Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Characterization of the Reaction Path and Transition States for RNA Transphosphorylation Models from Theory and Experiment**

Kin-Yiu Wong, Hong Gu, Shuming Zhang, Joseph A. Piccirilli,* Michael E. Harris,* and Darrin M. York**

anie_201104147_sm_miscellaneous_information.pdf

Supporting Information

Contents

Experimental Section

Up[NPO-^{18}O]G was synthesized by including $H_2{}^{18}O$ in the hydrolysis step during solid phase synthesis. Kinetic isotope effects were measured essentially as described in Ref. [1]. The 2′-*O*-transphosphorylation of 1-5 nmoles of UpG at pH 11-14 was carried out at 37°C at an ionic strength of 1 M. The reactant (UpG) and products (cUMP and G) were resolved by RP-HPLC on a 300 × 50 mm C18 column run isocratically at 1 mL/min of 0.1 M ammonium acetate and 4% acetonitrile. The fraction of reaction (*F* = [cUMP or G] / ([UpG] + [cUMP or G]) was quantified by integration of the chromatogram. The HPLC fractions containing the unreacted UpG were collected and dried under vacuum and resuspended in water at a concentration of 20-50 μM for subsequent isotopic analysis by MS.

Determination of the substrate $^{16}O/^{18}O$ ratio in UpG isolated by RP-HPLC was accomplished by tandem MS using an ABI QStar electrospray ionization quadrupole time-of-flight tandem mass spectrometer in negative ion mode as described in Ref. [1]. The entire deprotonated molecular ion isotopic cluster of UpG with monoisotopic *m/z* 588 was isolated and fragmented by inert gas collision. The resulting fragments at 100-600 *m/z* were resolved by TOF MS. Isotope ratios were determined from 476/478 and 211/213 *m/z* fragments. The KIEs reported in Table 2 were calculated from plots of ln($^{18}O/^{16}O$) versus F by fitting to:

$$\ln\left({}^{18}\mathrm{O}/{}^{16}\mathrm{O}\right) = \ln\left({}^{18}k - 1\right)/\left(1 - F\right) - \ln\left(I\right) \qquad (1)$$

where ^{18}k is the isotope effect, *F* is the fraction of substrate consumed as determined from integration of the HPLC chromatogram, and *I* is the initial $^{18}O/^{16}O$ ratio in the unreacted substrate.[1] Each experiment was performed in duplicate and the intensity data for both the 211/213 and 476/478 ions were analyzed yielding four determinations of the KIE for each experiment. The KIEs and errors reported in Table 2 are averages and standard deviations of six to ten different determinations coming from at least three independent experiments.

Computational Section

The molecular structures of the native and thio-substituted ionic reactants, and the transition-state structures in solution were minimized using density-functional theory (DFT) with the inclusion of the PCM model[2] for treating the solvent effects. The hybrid B3LYP[3] exchange-correlation functional was used together with the 6-31+G(d) basis set. Vibrational frequency analyses were performed to confirm the nature of the minimum and saddle points, and to compute quantum thermal corrections in the decoupled rigid-rotor harmonic-oscillator approximation. The software package Gaussian 09[4] was used for all the electronic structure calculations. An essential feature that enabled these calculations to be performed was the implementation of a continuous surface charge formalism[5] based on a smooth boundary element solvation method[6] that afforded continuity, smoothness, and robustness of the solvent reaction field and its derivatives with respect to nuclear positions. We have used the UAKS radii optimized at the PB0/6-31G(d) level of theory[7, 8] to define the molecular cavity for the PCM model, with special tuning of the O2′ and X5′ positions so as to be fixed along the reaction coordinate and close to the QM/MM and experimental barrier values. All atomic radii used in the present work and the PCM minimized structures are given below as part of the Supporting Information.

We used the following equation to determine the primary KIE values in this work:[9]

$$\mathrm{KIE} = \left(\frac{\omega^{*}_{l_0}}{\omega^{*}_{h_0}} \right) \frac{\exp\left[-\beta \left(W^{‡}_{l_0} - W_{h_0} \right) \right]}{\exp\left[-\beta \left(W^{\mathrm{R}}_{l_0} - W^{\mathrm{R}}_{h_0} \right) \right]}, \qquad (2)$$

where $\beta = 1/k_B T$ is the Boltzmann-weighted inverse temperature, the superscript ‡ and *R* denote the transition state and the reactant state, respectively, ω^{*} is the imaginary frequency at the transition state, l_0 indicates the light isotope and h_0 the heavy isotope, *W* is the centroid effective potential energy[10-12] calculated from the centroid position of path integrals. Note that Eq. (2) reduces to the widely-used Bigeleisen equation[13, 14] when the centroid potential is computed in the decoupled rigid-rotor harmonic-oscillator approximation (and neglecting tunneling effects; Appendix B in Ref. [9] provides a proof). Moreover the mass (isotope) and temperature dependent nature of the centroid potential energy distinguishes it from the *ab initio* Born-Oppenheimer potential energy, which is mass and temperature independent. With the treatment of solvent effects by a dielectric continuum, we determined the values of *W* in Eq. (2) using our automated integration-free path-integral (AIF-PI) method, which is based on Kleinert's variational perturbation (KP) theory, and in the decoupled instantaneous normal coordinate approximation (DINCA).[15-18] Our previous studies on a series of proton transfer reactions indicate that performing *ab initio* path-integral

calculations with our AIF-PI method in Eq. (2) can accurately and economically include anharmonicity and tunneling contributions to the calculated KIE values.[18] In this work, the entire system is quantized (more than 16 atoms) to compute harmonic KIE values. We further quantize the following six to seven nuclei to estimate the anharmonic KIE values using the AIF-PI method: the phosphorus atom and all atoms covalently bonded to it, including the O2′ of nucleophile and X5′ of leaving groups, in addition to the hydrogen connected with the O2′ for the case of neutral ethylene phosphate with methanol group. All our path-integral computations were performed using the AIF-PI program implemented for use with *Mathematica*.[19]

The computationally expensive density-functional QM/MM free-energy simulations for the native reaction with the isothermal-isobaric ensembles at 25°C and 1 atm were performed on the fly (i.e., the B3LYP 6-31+G(d) potential was computed at every time step) using a modified version of the CHARMM program (based on version c36a2)[20] interfacing with Q-Chem.[21] The entire solute was in the QM region, and was solvated by MM TIP3P water molecules[22] in a periodic cubic box. B3LYP QM/MM electrostatic interactions were calculated in the primary box, whereas all other periodic long-range electrostatic interactions were treated by a linear-scaling QM/MM-Ewald method at the semiempirical AM1 level.[23]

Bond Distances and Bond Orders of Transition States:

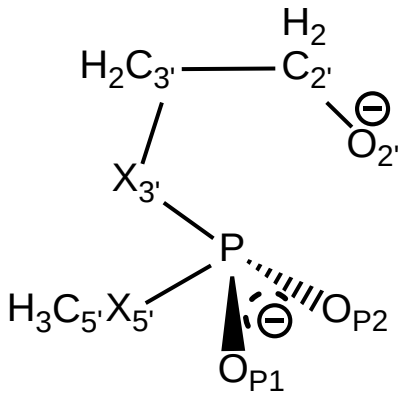
Table S1.

Bond lengths and bond orders for key phosphoryl positions in the rate-limiting transition state minimized in aqueous solution.[a]

Reaction	TS	$r_{PX2'}$	$r_{PX5'}$	$B_{PX2'}$	$B_{PX5'}$	$\sum B_{P2nd}$
Native	TS_1	2.127	1.769	0.294	0.501	2.684
Native	TS_2	1.766	2.306	0.518	0.207	2.762
S3′	TS_1	2.301	1.747	0.223	0.518	2.895
S3′	TS_2	1.755	2.298	0.528	0.202	2.897
S5′	TS_1	2.306	2.290	0.220	0.659	2.765

[a] Bond lengths r are in units of Å and bond order B is from the Wiberg bond index. $\sum B_{P2nd}$ is the summation of the bond orders over all three secondary oxygens (O_{P1}, O_{P2}, and O3′). Note: the calculated bond length and Wiberg bond order for the P—O single bond in ethylene phosphate are 1.664 Å and 0.629, respectively.

Atomic Radii Used for Building up Molecular Cavities in PCM Calculations:



Atomic numberings in a model for base-catalyzed RNA phosphate transesterification, in which X is either oxygen or sulfur.

Center for spherical cavity	Radius of cavity (Å)	Scaling factor
P	1.913	1.1
O_{P1}	1.568	1.1
O_{P2}	1.568	1.1
$C_{2'}$	1.703	1.1
$C_{3'}$	1.703	1.1
$C_{5'}$ when $X_{5'} = O_{5'}$	1.883	1.1
$C_{5'}$ when $X_{5'} = S_{5'}$	1.973	1.1
$O_{2'}$	1.485	1.1
$O_{3'}$	1.568	1.1
$O_{5'}$	1.485	1.1
$S_{3'}$	2.108	1.1
$S_{5'}$	1.990	1.1

Native Reactant Structure at the PCM-B3LYP/6-31+G(d) level:

C	-2.633113	1.086129	-1.400144
O	-1.438750	0.890879	-0.620336
P	-1.169278	-0.589396	0.074437
O	-1.205625	-1.661915	-1.001450
O	0.343608	-0.328697	0.618979
C	1.479067	-0.287291	-0.300413
C	2.708688	0.247811	0.448619
O	3.832592	0.282882	-0.373706
O	-2.051746	-0.745511	1.300549
H	1.669043	-1.312903	-0.674219
H	1.232495	0.372654	-1.156265
H	2.431788	1.259409	0.849287
H	2.850755	-0.405644	1.350778
H	-2.611103	2.123718	-1.751527
H	-2.650128	0.406207	-2.262691
H	-3.532000	0.928693	-0.786721

Native Transition-State 1 Structure at the PCM-B3LYP/6-31+G(d) level:

P	-0.296738	-0.526840	-0.045494
O	0.248936	1.016666	-0.448984
C	1.650248	1.345175	-0.466755
C	2.357923	0.372304	0.467854
H	1.741771	2.406757	-0.160954
H	2.023066	1.228358	-1.506496
H	2.199291	0.693509	1.530168
O	1.786809	-0.874894	0.203372
O	-0.512385	-0.804906	1.455352
O	-0.289964	-1.623413	-1.125659
O	-1.925824	0.031695	-0.449720
C	-2.545975	1.050623	0.322001
H	-2.029526	2.017673	0.213386
H	-2.575993	0.788516	1.390970
H	3.455698	0.392203	0.276064
H	-3.578065	1.164872	-0.039881

Native Intermediate Structure at the PCM-B3LYP/6-31+G(d) level:

P	-0.190611	-0.487173	-0.037149
O	0.176916	1.123947	-0.465903
C	1.540369	1.572636	-0.408145
C	2.288893	0.561823	0.453758
H	1.546638	2.599969	0.008432
H	1.950260	1.593153	-1.440318
H	2.136801	0.789369	1.537461
O	1.729208	-0.679582	0.100618
O	-0.405709	-0.765772	1.471240
O	-0.225556	-1.597805	-1.114527
O	-1.905424	-0.014707	-0.402475
C	-2.562586	0.950090	0.399512
H	-2.103953	1.948741	0.307762
H	-2.558800	0.667157	1.464736
H	3.379702	0.586212	0.246459
H	-3.608990	1.020582	0.064371

Native Transition-State 2 Structure at the PCM-B3LYP/6-31+G(d) level:

P	0.000794	-0.611563	-0.001465
O	0.176584	1.017048	-0.373111
C	1.547160	1.473592	-0.446032
C	2.340515	0.512768	0.427624
H	1.574780	2.522231	-0.091666
H	1.883482	1.425783	-1.503311
H	2.235429	0.773959	1.505384
O	1.754034	-0.759013	0.156229
O	-0.429845	-0.909581	1.437520
O	-0.190988	-1.603016	-1.150445
O	-2.171529	0.003603	-0.468323
C	-2.751216	1.012185	0.299285
H	-2.330631	2.020108	0.086848
H	-2.621853	0.848856	1.392181
H	3.415903	0.489862	0.162342
H	-3.846035	1.089618	0.125495

S3′ Reactant Structure at the PCM-B3LYP/6-31+G(d) level:

C	1.988183	2.080223	-0.076076
O	1.595520	0.990262	0.781378
P	1.337450	-0.518705	0.151685
O	1.120969	-1.398577	1.369331
S	-0.459282	-0.341334	-0.990230
C	-1.750322	0.005701	0.306216
C	-3.149784	0.120356	-0.340819
O	-4.134311	0.355176	0.612416
O	2.419302	-0.851878	-0.862831
H	-1.737422	-0.817352	1.035484
H	-1.489610	0.945083	0.816992
H	-3.092164	0.934678	-1.111411
H	-3.324306	-0.828007	-0.916022
H	2.076321	2.963037	0.566730
H	2.954126	1.868956	-0.552447
H	1.225908	2.260906	-0.846732

S3′ Transition-State 1 Structure at the PCM-B3LYP/6-31+G(d) level:

P	-0.523971	-0.592268	-0.129528
S	0.452064	1.359837	-0.388975
C	2.243421	1.067783	-0.071397
C	2.408010	-0.273016	0.662796
H	2.621668	1.917249	0.520391
H	2.766163	1.041738	-1.040081
H	2.092675	-0.129001	1.730200
O	1.672565	-1.260827	0.025726
O	-0.719496	-1.079927	1.311963
O	-0.596187	-1.528342	-1.341833
O	-2.034952	0.233313	-0.427019
C	-2.700862	0.940840	0.617473
H	-2.050593	1.699959	1.077863
H	-3.053973	0.257101	1.402697
H	3.502031	-0.505442	0.681638
H	-3.565816	1.446445	0.166684

S3′ Intermediate Structure at the PCM-B3LYP/6-31+G(d) level:

P	-0.363971	-0.623360	-0.090350
S	0.446550	1.444846	-0.392220
C	2.225115	1.149661	-0.038075
C	2.291455	-0.208050	0.673596
H	2.615732	1.962564	0.593640
H	2.789960	1.125065	-0.982950
H	1.989768	-0.083907	1.740706
O	1.452436	-1.121705	0.005999
O	-0.688167	-1.083950	1.351431
O	-0.582094	-1.530963	-1.324830
O	-1.964932	0.252853	-0.378796
C	-2.605546	0.936350	0.682802
H	-1.924632	1.625954	1.208574
H	-3.023608	0.240215	1.428401
H	3.333260	-0.592669	0.652329
H	-3.431958	1.529347	0.259902

S3′ Transition-State 2 Structure at the PCM-B3LYP/6-31+G(d) level:

P	-0.140045	-0.766342	-0.073662
S	0.277325	1.349658	-0.377342
C	2.104542	1.252274	-0.122927
C	2.356871	-0.052280	0.634773
H	2.438546	2.128616	0.453713
H	2.612654	1.244597	-1.099445
H	2.081381	0.065291	1.706885
O	1.583090	-1.085485	0.027511
O	-0.598825	-1.169086	1.332419
O	-0.442135	-1.628600	-1.302344
O	-2.202167	0.173398	-0.456632
C	-2.829093	0.932610	0.532805
H	-2.630874	2.023282	0.439188
H	-2.506697	0.643830	1.554866
H	3.423264	-0.347798	0.569312
H	-3.934058	0.815191	0.509127

S5′ Reactant Structure at the PCM-B3LYP/6-31+G(d) level:

C	-1.779542	2.094420	0.527947
S	-1.643000	0.891828	-0.858556
P	-0.855410	-0.857369	0.116041
O	-0.758780	-1.885371	-0.999108
O	0.631960	-0.414808	0.644914
C	1.737287	-0.200369	-0.284334
C	2.930733	0.391193	0.481771
O	4.021138	0.597393	-0.359536
O	-1.653628	-1.154053	1.373290
H	2.011931	-1.171997	-0.740241
H	1.409092	0.496459	-1.082665
H	2.570931	1.337435	0.967462
H	3.159556	-0.313484	1.325561
H	-2.200315	3.012730	0.104268
H	-2.448017	1.713410	1.305758
H	-0.794114	2.310814	0.951995

S5′ Transition-State 1 Structure at the PCM-B3LYP/6-31+G(d) level:

P	-0.007881	-0.286109	-0.007826
O	-0.862822	1.138190	-0.115833
C	-2.271991	1.231954	0.214184
C	-2.945626	-0.069860	-0.195791
H	-2.652671	2.123459	-0.322717
H	-2.359577	1.391693	1.309357
H	-3.047438	-0.086493	-1.313762
O	-2.144449	-1.103825	0.280169
O	0.075773	-1.097352	-1.303208
O	0.298955	-0.842250	1.385158
S	1.956283	0.882030	-0.152873
C	1.943062	1.648521	-1.824348
H	1.137516	2.385699	-1.913612
H	1.819080	0.883911	-2.598966
H	-3.982695	-0.082767	0.222382
H	2.903833	2.153850	-1.975962

References

[1] M. E. Harris, Q. Dai, H. Gu, D. L. Kellerman, J. A. Piccirilli, V. E. Anderson, *J. Am. Chem. Soc.* **2010**, *132*, 11613.
[2] M. Cossi, G. Scalmani, N. Rega, V. Barone, *J. Chem. Phys.* **2002**, *117*, 43.
[3] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.
[4] Gaussian 09, Revision A.1, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.
[5] G. Scalmani, M. J. Frisch, *J. Chem. Phys.* **2010**, *132*, 114110.
[6] D. M. York, M. Karplus, *J. Phys. Chem. A* **1999**, *103*, 11060.
[7] V. Barone, M. Cossi, J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 3210.
[8] Y. Takano, K. N. Houk, *J. Chem. Theory Comput.* **2005**, *1*, 71.
[9] K.-Y. Wong, Ph.D. thesis thesis, University of Minnesota (USA) (Minneapolis), **2008**.
[10] M. J. Gillan, *Phys. Rev. Lett.* **1987**, *58*, 563.
[11] J. Cao, G. A. Voth, *J. Chem. Phys.* **1994**, *101*, 6168.
[12] G. A. Voth, *Adv. Chem. Phys.* **1996**, *93*, 135.
[13] L. J. Schaad, L. Bytautas, K. N. Houk, *Can. J. Chem.* **1999**, *77*, 875.
[14] M. Wolfsberg, in *Isotope Effects in Chemistry and Biology* (Eds.: A. Kohen, H.-H. Limbach), Taylor & Francis, Boca Raton, **2006**, pp. 89.
[15] H. Kleinert, *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*, 3rd ed., World Scientific, Singapore; River Edge, NJ, **2004**.
[16] K.-Y. Wong, J. Gao, *J. Chem. Phys.* **2007**, *127*, 211103.
[17] K.-Y. Wong, J. Gao, *J. Chem. Theory Comput.* **2008**, *4*, 1409.
[18] K.-Y. Wong, J. P. Richard, J. Gao, *J. Am. Chem. Soc.* **2009**, *131*, 13963.
[19] Wolfram Research, Inc., Mathematica, Versions 5 and 6, Champaign, IL.
[20] B. R. Brooks, C. L. Brooks, III, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, *J. Comput. Chem.* **2009**, *30*, 1545.
[21] Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio, Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock, III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer, III, J. Kong, A. I. Krylov, P. M. W. Gill, M. Head-Gordon, *Physical Chemistry Chemical Physics* **2006**, *8*, 3172.
[22] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
[23] K. Nam, J. Gao, D. M. York, *J. Chem. Theory Comput.* **2005**, *1*, 2.