by the conventional transition state theory,[13]

which can be calculated from eq 2.

$$k = \frac{k_B T}{h} \exp(\frac{-\Delta G^{s\neq}}{RT}) \quad (1)$$

Where k_B is the Boltzmann constant, h is the Planck constant, T is the absolute temperature (298.15 K), R is the gas constant, and $\Delta G^{s\neq}$ is the activation free energy in the aqueous phase. $\Delta G^{s\neq}$ is also written as $\Delta G_f{}^{s\neq}$ and $\Delta G_r{}^{s\neq}$ to obtain the activation free energies for the forward and reverse reactions in the aqueous phase, which corresponding forward and reverse reaction rates are denoted as k_f and k_r, respectively.

Based on the Lowdin's early work,[14] there is the effect of proton tunneling in DNA on genetic information of mutations aging and tumors. In this work, quantum tunneling effect was accounted by Wigner's tunneling correction,[15] as it were successfully used for the double proton transfer rate constants in the formic acid dimmer [16] and tautomeric transition of DNA base pairs.[17,18] For that reason, conventional transition state theory (CTST) followed by Wigner tunneling correction is used to provide estimates of the rate constants in the absence and presence of catalysts, which can be calculated from eq 3.

$$k = \Gamma \frac{k_B T}{h} \exp(\frac{-\Delta G^{s\neq}}{RT}) \quad (3)$$

In eq. (3), Γ is correction factor that tend to decrease or increase the rate constant. Γ is obtained by eq 4.

$$\Gamma = 1 + \frac{1}{24}(\frac{h\nu^{\neq}}{k_B T})^2 \quad (4)$$

Where $\nu^{\neq}$ is the magnitude of the imaginary frequency associated with the vibrational mode at transition state. To test the validity of the Wigner correction, the rate constants for the tautomerism of the $Cyt2t^+$ to $CytN3^+$ as the H_2O, HCOOH catalysts by the Eckart correction have been studied (Table S1). The rate constants of Eckart correction (k_{Eckart} (298.15 K)=4.53×10^5 s^{-1}, 8.42×10^{12} s^{-1}) agree well with the corresponding Wigner correction (k_{Wigner} (298.15 K)=5.49×10^5 s^{-1}, 1.80×10^{12} s^{-1}), indicates that the Wigner correction is sufficient since it yields the same order of magnitude.

1.3 Dynamical stability of tautomers

According to literature,[19-22] the simplest estimation of the probability of the reaction process followed by the lifetime agrees satisfactorily with experimental data. The lifetime ($\tau_{99.9\%}$) needed to reach 99.9% of the equilibrium concentration of the reactant and product in the course of their

mutual tautomerisation, which can be estimated by eq 5.[23]

$$\tau_{99.99\%} = \frac{\ln 10^3}{k_f^{corr} + k_r^{corr}} \quad (5)$$

where k_f^{corr} *and* k_r^{corr} are the rate constants for the forward and reverse reaction by the Wigner correction. The lifetime τ of the product complexes is given by $1/ k_r^{corr}$.

1.4 Natural Population Analysis

The natural population analysis (NPA) charges have been used to analyze the difference of the reaction trend in the gas and aqueous phases. They were calculated at the CBS-QB3 method, 38adopting natural population analysis. [24,25]

References

(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. J. Zheng, 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. J. A. Montgomery, 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, O. Farkas, J. B. Foresman, J.V. Ortiz, J. Cioslowski,; D. J. Fox, Gaussian 09, Revision A.02. Gaussian, Inc., Wallingford, CT, 2009.

(2) J. A. Montgomery, M. J. Frisch, J. W. Ochterski and G. A. Petersson, *J. Chem. Phys*., 1999, **110**, 2822.

(3) B. Sirjean and R. Fournet, *J. Phys. Chem. A*, 2012, **116**, 6675.

(4) B. Sirjean, P. A. Glaude, M. F. Ruiz-Lòpez and R. Fournet, *J. Phys. Chem. A*, 2006, **110**, 12693.

(5) B. Sirjean, P. A. Glaude, M. F. Ruiz-Lòpez and R. Fournet, *J. Phys. Chem. A*, 2009, **113**, 6924.

(6) W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257.

(7) E. Cancès, B. Mennucci and J. Tomasi, *J. Chem. Phys*., 1997, **107**, 3032.

(8) A. Toro-Labbé, *J. Phys. Chem. A*, 1999, **103**, 4398.

(9) B. Herrera and A.Toro-Labbé, *J. Phys. Chem. A*, 2007, **111**, 5921.

(10) F. Bulat and A. Toro-Labbé, *Chem. Phys. Lett.*, 2002, **354**, 508.

(11) P. Politzer, J. S. Murray, P. Lane and A. Toro-Labbé, *Int. J. Quantum Chem.*, 2007, **107**, 2153.

(12) V. Labet, C. Morell, A. Grand and A. Toro-Labbé, *J. Phys. Chem. A*, 2008, **112**, 11487.

(13) H. Eyring, *Chem. Rev*., 1935, **17**, 65.

(14) P. O. Löwdin, Rev. *Mod. Phys*., 1963, **35**, 724.

(15) E. P. Wigner, Z. *Phys. Chem*., Abt. B, 1932, **19**, 203.

(16) A. A. Arabi and C. F. Matta, *Phys. Chem. Chem. Phys*., 2011, **13**, 13738.

(17) a) O. O. Brovarets' and D. M. Hovorun, *J. Biomol. Struct. Dyn*., 2014, **32**, 127; b) O. O. Brovarets' and D. M. Hovorun, *J. Biomol. Struct. Dyn*., 2014, **32**, 1474; c) O. O. Brovarets' and D. M. Hovorun, *J. Comput. Chem*., 2013, **34**, 2577; d) O. O. Brovarets', R. O. Zhurakivsky and D. M. Hovorun, *J. Mol. Model*., 2013, **19**, 4223; e) O. O. Brovarets', R. O. Zhurakivsky and D. M. Hovorun, *Phys. Chem. Chem. Phys*., 2014, **16**, 3715.

(18) O. O. Brovarets' and D. M. Hovorun, *Phys. Chem. Chem. Phys*., 2013, **15**, 20091.

(19) O. O. Brovarets' and D. M. Hovorun, *RSC Adv*., 2015, **5**, 99594.
(20) O. O. Brovarets' and D. M. Hovorun, *Phys. Chem. Chem. Phys*., 2015, **17**, 15103.
(21) O. O. Brovarets' and D. M. Hovorun, *J. Biomol. Struct. Dyn*., 2015, **33**, 28.
(22) O. O. Brovarets' and D. M. Hovorun, *Phys. Chem. Chem. Phys*., 2015, **17**, 21381.
(23) P. W. Atkins, *Physical chemistry*, Oxford University Press, Oxford, 1998.
(24) A. E. Reed, R. B.Weinstock and F. Weinhold, *J. Chem. Phys.*, 1985, **83**, 735.
(25) A. E. Reed, L. A. Curtiss and F. Weinhold, *Chem. Rev*., 1988, **88**, 899.

Table S1 Rate constants of Eckart correction (k_{Eckart}) and Wigner correction (k_{Wigner}) for the tautomerism of the $Cyt2t^+$ to $CytN3^+$ isomer as the H_2O, HCOOH catalysts

Paths	k_{Wigner} /s^{-1}	k_{Eckart} /s^{-1}
B-RC→B-P(path B)	5.49×10^5	4.53×10^5
C-RC→C-P(path C)	1.80×10^{12}	8.42×10^{12}

Table S2 Relative energies (in $kJ\cdot mol^{-1}$) of different protonated cytosine isomers both in the gas and aqueous phases

Species	ΔE^g /($kJ\cdot mol^{-1}$)	ΔG^g /($kJ\cdot mol^{-1}$)	ΔG^s /($kJ\cdot mol^{-1}$)
$Cyt2t^+$	0.00	0.00	0.00
$CytN3^+$	6.05	5.16	-25.79
$Cyt2c^+$	34.00	34.00	13.43
$Cyt23t^+$(a)	110.98	109.43	82.82
$Cyt23t^+$(b)	128.77	126.55	107.81
$Cyt23c^+$(a)	115.40	113.15	82.85
$Cyt23c^+$(b)	129.30	126.89	108.06
$CytN4^+$	134.30	131.07	49.47

Table S3 NPA charge (e) on N3 of $Cyt2t^+$ isomer for path C in the gas (a) and aqueous phases (b)

	a	b
ρN	-0.54	-0.600

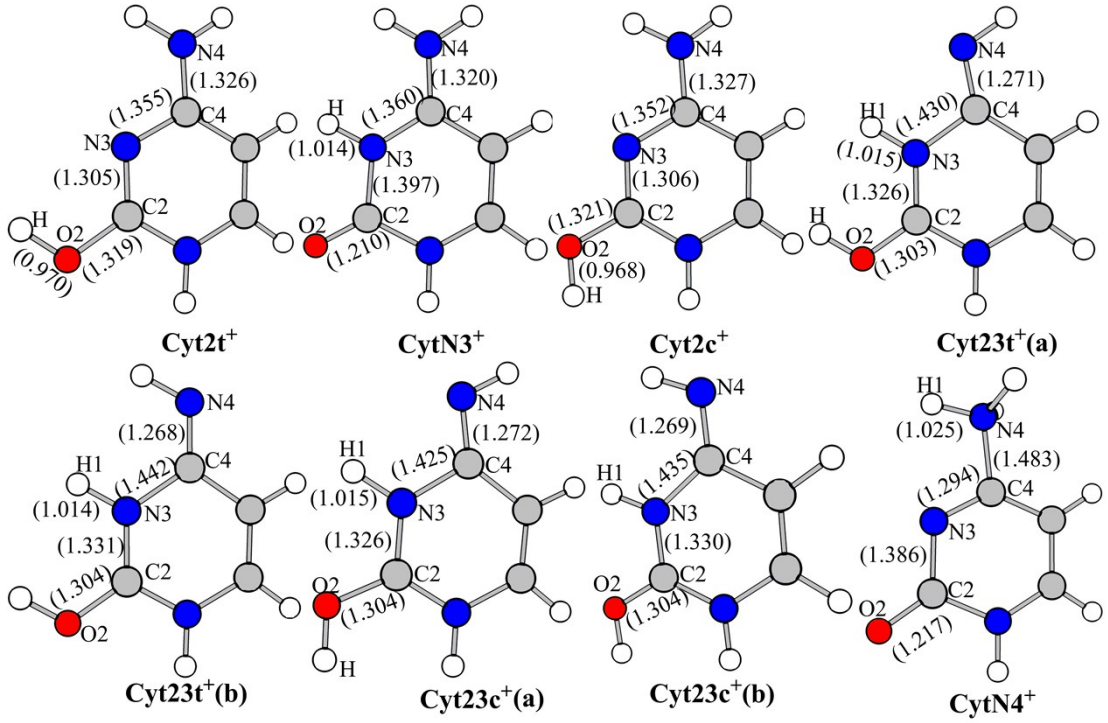

N4
(1.326)
(1.355)
C4
N3
(1.305)
C2
H
O2
(0.970)
(1.319)
Cyt2t+
N4
(1.320)
H
(1.360)
C4
(1.014)
N3
(1.397)
O2
C2
(1.210)
CytN3+
N4
(1.327)
(1.352)
C4
N3
(1.306)
(1.321)
C2
O2
(0.968)
H
Cyt2c+
N4
(1.271)
H1
(1.430)
C4
(1.015)
N3
(1.326)
H
C2
O2
(1.303)
Cyt23t+(a)
N4
(1.268)
H1
(1.442)
C4
(1.014)
N3
(1.331)
(1.304)
C2
O2
Cyt23t+(b)
N4
(1.272)
H1
(1.425)
C4
(1.015)
N3
(1.326)
O2
C2
(1.304)
H
Cyt23c+(a)
N4
(1.269)
H1
(1.435)
C4
N3
(1.330)
O2
C2
(1.304)
Cyt23c+(b)
H1
(1.025)
N4
(1.483)
(1.294)
C4
N3
(1.386)
O2
C2
(1.217)
CytN4+

Fig. S1

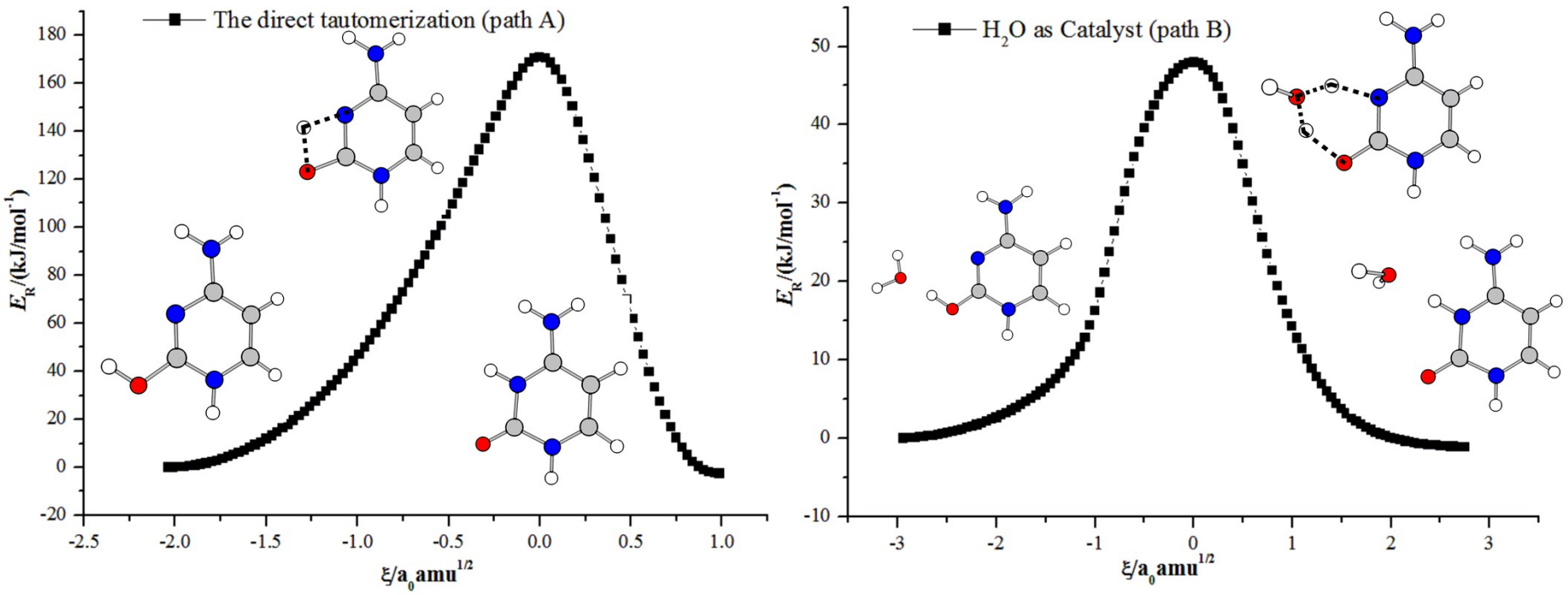
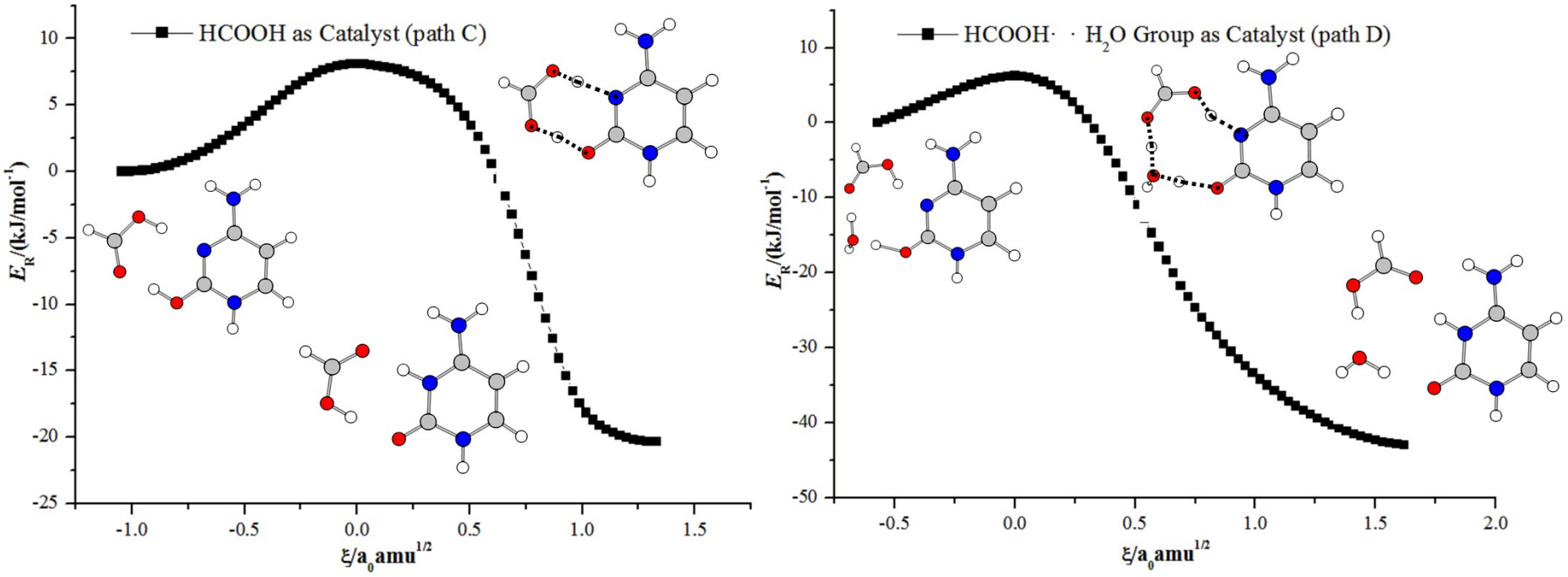
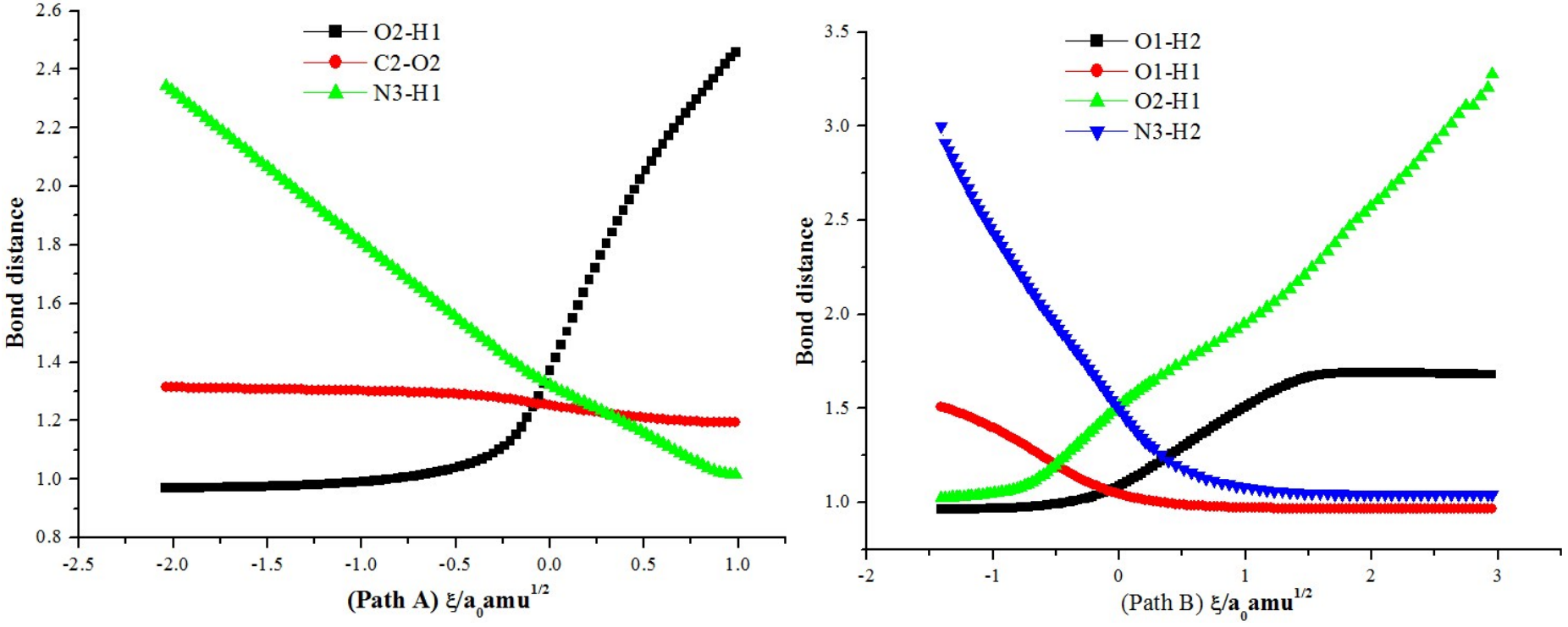

The direct tautomerization (path A)
E_R/(kJ/mol^{-1})
ξ/a$_0$amu$^{1/2}$
(a)
H$_2$O as Catalyst (path B)
E_R/(kJ/mol^{-1})
ξ/a$_0$amu$^{1/2}$
(b)

Fig. S2


HCOOH as Catalyst (path C)
E_R/(kJ/mol^{-1})
ξ/$a_0$amu$^{1/2}$
HCOOH· · H_2O Group as Catalyst (path D)
E_R/(kJ/mol^{-1})
ξ/$a_0$amu$^{1/2}$

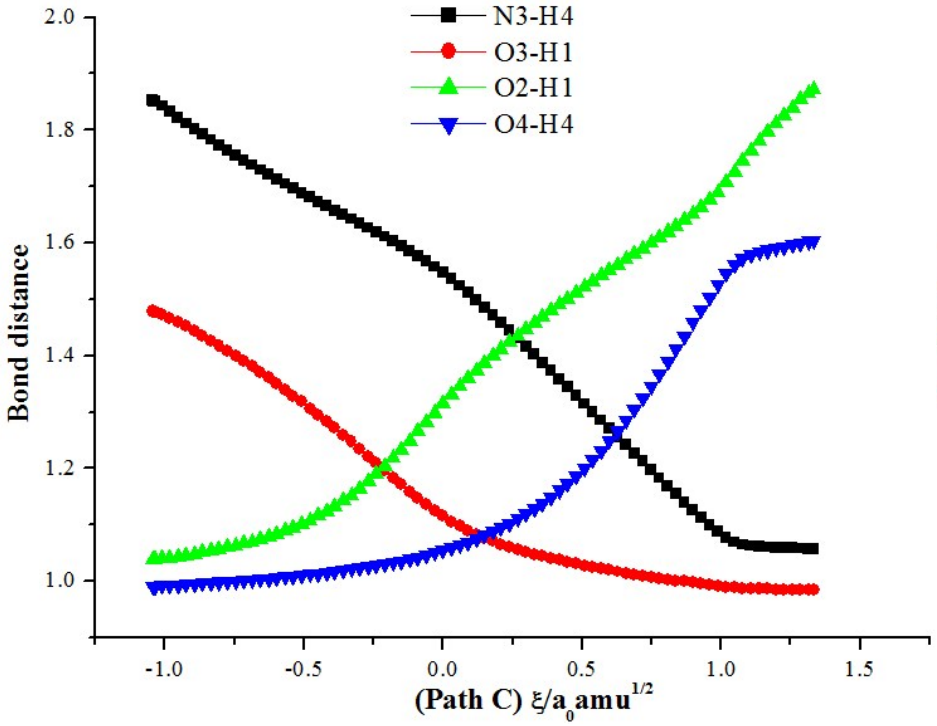
(c) (d)

Fig. S3

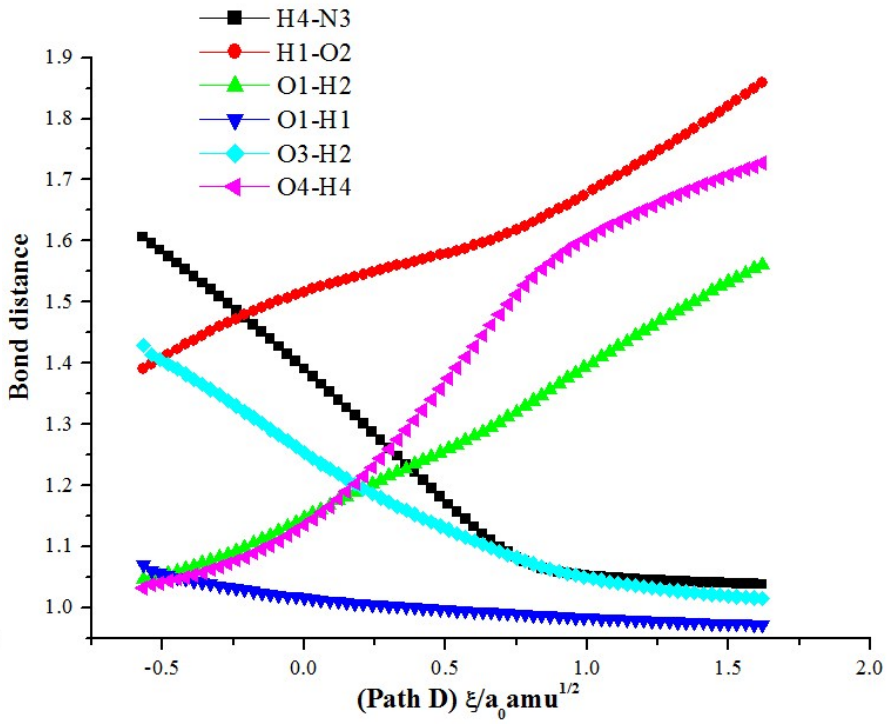

O2-H1
C2-O2
N3-H1
Bond distance
(Path A) ξ/a0amu1/2
O1-H2
O1-H1
O2-H1
N3-H2
Bond distance
(Path B) ξ/a0amu1/2

(a) (b)

Fig. S4


N3-H4
O3-H1
O2-H1
O4-H4
Bond distance
(Path C) ξ/a0amu1/2

(c)


H4-N3
H1-O2
O1-H2
O1-H1
O3-H2
O4-H4
Bond distance
(Path D) ξ/a0amu1/2

(d)

Fig. S5

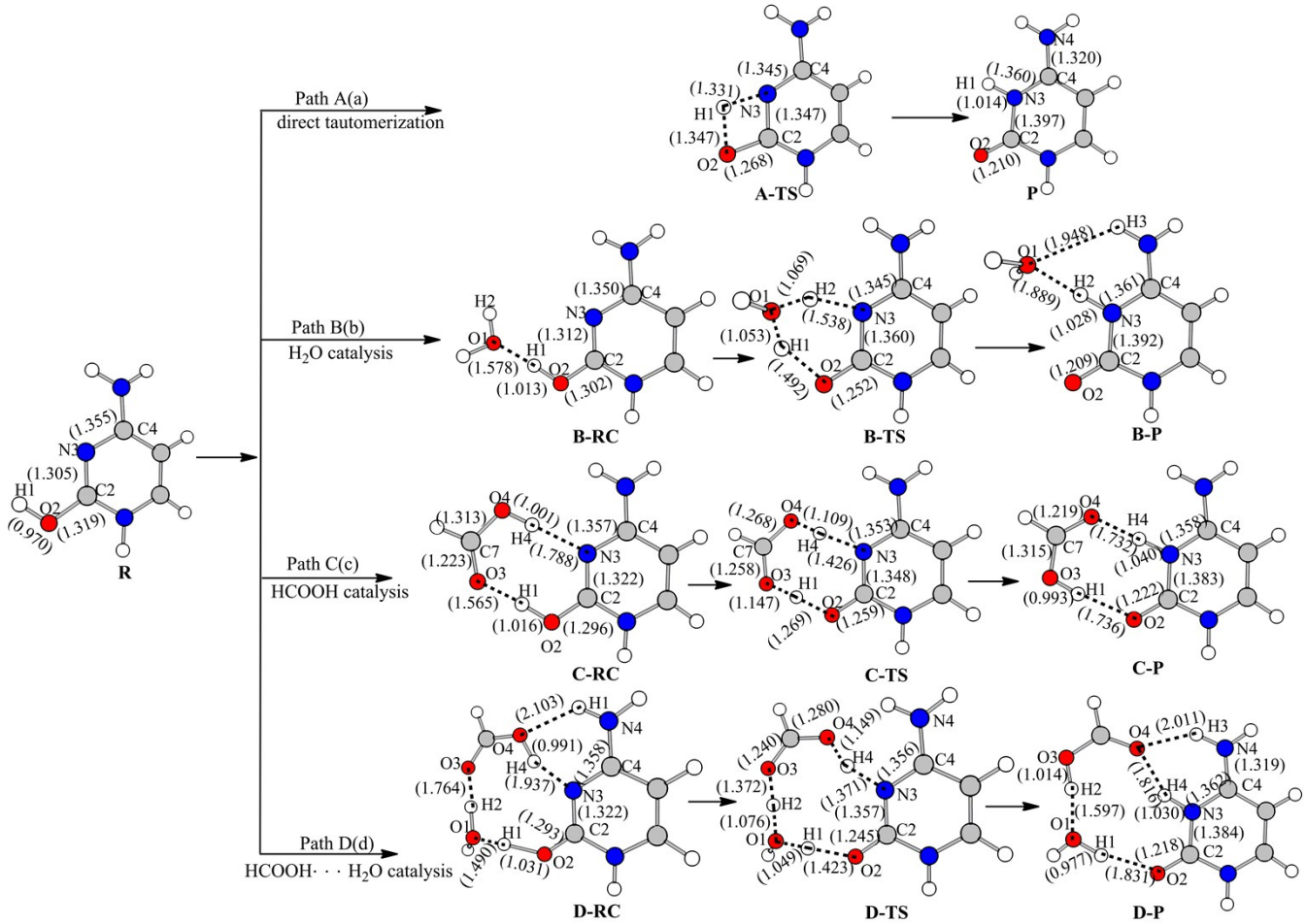

Path A(a)
direct tautomerization
Path B(b)
H2O catalysis
Path C(c)
HCOOH catalysis
Path D(d)
HCOOH· · · H2O catalysis
R
A-TS
P
B-RC
B-TS
B-P
C-RC
C-TS
C-P
D-RC
D-TS
D-P

Fig. S6