Linear Free Energy Relationships in RNA Transesterification: Theoretical Models to Aid Experimental Interpretations

Supporting Information

Ming Huang [*][†] and Darrin M. York [†][‡]

May 14, 2014

UAKS Solute Radii for Intermediates

The UAKS original radii ($R_{O_p}^I$) and adjusted radii ($\tilde{R}_{O_p}^I$) for oxygen atoms O_p connected to phosphorus in the intermediates (I) are listed in Table S1, compared with those ($R_{O_{Lg}}^{TS}$ and $R_{O_{Nu}}^{TS}$) of oxygen atoms O_{Nu} and O_{Lg} in the nucleophile and leaving groups of transition states. During the calculation, we note that the original UAKS solute radii $R_{O_p}^I$ in the intermediates have the same values of more charged O_{Nu} in TS1 and O_{Lg} in TS2, i.e, $R_{O_p}^I = \text{MIN}(R_{O_{Lg}}^{TS}, R_{O_{Nu}}^{TS})$, where TS is TS1 or TS2. In order to make these radii consistent, we used an averaging strategy[1] whereby the radii of oxygen atoms connected to phosphorus in the intermediates were set to the average radii

---

[*]Scientific Computation, University of Minnesota, 207 Pleasant St. SE, Minneapolis, MN 55455–0431, USA.
[†]Center for Integrative Proteomics Research, BioMaPS Institute and Department of Chemistry & Chemical Biology, Rutgers University, 174 Frelinghuysen Road, Piscataway, NJ 08854–8076, USA.
[‡]Electronic mail: york@biomaps.rutgers.edu

	I		TS1		TS2	
Leaving Group	R_{O_p}	$\tilde{R}_{O_p}$	$R_{O_{Nu}}$	$R_{O_{Lg}}$	$R_{O_{Nu}}$	$R_{O_{Lg}}$
$(CH_3)_2CHO^-$	1.445	1.513	1.445	1.580	1.580	1.445
$CH_3CH_2CH_2O^-$	1.445	1.513	1.445	1.580	1.580	1.445
$CH_3CH_2O^-$	1.440	1.507	1.440	1.575	1.575	1.440
CH_3O^-	1.433	1.501	1.433	1.568	1.568	1.433
$HOCH_2CH_2O^-$	1.445	1.513	1.445	1.580	1.580	1.445
$ClCH_2CH_2O^-$	1.445	1.513	1.445	1.580	1.580	1.445
$FCH_2CH_2O^-$	1.445	1.513	1.445	1.580	1.580	1.445
$CHCCH2O^-$	1.445	1.513	1.445	1.580	1.580	1.445
$CF_3CH_2O^-$	1.454	1.522	1.454	1.589	1.589	1.454

Table S1: Original and adjusted UAKS solute radii ($R_{O_p}^I$ and $\tilde{R}_{O_p}^I$, Å) of oxygen atoms (O_p) connected to phosphorus in the intermediates (I), compared with those ($R_{O_{Lg}}^{TS}$ and $R_{O_{Nu}}^{TS}$) of O_{Nu} and O_{Lg} in transition states (TS1 and TS2). O_{Nu} and O_{Lg} are oxygen atoms in the nucleophile and leaving group, respectively.

of oxygen atoms of leaving group and nucleophile in transition states, i.e, $\tilde{R}_{O_p}^I = \frac{1}{2}(R_{O_{Lg}}^{TS} + R_{O_{Nu}}^{TS})$.

Free Energies and Brønsted Correlations

	UFF				UAKS				Expt.
Leaving Group	$\Delta G_1^{\neq}$	ΔG_i	$\Delta G_2^{\neq}$	ΔG	$\Delta G_1^{\neq}$	ΔG_i	$\Delta G_2^{\neq}$	ΔG	pK_a
$(CH_3)_2CHO^-$	20.63	20.62	26.36	-3.65	20.40	18.09	24.15	-12.94	17.1
$CH_3CH_2CH_2O^-$	19.26	19.46	24.52	-3.45	18.84	15.80	23.00	-11.45	16.1
$CH_3CH_2O^-$	20.76	20.73	26.37	-1.87	20.20	17.24	23.56	-9.37	16
CH_3O^-	19.33	19.50	24.75	-2.53	19.45	16.35	22.67	-8.37	15.54
$HOCH_2CH_2O^-$	18.71	17.05	20.10	-8.37	19.49	17.59	23.52	-14.62	15.07
$ClCH_2CH_2O^-$	17.57	16.45	18.92	-8.89	18.57	16.07	20.67	-21.04	14.31
$FCH_2CH_2O^-$	18.13	16.93	20.51	-7.48	20.64	16.29	21.75	-19.02	14.2
$CHCCH2O^-$	17.30	15.78	18.23	-9.29	17.69	14.36	18.11	-18.50	13.55
$CF_3CH_2O^-$	16.37	12.95	14.00	-13.16	17.14	12.45	15.66	-22.81	12.37
$CH_3CH_2S^-$	14.64	$\cdots$	$\cdots$	-35.14	16.46	$\cdots$	$\cdots$	-31.89	10.61
CH_3S^-	15.32	$\cdots$	$\cdots$	-34.42	15.94	$\cdots$	$\cdots$	-31.30	10.33
PhO^-	15.14	$\cdots$	$\cdots$	-25.07	16.24	$\cdots$	$\cdots$	-25.42	9.95
$3\text{-}CNPhO^-$	14.42	$\cdots$	$\cdots$	-28.81	14.53	$\cdots$	$\cdots$	-30.15	8.61
$4\text{-}CNPhO^-$	13.03	$\cdots$	$\cdots$	-32.53	14.55	$\cdots$	$\cdots$	-31.60	7.95
$p\text{-}NO_2PhO^-$	12.41	$\cdots$	$\cdots$	-35.02	14.50	$\cdots$	$\cdots$	-32.68	7.14
$2,3,5,6\text{-}F_4PhO^-$	11.38	$\cdots$	$\cdots$	-38.08	11.39	$\cdots$	$\cdots$	-39.24	5.53
CH_3COO^-	12.31	$\cdots$	$\cdots$	-34.64	11.58	$\cdots$	$\cdots$	-36.70	4.46

Table S2: Calculated free energies (kcal/mol) at 80° for transition states ($\Delta G^{\neq}$), intermediates (ΔG_i) and products (ΔG) relative to reactants in RNA transesterification model reactions with different leaving groups as well as their experimental pK_as in water at 25°C. $\Delta G_1^{\neq}$ and $\Delta G_2^{\neq}$ are free energy barriers of early and late transition states (TS1 and TS2), respectively. All the experimental pK_as are taken from IUPAC chemical data series (No. 23),[2] except those of ethylene glycol and 2,3,5,6-tetrafluorophenol, which are from CRC Handbook[3] and Bourne $et\ al.$,[4] respectively. The pK_as of ethylene glycol and acetic acid have been corrected for statistical factors by adding log(p/q), where p and q are the number of reactive positions available in the acid and in the base, respectively.[5,6]
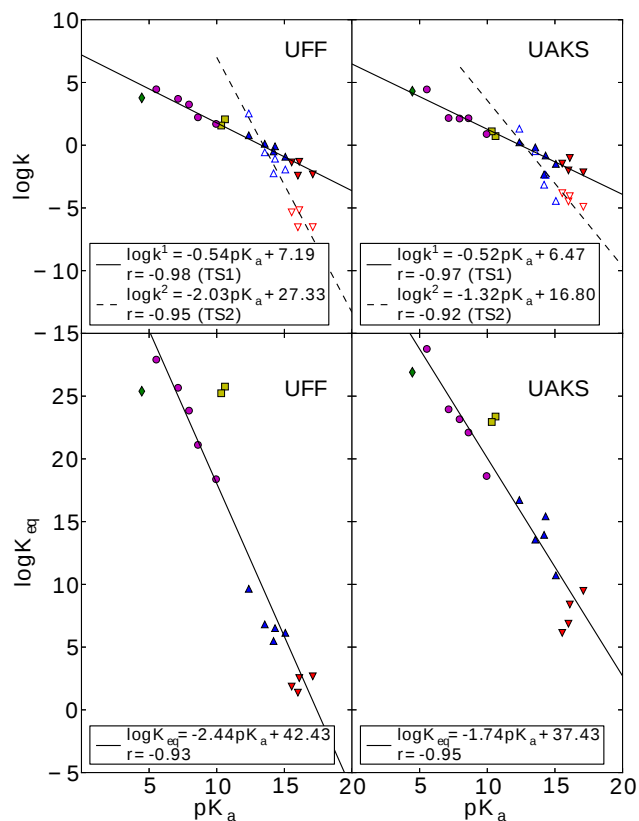
Figure S1: Linear free energy relationships between calculated logarithmic rate constants ($\log k^1$ and $\log k^2$) at 80°C and experimental pK_as at 25°C of leaving groups (top) and between calculated logarithmic equilibrium constants ($\log K_{eq}$) and experimental pK_as of leaving groups (bottom). k^1 and k^2 are rate constants for early and late transition states (TS1 and TS2), respectively. Rate constants are obtained from density-functional calculations with PCM solvation model and UFF (left) and UAKS (right) radii. Red down triangle, blue up triangle, yellow square, purple round and green diamond symbols correspond to alkyl, hetero-alkyl, thio, phenyl and acid leaving groups, respectively. Filled and empty symbols stand for TS1 and TS2, respectively. Regression parameters for R are given as well as linear correlation coefficient for r.

Relationship of Bond length and Bond order

Bond order analysis for bonds in the transition states of model reactions has been done using the NBO program[7,8] of the Gaussian09 in the M06-2X/6-31++G(d,p) level of theory with PCM implicit solvation and UFF and UAKS radii. The bond lengths (R) and Wiberg bond indices (N) of the forming bonds (P-Nu) and breaking bonds (P-Lg) in the transition states are given in the Table S3 and S4 for UFF and UAKS models, respectively. The data in the tables show that UFF and UAKS give almost the same bond lengths and bond orders for forming bonds and breaking bonds in the transition states. The maximal difference of bond length and bond order between the two radii models are very small, 0.10 and 0.04 Å, respectively, so data obtained with both UFF and UAKS models are combined for bond order analysis.

Leaving Group	TS1				TS2			
	R_{P-Nu}	R_{P-Lg}	N_{P-Nu}	N_{P-Lg}	R_{P-Nu}	R_{P-Lg}	N_{P-Nu}	N_{P-Lg}
$(CH_3)_2CHO^-$	2.29	1.72	0.31	0.82	1.73	2.42	0.82	0.21
$CH_3CH_2CH_2O^-$	2.30	1.72	0.30	0.82	1.73	2.43	0.82	0.21
$CH_3CH_2O^-$	2.29	1.72	0.31	0.81	1.73	2.45	0.82	0.20
CH_3O^-	2.30	1.72	0.30	0.81	1.72	2.46	0.83	0.19
$HOCH_2CH_2O^-$	2.32	1.73	0.29	0.78	1.73	2.37	0.82	0.22
$ClCH_2CH_2O^-$	2.33	1.73	0.29	0.78	1.73	2.38	0.82	0.22
$FCH_2CH_2O^-$	2.35	1.72	0.27	0.80	1.73	2.39	0.82	0.22
$CHCCH2O^-$	2.37	1.72	0.26	0.79	1.73	2.39	0.82	0.22
$CF_3CH_2O^-$	2.36	1.74	0.27	0.76	1.73	2.35	0.82	0.24
PhO^-	2.38	1.76	0.26	0.72	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$3\text{-}CNPhO^-$	2.40	1.77	0.25	0.70	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$4\text{-}CNPhO^-$	2.43	1.78	0.24	0.69	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$p\text{-}NO_2PhO^-$	2.45	1.78	0.23	0.68	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$2,3,5,6\text{-}F_4PhO^-$	2.45	1.80	0.23	0.66	$\cdots$	$\cdots$	$\cdots$	$\cdots$
CH_3COO^-	2.46	1.79	0.22	0.67	$\cdots$	$\cdots$	$\cdots$	$\cdots$

Table S3: The bond lengths, R_{P-Nu} and R_{P-Lg} (Å) and Wiberg bond indices, N_{P-Nu} and N_{P-Lg} of forming bonds (P-Nu) and breaking bonds (P-Lg) in the transition states of model reactions obtained from density-functional calculations with PCM solvation and UFF radii.

The bond length and Wiberg bond index of P-O bond in dimethyl phosphate anion were also calculated with the same method, which gives the reference as fully formed single P-O bond. The

| | TS1 | | | | TS2 | | | |
Leaving Group	R_{P-Nu}	R_{P-Lg}	N_{P-Nu}	N_{P-Lg}	R_{P-Nu}	R_{P-Lg}	N_{P-Nu}	N_{P-Lg}
$(CH_3)_2CHO^-$	2.22	1.72	0.35	0.83	1.72	2.33	0.83	0.26
$CH_3CH_2CH_2O^-$	2.22	1.72	0.35	0.82	1.73	2.33	0.83	0.27
$CH_3CH_2O^-$	2.24	1.72	0.34	0.83	1.72	2.36	0.84	0.25
CH_3O^-	2.24	1.72	0.34	0.82	1.72	2.36	0.84	0.25
$HOCH_2CH_2O^-$	2.23	1.72	0.35	0.82	1.72	2.34	0.84	0.25
$ClCH_2CH_2O^-$	2.28	1.72	0.32	0.81	1.73	2.29	0.83	0.28
$FCH_2CH_2O^-$	2.27	1.72	0.32	0.81	1.72	2.31	0.83	0.27
$CHCCH2O^-$	2.29	1.72	0.31	0.80	1.72	2.32	0.84	0.26
$CF_3CH_2O^-$	2.30	1.73	0.31	0.78	1.72	2.27	0.83	0.28
PhO^-	2.30	1.76	0.31	0.73	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$3\text{-}CNPhO^-$	2.36	1.76	0.27	0.72	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$4\text{-}CNPhO^-$	2.38	1.76	0.27	0.71	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$p\text{-}NO_2PhO^-$	2.39	1.77	0.26	0.70	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$2,3,5,6\text{-}F_4PhO^-$	2.39	1.79	0.25	0.68	$\cdots$	$\cdots$	$\cdots$	$\cdots$
CH_3COO^-	2.42	1.79	0.24	0.67	$\cdots$	$\cdots$	$\cdots$	$\cdots$

Table S4: The bond lengths, R_{P-Nu} and R_{P-Lg} (Å) and Wiberg bond indices, N_{P-Nu} and N_{P-Lg} of forming bonds (P-Nu) and breaking bonds (P-Lg) in the transition states of model reactions obtained from density-functional calculations with PCM solvation and UAKS radii.

calculated bond length and Wiberg bond index of P-O bond are 1.640 Å and 0.616 and 1.634 Å and 0.641 with UFF and UAKS model, respectively. Here bond length and bond order from UAKS model are used as reference to normalize calculated Wiberg bond indices.

The correlation between bond lengths and bond orders of P-Nu and P-Lg bonds in the transition states except ones with leaving alkanethiolate anions, was investigated under the assumption that bond orders vary exponentially with bond distances, as suggested by Linus Pauling,[9,10] according to the following model:

$$N = N_0 \cdot exp\left(\frac{R_0 - R}{C}\right) \tag{1}$$

where R and N are the bond length (Å) and bond order, respectively, R_0 and N_0 are the bond length and bond order of fully formed bond, respectively, and C is a constant to determine how fast the bond order changes with bond length. Here R_0 is chosen from the dimethyl phosphatethe

(R_0=1.64). N_0 and C are treated as fitting parameters. The correlation equation of bond lengths and bond orders of P-Nu and P-Lg bonds was obtained from curve fitting as follows,

$$N = 0.94 \cdot exp\left(\frac{1.64 - R}{0.56}\right) \tag{2}$$

The overall root-mean-square error (RMSE) was 0.02. The bond lengths and bond orders of P-Nu and P-Lg are plotted in the Figure S2, together with the correlation curve of bond lengths and bond orders.
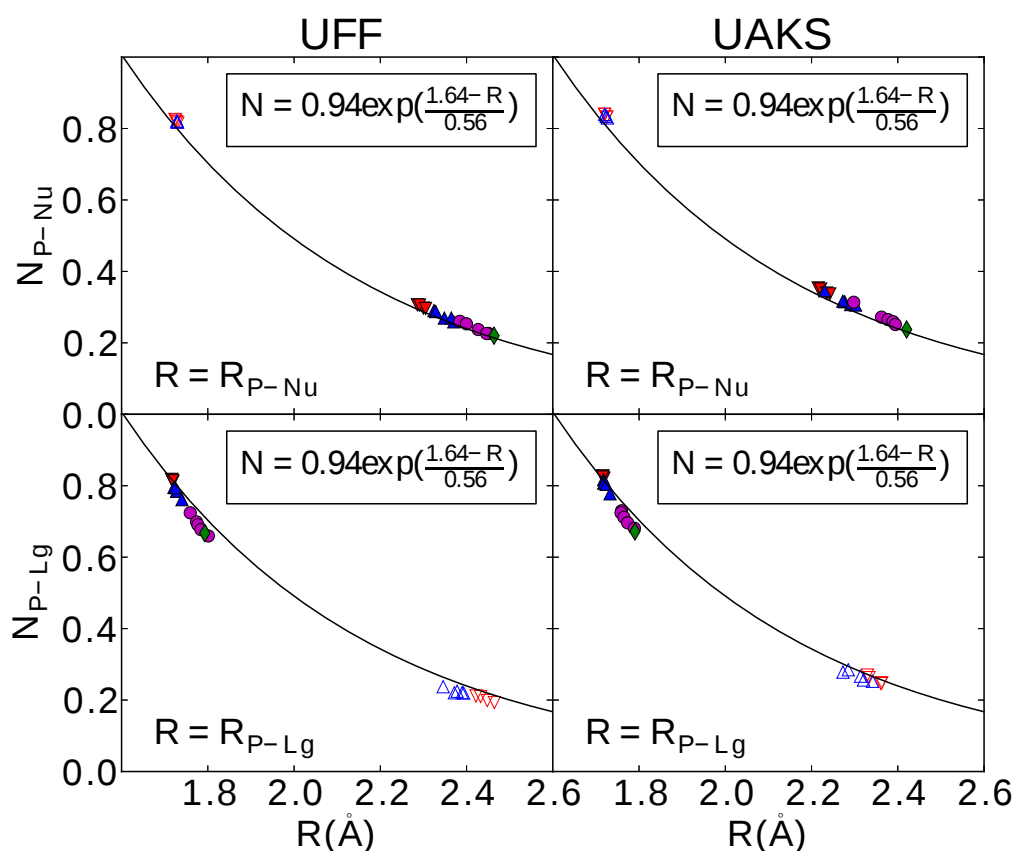

Figure S2: The relationship between bond lengths (R, Å) and bond orders (N) of P-Nu (top) and P-Lg (bottom) bonds in the rate limiting transition states of model reactions. Triangle, round and diamond symbols correspond to alkyl, phenol and acid leaving groups, respectively.

The figure S2 and the equation 2 show that bond lengths and bond orders of P-O bonds have a good exponential correlation with a small Root-Mean-Square Error, 0.02. The equation 2 is useful

for prediction of bond order of P-O bond by knowing the corresponding bond length or visa versa.

References

[1] K.-Y. Wong, H. Gu, S. Zhang, J. A. Piccirilli, M. E. Harris and D. M. York, *Angew. Chem. Int. Ed.*, 2012, **51**, 647–651.

[2] E. P. Serjeant and B. Dempsey, *Ionisation Constants of Organic Acids in Aqueous Solution*, Pergamon Press, New York, 1979.

[3] *CRC handbook of chemistry and physics*, ed. D. R. Lide, CRC Press LLC, Boca Raton, FL, 90th edn., 2010.

[4] N. Bourne, E. Chrystiuk, A. M. Davis and A. Williams, *J. Am. Chem. Soc.*, 1988, **110**, 1890–1895.

[5] P. Ballinger and F. A. Long, *J. Am. Chem. Soc.*, 1960, **82**, 795–798.

[6] D. M. Brown and D. A. Usher, *J. Chem. Soc.*, 1965, **87**, 6558–6564.

[7] J. P. Foster and F. Weinhold, *J. Am. Chem. Soc.*, 1980, **102**, 7211–7218.

[8] A. E. Reed, R. B. Weinstock and F. Weinhold, *J. Chem. Phys.*, 1985, **83**, 735–746.

[9] L. Pauling, *J. Am. Chem. Soc.*, 1947, **69**, 542–553.

[10] K. N. Houk, S. M. Gustafson and K. A. Black, *J. Am. Chem. Soc.*, 1992, **114**, 8565–8572.