HHS Public Access

Author manuscript

Phys Chem Chem Phys. Author manuscript; available in PMC 2015 August 14.

Published in final edited form as:

Phys Chem Chem Phys. 2014 August 14; 16(30): 15846–15855. doi:10.1039/c4cp01050g.

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

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

[*]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.

Abstract

RNA cleavage transesterification is of fundamental reaction in biology that is catalyzed by both protein and RNA enzymes. In this work, a series of RNA transesterification model reactions with a wide range of leaving groups are investigated with density-functional calculations in an aqueous solvation environment in order to study linear free energy relationships (LFERs) and their connection to transition state structure and bonding. Overall, results obtained from the polarizable continuum solvation model with UAKS radii produce the best linear correlations and closest overall agreement with experimental results. Reactions with a poor leaving group are predicted to proceed via a stepwise mechanism with a late transition state that is rate controlling. As leaving group becomes more acidic and labile, the barriers of both early and late transition states decrease. LFERs for each transition state are computed, with the late transition state barrier showing greater sensitivity to leaving group pK_a. For sufficiently enhanced leaving groups, the reaction mechanism transits to a concerted mechanism characterized by a single early transition state. Further linear relationships were derived for bond lengths and bond orders as a function of leaving group pK_a and rate constant values that can be used for prediction. This work provides important benchmark linear free energy data that allows a molecular-level characterization of the structure and bonding of the transition states for this important class of phosphoryl transfer reactions. The relations reported herein can be used to aid in the interpretation of data obtained from experimental studies of non-catalytic and catalytic mechanisms.

Introduction

Cleavage of the phosphodiester bond of RNA[1] is a fundamental phosphoryl transfer reaction in biology[2] that is catalyzed by both protein enzymes such as RNase A,[3–5] and RNA enzymes such as the class of small self-cleaving nucleolytic ribozymes[6,7] that include the hammerhead,[8–10] hairpin,[11,12] hepatitis delta virus,[13–16] varkud satellite[17–19] and *glmS*[20–23] ribozymes. The first step in this reaction involves a cleavage transesterification whereby the 2′-OH position on the RNA ribose ring becomes activated, and makes an in-line attack on

[‡]york@biomaps.rutgers.edu.

the adjacent phosphate, proceeding through a pentavalent transition state or intermediate, and resulting in a 2′,3′-cyclic phosphate and a cleaved 5′-leaving group (Scheme 1). Consequently, there is great interest in understanding the mechanisms whereby proteins and RNA enzymes are able to catalyze this reaction.

A powerful experimental method to study catalytic mechanism is to examine linear free energy relationships (LFERs) that provide insight into the nature of the transition state through examination of the sensitivity of the reaction rate constant (or equilibrium constant) to chemical modifications at key positions such as the nucleophile and leaving group.[24–32] For example, Brønsted coefficients have been utilized to estimate effective charge developed on the leaving group in the transition states and measure the effect of leaving group on the reaction rates.[29,33–37] The Leffler index[29,38–41] serves as an indicator of the extent of bond formation and bond fission in the transition state, and to locate its position along reaction coordinate. In this way, LFERs are used to make qualitative inferences about the nature of the transition state geometry, bonding and charge distribution.

The goal of this work is to establish a *quantitative connection* between LFER data and molecular structure and bonding relevant for RNA cleavage transesterification reactions. Toward this end, we have performed density-functional calculations for a set of RNA transesterification model reactions illustrated in Scheme 1, with different leaving groups shown in Scheme 2. The results are compared with available experimental data, and provide a detailed atomic level picture of mechanism. Further, relationships are established that allow prediction of bond lengths and bond orders in the rate controlling transition states that can be used to aid in the quantitative interpretation of LFERs in enzymes and ribozymes.

Computational Methods

To explore the in-line mechanisms of RNA transesterification model reactions and analyze linear free energy relationships, stationary points (minima and transition states) along the reaction coordinate for the in-line mechanisms were identified using GAUSSIAN09[42] suite of programs with the M06-2X[43] density-functional model and an ultrafine numerical integration grid (pruned from 99/590 radial/angular points). Geometry optimizations were carried out with the 6-31++G(d,p) basis set with default convergence criteria. Frequency analysis at the same theoretical level were performed to establish the nature of all the stationary points and to allow evaluation of thermodynamic quantities. Reaction pathways were verified with intrinsic reaction path calculations. Electronic energies for optimized geometries were further refined by single point calculation using the 6-311++G(3df,2p) basis set. This protocol for geometry optimization and energy refinement is designated herein by the abbreviated notation M06-2X/6-311++G(3df,2p)//M06-2X/6-31++G(d,p). Thermodynamic properties at 298.15K were obtained from the theoretical calculations using standard statistical mechanical expressions for separable vibrational, rotational and translational contributions in the canonical ensemble[44] and have been described in detail elsewhere.[45]

Bond orders for bonds in the transition states were investigated using Natural Bond Order (NBO) analysis[46,47] at the same level of theory and basis set as for the geometry

optimization. Calculated Wiberg bond orders are normalized by dividing 0.64, the Wiberg bond order calculated for a bridging P-O single bond in dimethyl phosphate (taken to be a fully formed P-O single bond).

Solvent effects are important in phosphoryl transfer reactions[48] and were examined using the same geometry optimization/energy refinement protocol, M06-2X/6-311++G(3df,2p)// M06-2X/6-31++G(d,p), as for the gas phase calculations, but with solvation effects included self-consistently and with full geometry optimization using the polarizable continuum model (PCM)[49,50] with two different sets of solvation radii (UFF[51] and UAKS[52]). The UFF radii is the default used in the Gaussian PCM solvation model, whereas the UAKS radii were optimized for use with Kohn-Sham density functional calculations at the PBE0/6-31G(d) level of theory.[42] Our experience has been that the UAKS radii are generally reliable for stationary points that are stable minima, and fairly transferable to density-functional models that give similar densities (including the M06-2X functional used here, see for example references 53, 54). A limitation of the model in its current form, however, is that the radii do not necessarily adjust smoothly along the reaction path, and are not necessarily reliable for all transition states or transient intermediates. During the calculation, the default UAKS radii in the intermediates differed from those in the transition states. In order to make these radii consistent, we used an averaging strategy[48] whereby the radii of oxygen atoms connected to phosphorus in the intermediates were set to the average radii of oxygen atoms of the leaving group and nucleophile in the nearyby transition states. (see Supporting Information for details).

Results and Discussion

A series of RNA transesterification model reactions (Scheme 1) are herein studied with 17 different leaving groups (Scheme 2) that fall into the following five general classes: 1) alkoxide, 2) hetero-alkoxide, 3) alkanethiolate, 4) aryloxide and 5) carboxylate anions, and range from electron-donating poor leaving groups to electron-withdrawing enhanced leaving groups. In all of the model reactions studied here, the first step involves the association of the nucleophile - an intramolecular process that initiates formation of a five membered ring involving a pentavalent phosphorane species. There are two formal associative mechanisms that lead to the same products: a concerted mechanism that proceeds through a single transition state, and a stepwise mechanism that proceeds through two transition states that are separated by an intermediate. A concerted mechanism is described as either synchronous (having similar degrees of bonding to the nucleophile and leaving group in the transition state) or asynchronus (having differing degrees of bonding to the nucleophile and leaving group in the transition state). For either stepwise mechanisms or concerted asynchronus mechanisms, the transition states can further be designated as either “early” or “late”, depending on where along the reaction coordinate they occur. We denote a transition state as being “early” if it is characterized by a small degree of bond formation/cleavage with the nucleophile/leaving group, respectively. Conversely, a “late” transition state involves a nearly fully formed bond with the nucleophile and a nearly cleaved bond with the leaving group. For the “enhanced” leaving groups (with pK_a values less than ≈ 13) considered here, the mechanisms are concerted asynchronous and proceed through an *early* transition state (TS1). Alternatively, for relatively poor leaving groups (with pK_a values greater than ≈ 13)

the mechanisms are stepwise and proceed through both *early* and *late* transition states (TS1 and TS2, respectively) separated by a shallow metastable intermediate (I). As described in detail below, *early* and *late* refer to the location of the transition state along a reaction coordinate that involves the difference in leaving group and nucleophile distances with the reactive phosphorus (negative and positive reaction coordinate values indicate early and late transition states, respectively).

Representative examples of rate-limiting transition states for each of these leaving group classes are presented in Figure 1. Free energies of transition states, intermediates and products relative to reactants of model reactions, calculated using both UFF and UAKS solvation radii, are listed in Table 1 along with experimental pK_as of the leaving group conjugate acids in water at 25°C. Activation free energies for rate-limiting transition states of model reactions calculated at 25° and 80° are presented in Table 3 together with available experimental barriers.

In this work, we consider three following Brønsted-type correlations[24,25]

$$\beta_{lg}^{1}=\partial log k^{1}/\partial \mathrm{p}K_{\mathrm{a}} \quad (1)$$

$$\beta_{lg}^{2}=\partial log k^{2}/\partial \mathrm{p}K_{\mathrm{a}} \quad (2)$$

$$\beta_{eq}=\partial log K_{eq}/\partial \mathrm{p}K_{\mathrm{a}} \quad (3)$$

where k^1 and k^2 are the rate constants estimated from the free energy barriers for TS1 and TS2, respectively, K_{eq} is the equilibrium constant estimated from the reaction free energy, and pK_a refers to the leaving group conjugate acid (LgH). Also of interest is the so-called Leffler index[38] ($\alpha_{fission}$) which is defined as a ratio between the Brønsted correlations as

$$\alpha_{fission}=\partial log k/\partial log K_{eq}=\beta_{lg}/\beta_{eq} \quad (4)$$

where β_{lg} and k are the Brønsted correlation and the rate constant, respectively, for the rate-controlling transition state.

The calculated Brønsted correlations β_{lg}^{1}, β_{lg}^{2} and β_{eq} are plotted in Figure 2 for both UFF and UAKS solvation radii. The β_{lg} values from available experiments are summarized in Table 2, together with our calculated β_{lg} at 25° and 80° for comparison. Finally, we examine correlations between bond lengths or bond orders and *logk* or pK_a values in Figures 3 and 4, respectively. All quantitative correlation relationships have been collected and are compiled in a single table (Table 4) for reference.

Experimental Reference Data

Several model reactions with various aryl and alkyl leaving groups have been investigated experimentally at different temperatures to analyze LFERs of RNA transesterification; the measured Brønsted coefficients are shown in Table 2. Lönnberg *et al.*[55] analyzed original measured kinetic data[33,56] of uridine 3′-phosphate diester cleavage and derived a non-linear

Brønsted correlation with a convex break at pK_a of 12.58; two significantly different Brønsted values (β_{lg}^1 and β_{lg}^2) of -0.52 and -1.34 were obtained for model reactions with aryl and alkyl leaving groups, respectively. The reference β_{eq} value of -1.74 is taken from phosphoryl transfer of phosphono monoanion;[57] the β_{eq} value has been widely used as an estimation for β_{eq} of RNA transesterification model reactions.[2,33]

It should be mentioned that the experimental model system in Table 2 that is closest in structure to the system examined in the present computational work is 2-hydroxypropyl phosphate diester.[58] This data was analyzed[55] by fitting to a non-linear Brønsted correlation model, although the authors note that the data was not sufficient to obtain definitive Brønsted parameters. Nonetheless, the break in the Brønsted plot appears to occur at a higher pK_a value than that measured for uridine 3′-phosphate diester cleavage[33,56] and predicted from the present calculations on a very similar model system.

Comparison of Brønsted Coefficients

Calculated and experimental activation free energies for rate-limiting transition states are listed in Table 3. The barrier differences between UFF and UAKS calculations are on average 0.25±1.58 and 0.29±1.66 kcal/mol at 25°C and 80°C, respectively. UFF and UAKS barriers are on average 5.05±2.46 and 5.13±1.43 kcal/mol lower than experimental values at 80° for cyclization of 2-hydroxypropyl phosphate diesters. The calculated barriers are generally closer to the experimental values for the transesterification of uridine 3′-phosphates (mean signed error of 0.47±1.80 and 0.58±0.96 kcal/mol for UFF and UAKS radii, respectively).

Linear relationships between calculated logarithmic rate constants for each transition state (TS1 and TS2) and the experimental pK_a of leaving groups are illustrated in Figure 2. Linear correlations from both UFF and UAKS solvation radii were overall similar, with correlation coefficients ranging in magnitude between 0.93 and 0.98. The β_{lg}^1 values predicted by UFF and UAKS (-0.54 and -0.52, respectively) are in close agreement with each other, whereas the β_{lg}^2 (-2.04 and -1.37, respectively) and β_{eq} (-2.44 and -1.74, respectively) values show greater variation. As expected, the β_{lg}^2 values are considerably greater than for β_{lg}^1, illustrating the increased sensitivity of the late transition state TS2 to nature of the leaving group. The sulfur-containing leaving groups appear as outliers for fits of logK_{eq}, particularly for the UFF model, due largely to overstabilization of the thiolate anion in solution by these models.

The LFER results from the UAKS radii are in excellent agreement with available experimental values for similar reactions. Calculated and experimental values for β_{lg}^1, β_{lg}^2 and β_{eq} are compared in Table 2. The β_{lg}^1 value for predicted by both UFF and UAKS (-0.54 and -0.52, respectively) are in close agreement with experimental value (-0.52) for the hydroxide-ion catalyzed transesterifications of uridine 3′-phosphates with good leaving groups.[55] The β_{lg}^2 value predicted from the UAKS radii (-1.37), also agrees well with the experimental value of -1.34 for uridine 3′-phosphates with poor leaving groups.[55] The

UAKS value for β_{eq} (-1.74) is in excellent agreement with the value measured for the phosphono monoanion transfer reaction.[57]

Another important quantity to analyze is the break point in the LFERs corresponding to TS1 and TS2 derive; i.e., the intersection of the fitted lines for β_{lg}^{1} and β_{lg}^{2}. The break point predicted from the calculation with UAKS radii is 12.98, which is close to the analogous experimentally predicted value of 12.58.[55] The calculated value also falls within the range of the experimental pK_a values for the 2′-hydroxyl group in uridine 3′-phosphate ethyl ester (12.85),[59] and for the 2′-hydroxyl group in chimeric oligonucleotide (13.1).[60]

Overall, the UFF radii are in poorer quantitative agreement with experiment. This is not terribly surprising since the UAKS model radii were optimized to reproduce solvation free energies when used with density functional methods, having been developed using PBE0/6-31G(d) parameterization level.[42] Nonetheless, the very close alignment of the UAKS results with experiments provides support for the supposition that one can use this model as a basis from which to determine reaction mechanism, and make a quantitative connection between experimental data and transition state geometry and bonding. These connections will be the focus of the following sections.

Reaction Mechanism

The results in Figures 1 and 2 illustrate that leaving groups have strong impact on the mechanisms of model reactions. Table 1 lists the calculated relative free energies of transition states, intermediates and products relative to reactants for both UAKS and UFF radii models, as well as the experimental pK_a values for the leaving groups.

The pK_a values listed in the table and discussed herein correspond to the equilibrium

$$LgH\,(aq) \rightleftharpoons Lg^{-}\,(aq) + \mathrm{H}^{+}\,(aq) \quad (5)$$

where the leaving group anion is $Lg^{-}(aq)$. Consequently, decreasing pK_a indicates a shift in the equilibrium so as to favor the state where the Lg-H bond is broken leading to the $Lg^{-}(aq)$ species. Decreasing leaving group pK_a is thus expected to be correlated with the transesterification reaction free energy, since the reaction product involves a similar state whereby a Lg-P bond is broken leading to the $Lg^{-}(aq)$ species. Indeed, we observe that the reaction free energies in Table 1, ΔG, generally trend toward more negative values with decreasing leaving group pK_a. Stability of the anion leaving group depends on its ability to electronically withdraw electron density up to a full -1 charge, and its solvation free energy. Hence, high pK_a values correspond to poor leaving groups, and low pK_a values correspond to good (enhanced) leaving groups.

As is typical for phosphoryl transfer reactions of phosphate diesters,[2] all of the transition states predicted here are associative in nature, meaning that the approach of the nucleophile to the phosphorus generally precedes cleavage of the bond to the leaving group. For the purposes of discussion, therefore, we introduce a coordinate, ζ, that describes the overall reaction progression as the difference in leaving group and nucleophile bond distances to phosphorus, i.e., $\zeta = R_{P-Lg} - R_{P-Nu}$. With this definition of reaction coordinate, ζ for reactant

and product states would have large negative and positive values, respectively. Referring to Scheme 1, the ζ value for the "early" transition state (TS1) would have a small negative value, whereas the "late" transition state (TS2) would have a small positive ζ value, and the intermediate (I) would have a near zero value. Here by "large", "small" and "near zero" we mean that magnitudes are roughly greater than 1 Å, between 1 and 0.5 Å, and less than 0.5 Å, respectively.

If the reaction is native transesterification where the nucleophile and leaving group have similar pK_a values, the reaction will proceed through both an early and a late transition states, separated by a shallow, metastable intermediate.[61] For a dianionic transition state, the intermediate is high in energy,[62] and not sufficiently long lived[55] to undergo pseudorotation.[63,64] The UAKS results shown in Table 1 which are in good agreement with available experimental data indicate that the barriers to decomposition of the intermediates are fairly small (on average 3.00±0.58 kcal/mol). For reactions involving leaving groups that have comparable pK_a values to that of the nucleophile, the two transition states are expected to be somewhat similar in energy. However, the endocyclic bond between the nucleophile and phosphorus must also take into account some degree of strain energy to form a five-membered ring in the pentavalent transition state, although this is expected to be quite small. More significant is that cleavage of the exocyclic bond results in a considerably more strained tetravalent phosphorus species which is higher in energy. These effects have been discussed in detail elsewhere.[65–67] In addition to ring strain, differential solvation effects[62,68,69] of the acyclic reactant and cyclic product states also play a role. The overall result, for this series of reactions, is that for stepwise mechanisms (e.g., with poor leaving groups having high pK_a values), cleavage of the exocyclic bond (TS2) is typically rate controlling except the case that trifluoroethoxide serves as a leaving group where the barrier of TS1 is just slightly higher than that of TS2 by 1.51 kcal/mol. It should be noted that the present series of model reactions differ from native RNA transesterification reactions in that 1) the nucleophile of the former (a primary alcohol) has a higher pK_a value than the secondary 2′OH group of the latter, and 2) the strain energy of ring formation is expected to differ due to coupling of the second ribose ring in the case of RNA.

The trends in reactivity modeled by the LFERs, and the ultimate shift in mechanism from a stepwise pathway involving a rate-controlling late (TS2) transition state to a concerted pathway with a single early (TS1) transition state, can be easily understood through consideration of the Hammond effect.[24,70] As the pK_a of the leaving group decreases, the reaction equilibrium favors the product state that contains the solvated anionic leaving group. As a consequence, the TS2 barrier height becomes lower, and shifts away from the products (i.e., toward less positive ζ values). Since chemical modification of the leaving group has a direct and profound effect on the stability of the exocyclic chemical bond to phosphorus, the cleavage of which is characteristic of the rate-controlling TS2, we expect β_{lg}^2 to have a large magnitude, indicating that the reaction rate is highly sensitive to modifications of the leaving group.

At some point, as the pK_a of the leaving group becomes lower, the shift and lowering of TS2 is such that it becomes only a decaying shoulder in the reaction profile and ultimately

vanishes. This causes the reaction to revert to a single-step mechanism with only an early transition state (TS1) corresponding to formation of the endocyclic bond between the nucleophile and phosphorus. At this point a convex break point in the LFERs occurs, and for reactions involving leaving groups with pK_a values lower than the break point, the value of β_{lg}^1 is smaller in magnitude than β_{lg}^2, reflecting a diminished sensitivity to variation of the leaving group. Indeed, the β_{lg}^2 is more than 2.5 times larger than β_{lg}^1, which suggests Brønsted coefficients are useful indexes to identify mechanisms of RNA transesterification reactions. The mechanism predicted from our calculations is consistent with the interpretation of experimental data of Davis *et al.*[33] and Kosonen *et al.*[56] on transesterification of uridine 3′-phosphate diesters. The computational results presented here indicate, for the series of model reactions considered, sufficiently enhanced leaving groups lead to concerted mechanisms that all proceed through a single early TS, and LFER analysis predicts a β_{lg} value with a small magnitude. For less enhanced and poor leaving groups, mechanisms are stepwise, and in almost all cases the rate-controlling transition state is late, leading to a large negative β_{lg} value.

Further support for the supposition that the rate-limiting transition states for transesterification of uridine 3′-phosphate diesters with good leaving groups should be early (TS1) is given by calculation of the Leffler index,[2,38] $\alpha_{fission}$. The $\alpha_{fission} = \beta_{lg}/\beta_{eq}$=0.30, where the β_{lg} value of -0.52 is used.[55] The $\alpha_{fission}$ value is consistent with the average bond order of breaking P-Lg bond in early transition state from calculation with UAKS radii, 0.31.

Transition State Structure and Bonding

Experimental measurement of LFERs provides insight into mechanism, and qualitative inferences can be made with regard to characterization of the transition state. Theoretical calculations, on the other hand, can provide detailed information about the structure and bonding in the transition state. Of course, in order for this detailed information to be useful, the theoretical models must be sufficiently validated with respect to experiment. We have demonstrated that the present electronic structure calculations and PCM solvation model with UAKS radii agree well with available experiments on similar RNA transesterification model reactions. Consequently, we may proceed to derive relations that allow the prediction of structure and bonding in the transition state based on these results.

Calculated bond length (R) and bond order (N) results for the forming (P-Nu) bond and breaking (P-Lg) bond for both TS1 and TS2 are listed in Supporting Information. Linear relationships have been identified between these bond quantities and the calculated logarithmic rate constants or experimental leaving group pK_a values. These correlations are illustrated in Figures 3-4 and the regression values, along with linear correlation coefficients, are listed in Table 4 and can be used for prediction. Analogous correlations for TS2 were not evident from the calculations. Nonetheless, we were able to reasonably relate the bond lengths and bond orders in the transition states for this reaction with a simple 3-parameter exponential model:

$$N(R) = Ae^{(B-R)/C} \quad (6)$$

where R is the P-Nu or P-Lg bond length (Å) in the transition state, and the fitted parameters are A=0.94, B=1.64 Å and C=0.56 Å. This relation allows one to infer transition state bonding from geometry or visa versa. Plots of these relations are given in Supporting Information. The rich bonding information for the rating-limiting transition state is not otherwise quantitatively interpretable from Leffler indices.[2]

Conclusions

Herein a series of RNA transesterification model reactions with a wide range of leaving groups have been investigated with density-functional calculations in an aqueous solvation environment modeled with two different sets of solvation radii (UFF and UAKS). Linear free energy relationships are derived from the calculations for both early and late transition states. Results using the UAKS radii agree closely with available experiments, and provide a model from which quantitative information about transition state structure and bonding can be derived. Depending on the nature of the leaving group, reactions may proceed via a stepwise mechanism that passes through both an early and late transition states separated by a transient intermediate, or through a single early transition state. Brønsted correlations, β_{lg}^{1} and β_{lg}^{2} can be used to distinguish these two mechanisms. Further correlations are derived that connect transition state bond lengths and bond orders with experimental reaction rate constants and leaving group pK_a values, and between transition state bond lengths and bond orders. Together, these results provide models from which to aid in the interpretation of experimental LFER data, and make predictions about RNA cleavage transesterification reactions catalyzed by proteins and RNA enzymes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

The authors are grateful for financial support provided by the National Institutes of Health (GM62248). Computational resources from the Minnesota Supercomputing Institute for Advanced Computational Research (MSI) were utilized in this work. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575.

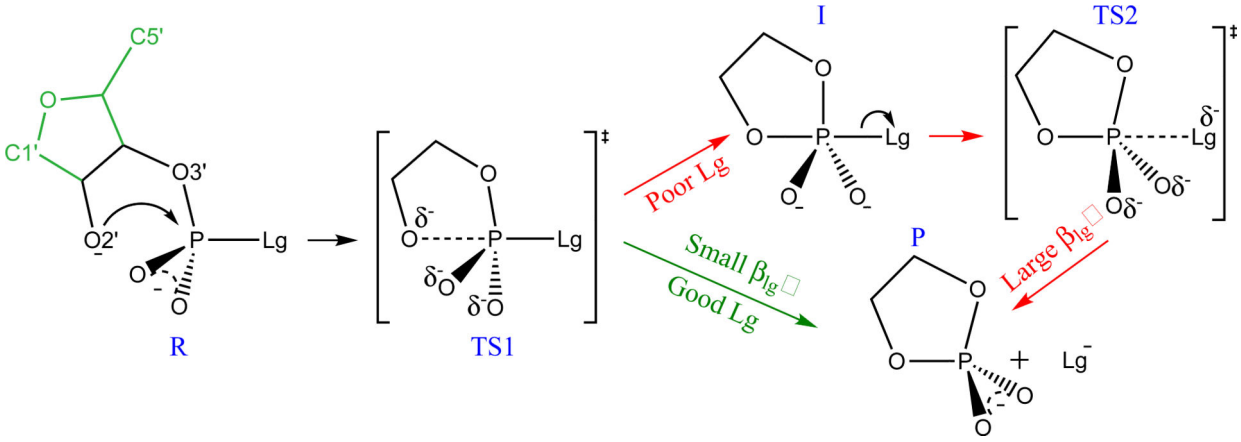
References

1. Perreault DM, Anslyn EV. Angew. Chem. Int. Ed. 1997; 36:432–450.
2. Lassila JK, Zalatan JG, Herschlag D. Annu. Rev. Biochem. 2011; 80:669–702. [PubMed: 21513457]
3. Raines RT. Chem. Rev. 1998; 98:1045–1065. [PubMed: 11848924]
4. Formoso E, Matxain JM, Lopez X, York DM. J. Phys. Chem. B. 2010; 114:7371–7382. [PubMed: 20455590]
5. Gu H, Zhang S, Wong K-Y, Radak BK, Dissanayake T, Kellerman DL, Dai Q, Miyagi M, Anderson VE, York DM, Piccirilli JA, Harris ME. Proc. Natl. Acad. Sci. USA. 2013; 110:13002–13007. [PubMed: 23878223]

6. Ferré-D'Amaré AR, Scott WG. Cold Spring Harb Perspect Biol. 2010; 2:a003574. [PubMed: 20843979]
7. Scott WG. Curr. Opin. Struct. Biol. 2007; 17:280–286. [PubMed: 17572081]
8. Scott WG, Murray JB, Arnold JRP, Stoddard BL, Klug A. Science. 1996; 274:2065–2069. [PubMed: 8953035]
9. Scott WG. Q. Rev. Biophys. 1999; 32:241–294. [PubMed: 11194566]
10. Martick M, Lee T-S, York DM, Scott WG. Chem. Biol. 2008; 15:332–342. [PubMed: 18420140]
11. Walter NG, Burke JM. Curr. Opin. Chem. Biol. 1998; 2:24–30. [PubMed: 9667918]
12. Rupert PB, Massey AP, Sigurdsson ST, Ferré-D'Amaré AR. Science. 2002; 298:1421–1424. [PubMed: 12376595]
13. Sharmeen L, Kuo MY, Dinter-Gottlieb G, Taylor J. J. Virol. 1988; 62:2674–2679. [PubMed: 2455816]
14. Ferré-D'Amaré AR, Zhou K, Doudna JA. Nature. 1998; 395:567–574. [PubMed: 9783582]
15. Cerrone-Szakal AL, Siegfried NA, Bevilacqua PC. J. Am. Chem. Soc. 2008; 130:14504–14520. [PubMed: 18842044]
16. Chen J-H, Yajima R, Chadalavada DM, Chase E, Bevilacqua PC, Golden BL. Biochemistry. 2010; 49:6508–6518. [PubMed: 20677830]
17. Hoffmann B, Mitchell GT, Gendron P, Major F, Andersen AA, Collins RA, Legault P. Proc. Natl. Acad. Sci. USA. 2003; 100:7003–7008. [PubMed: 12782785]
18. Jaikaran D, Smith MD, Mehdizadeh R, Olive J, Collins RA. RNA. 2008; 14:938–949. [PubMed: 18356538]
19. Wilson TJ, Lilley DM. RNA. 2011; 17:213–221. [PubMed: 21173201]
20. Klein DJ, Ferré-D'Amaré AR. Science. 2006; 313:1752–1756. [PubMed: 16990543]
21. Klein DJ, Been MD, Ferré-D'Amaré AR. J. Am. Chem. Soc. 2007; 129:14858–14859. [PubMed: 17990888]
22. Cochrane JC, Lipchock SV, Smith KD, Strobel SA. Biochemistry. 2009; 48:3239–3246. [PubMed: 19228039]
23. Viladoms J, Scott LG, Fedor MJ. J. Am. Chem. Soc. 2011; 133:18388–18396. [PubMed: 21936556]
24. Jencks DA, Jencks WP. J. Am. Chem. Soc. 1977; 99:7948–7960.
25. Jencks WP. Chem. Rev. 1985; 85:511–527.
26. Warshel A, Schweins T, Fothergill M. J. Am. Chem. Soc. 1994; 116:8437–8442.
27. Åqvist J, Kolmodin K, Florian J, Warshel A. Chem. Biol. 1999; 6:R71–R80. [PubMed: 10074472]
28. Mihai C, Kravchuk AV, Tsai M-D, Bruzik KS. J. Am. Chem. Soc. 2003; 125:3236–3242. [PubMed: 12630878]
29. Onyido I, Swierczek K, Purcell J, Hengge AC. J. Am. Chem. Soc. 2005; 127:7703–7711. [PubMed: 15913360]
30. Florián J, Åqvist J, Warshel A. J. Am. Chem. Soc. 1998; 120:11524–11525.
31. Plotnikov NV, Prasad BR, Chakrabarty S, Chu ZT, Warshel A. J. Phys. Chem. B. 2013; 117:12807–12819. [PubMed: 23601038]
32. Prasad BR, Plotnikov NV, Warshel A. J. Phys. Chem. B. 2013; 117:153–163. [PubMed: 23198768]
33. Davis AM, Hall AD, Williams A. J. Am. Chem. Soc. 1988; 110:5105–5108.
34. Davis AM, Regan AC, Williams A. Biochemistry. 1988; 27:9042–9047. [PubMed: 2852962]
35. Herschlag D, Jencks WP. J. Am. Chem. Soc. 1989; 111:7587–7596.
36. Hollfelder F, Herschlag D. Biochemistry. 1995; 34:12255–12264. [PubMed: 7547968]
37. Holtz KM, Catrina IE, Hengge AC, Kantrowitz ER. Biochemistry. 2000; 39:9451–9458. [PubMed: 10924140]
38. Leffler JE. Science. 1953; 117:340–341. [PubMed: 17741025]
39. Williams A. Acc. Chem. Res. 1984; 17:425–430.
40. Almer H, Strömberg R. J. Am. Chem. Soc. 1996; 118:7921–7928.

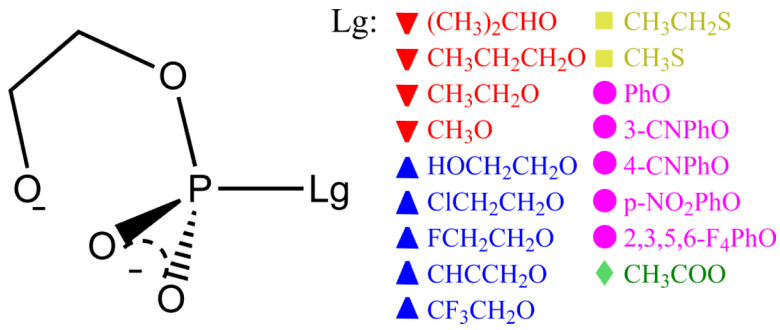
41. Younker JM, Hengge AC. J. Org. Chem. 2004; 96:9043–9048. [PubMed: 15609936]
42. Frisch, MJ.; Trucks, GW.; Schlegel, HB.; Scuseria, GE.; Robb, MA.; Cheeseman, JR.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, GA.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, HP.; Izmaylov, AF.; Bloino, J.; Zheng, G.; Sonnenberg, JL.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; T., V.; Montgomery, JA., Jr.; Peralta, JE.; Ogliaro, F.; Bearpark, M.; Heyd, JJ.; Brothers, E.; Kudin, KN.; Staroverov, VN.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, JC.; Iyengar, SS.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, JM.; Klene, M.; Knox, JE.; Cross, JB.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, RE.; Yazyev, O.; Austin, AJ.; Cammi, R.; Pomelli, C.; Ochterski, JW.; Martin, RL.; Morokuma, K.; Zakrzewski, VG.; Voth, GA.; Salvador, P.; Dannenberg, JJ.; Dapprich, S.; Daniels, AD.; Farkas, O.; Foresman, JB.; Ortiz, JV.; Cioslowski, J.; Fox, DJ. Gaussian 09, Revision A.02. Gaussian, Inc.; Wallingford, CT: 2009.
43. Zhao Y, Truhlar DG. Theor. Chem. Acc. 2008; 120:215–241.
44. Cramer, CJ. Essentials of Computational Chemistry: Theories and Models. 2nd edn.. John Wiley & Sons; Chichester, England: 2002.
45. Range K, McGrath MJ, Lopez X, York DM. J. Am. Chem. Soc. 2004; 126:1654–1665. [PubMed: 14871095]
46. Foster JP, Weinhold F. J. Am. Chem. Soc. 1980; 102:7211–7218.
47. Reed AE, Weinstock RB, Weinhold F. J. Chem. Phys. 1985; 83:735–746.
48. Wong K-Y, Gu H, Zhang S, Piccirilli JA, Harris ME, York DM. Angew. Chem. Int. Ed. 2012; 51:647–651.
49. Tomasi J, Mennucci B, Cammi R. Chem. Rev. 2005; 105:2999–3093. [PubMed: 16092826]
50. Scalmani G, Frisch MJ. J. Chem. Phys. 2010; 132:114110–114124. [PubMed: 20331284]
51. Rappé AK, Casewit CJ, Colwell KS, Goddard WA III, Skiff WM. J. Am. Chem. Soc. 1992; 114:10024–10035.
52. Barone V, Cossi M, Tomasi J. J. Chem. Phys. 1997; 107:3210–3221.
53. Liu Y, Gregersen BA, Hengge A, York DM. Biochemistry. 2006; 45:10043–10053. [PubMed: 16906762]
54. Liu, Y. Ph.D. thesis. University of Minnesota; 2008.
55. Lönnberg H, Strömberg R, Williams A. Org. Biomol. Chem. 2004; 2:2165–2167. [PubMed: 15280948]
56. Kosonen M, Youseti-Salakdeh E, Strömberg R, Lönnberg H. J. Chem. Soc. Perkin Trans. 1997; 2:2611–2666.
57. Bourne N, Williams A. J. Org. Chem. 1984; 49:1200–1204.
58. Brown DM, Usher DA. J. Chem. Soc. 1965; 87:6558–6564.
59. Acharya S, Földesi A, Chattopadhyaya J. J. Org. Chem. 2003; 68:1906–1910. [PubMed: 12608809]
60. Li Y, Breaker RR. J. Am. Chem. Soc. 1999; 121:5364–5372.
61. Radak BK, Harris ME, York DM. J. Phys. Chem. B. 2013; 117:94–103. [PubMed: 23214417]
62. Liu Y, Gregersen BA, Lopez X, York DM. J. Phys. Chem. B. 2005; 109:19987–20003. [PubMed: 16853584]
63. López CS, Faza ON, Gregersen BA, Lopez X, de Lera AR, York DM. Chem. Phys. Chem. 2004; 5:1045–1049. [PubMed: 15298394]
64. López CS, Faza ON, de Lera AR, York DM. Chem. Eur. J. 2005; 11:2081–2093. [PubMed: 15714539]
65. Kluger R, Taylor SD. J. Am. Chem. Soc. 1990; 112:6669–6671.
66. Kluger R, Taylor SD. J. Am. Chem. Soc. 1991; 113:5714–1719.
67. Taylor SD, Kluger R. J. Am. Chem. Soc. 1992; 114:3067–3071.
68. Dejaegere A, Karplus M. J. Am. Chem. Soc. 1993; 115:5316–5317.
69. Lopez X, Dejaegere A, Leclerc F, York DM, Karplus M. J. Phys. Chem. B. 2006; 110:11525–11539. [PubMed: 16771429]
70. Hammond GS. J. Am. Chem. Soc. 1955; 77:334–338.

71. Serjeant, EP.; Dempsey, B. Ionisation Constants of Organic Acids in Aqueous Solution. Pergamon Press; New York: 1979.
72. Lide, DR., editor. CRC handbook of chemistry and physics. 90th edn.. CRC Press LLC; Boca Raton, FL: 2010.
73. Bourne N, Chrystiuk E, Davis AM, Williams A. J. Am. Chem. Soc. 1988; 110:1890–1895.
74. Ballinger P, Long FA. J. Am. Chem. Soc. 1960; 82:795–798.
75. Kosonen M, Yousefi-Salakdeh E, Strömberg R, Lönnberg H. J. Chem. Soc., Perkin Trans. 1998; 2:1589–1595.
76. Satoh K, Inoue Y. Chem. Lett. 1972; 1:1097–1100.
77. Eyring H. J. Chem. Phys. 1935; 3:107–115.



Scheme 1.

The mechanisms of RNA transesterification model reactions with different leaving groups (Lg^-). R, TS1, I, TS2 and P stand for reactant, early transition state, intermediate, late transition state and product, respectively. Comparison of the RNA numbering scheme is shown in the leftmost frame (RNA atoms that are not present in the model reactions are indicated in green).



Scheme 2.

2-hydroxyethyl phosphate diester with various leaving groups (Lg^{-}), color/shape coded into the following five general classes: alkoxide (red/down triangle), hetero-alkoxide (blue/up triangle), alkanethiolate (yellow/square), aryloxide (purple/circle) and carboxylate (green/diamond) anions.

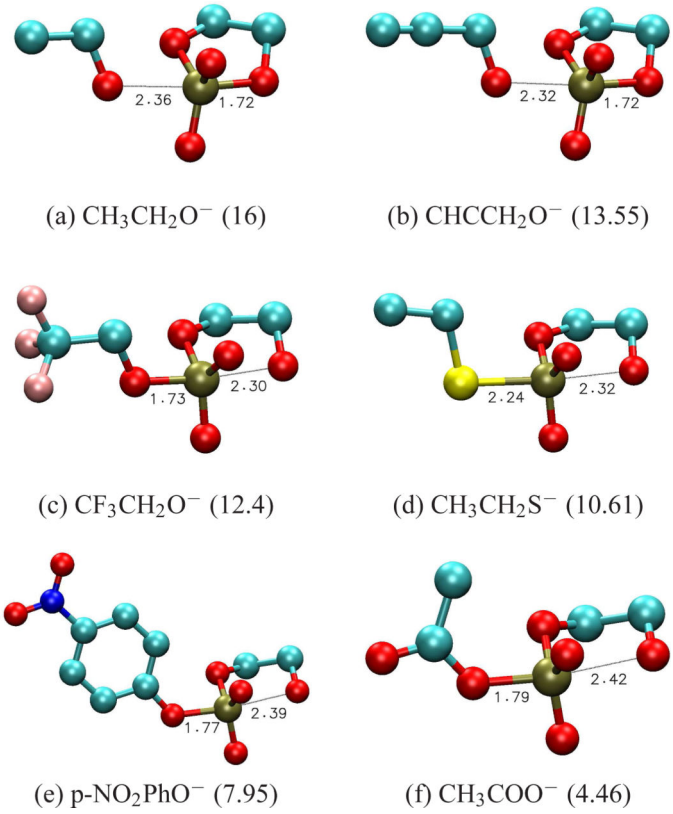


Figure 1.

Structures of representative rate-limiting transition states for model reactions. Leaving groups (Lg^-) and the pK_as associated with their conjugate acids (LgH) are indicated immediately below the structures.

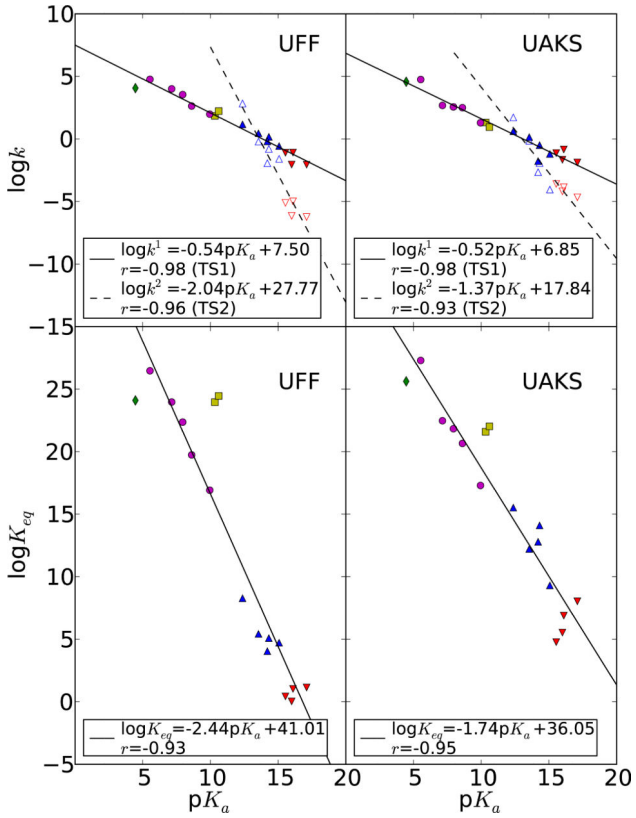


Figure 2.

Linear free energy relationships between calculated logarithmic rate constants (logk^1 and logk^2) and experimental pK_as of leaving groups (top) and between calculated logarithmic equilibrium constants (logK_{eq}) and experimental pK s of leaving 1 a groups (bottom). k and k^2 are calculated 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 logk^1, logk^2 and logK_{eq} are given as well as linear correlation coefficient for r.

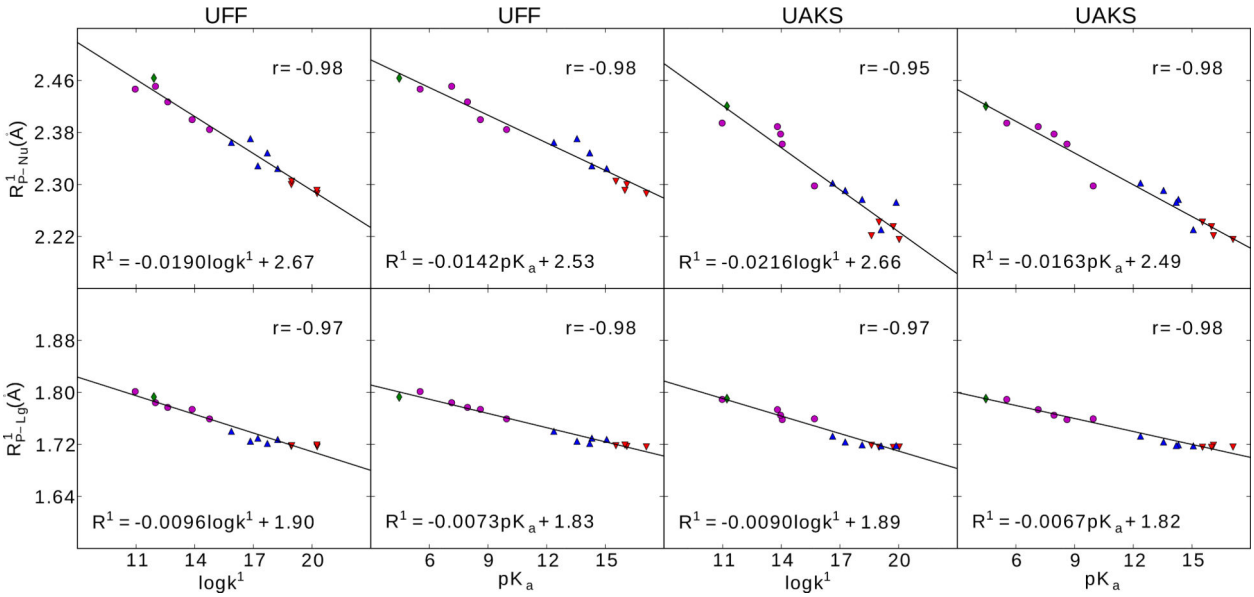


Figure 3.

Linear relationships between bond lengths ($R^1_{\mathrm{P}-Nu}$ and $R^1_{\mathrm{P}-Lg}$ P-Nu P-Lg, Å) of early transition states (TS1) and the corresponding calculated rate constants (logk^1) and experimental pK_as of leaving groups. The calculation were performed using density-functional method with PCM solvation model and UFF (left) and UAKS (right) radii. Red down triangle, blue up triangle, purple round and green diamond symbols correspond to alkyl, hetero-alkyl, phenyl and acid leaving groups, respectively. Regression parameters for R^1 are given as well as linear correlation coefficient for r.

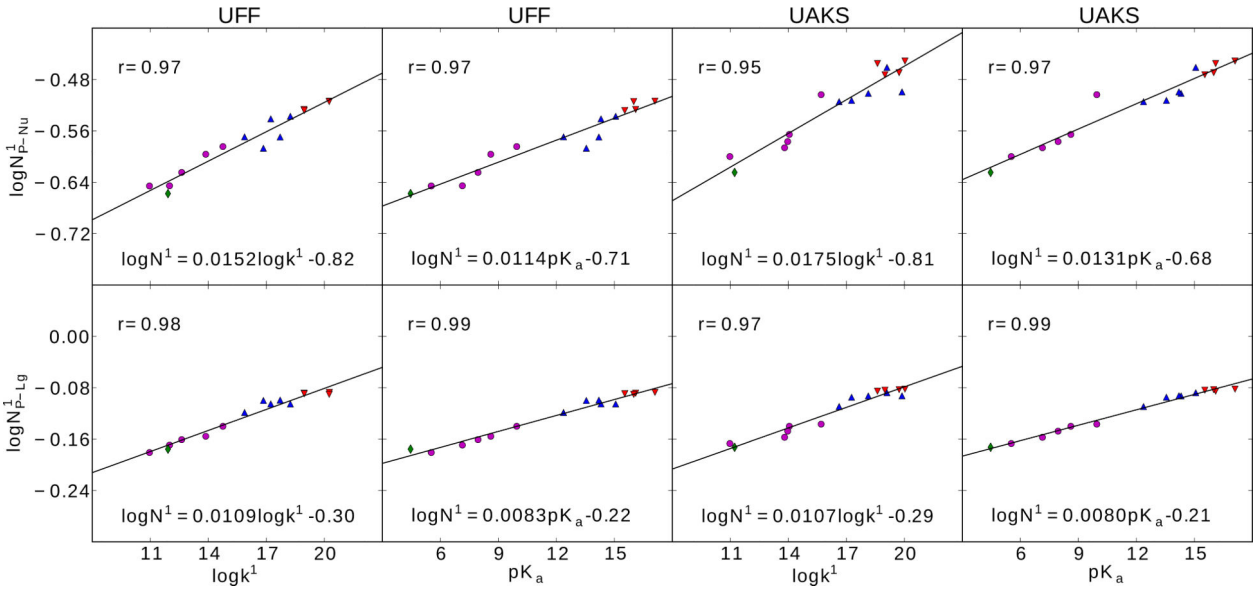


Figure 4.

Linear relationships between logarithmic bond orders ($logN^1_{P-Nu}$ and $logN^1_{P-Lg}$) of early transition states (TS1) and the corresponding rate constants (logk^1) and experimental pK_as of leaving groups. The calculation were performed using density-functional method with PCM solvation model and UFF (left) and UAKS (right) radii. Red down triangle, blue up triangle, purple round and green diamond symbols correspond to alkyl, hetero-alkyl, phenyl and acid leaving groups, respectively. Regression parameters for logN^1 are given as well as linear correlation coefficient for r.

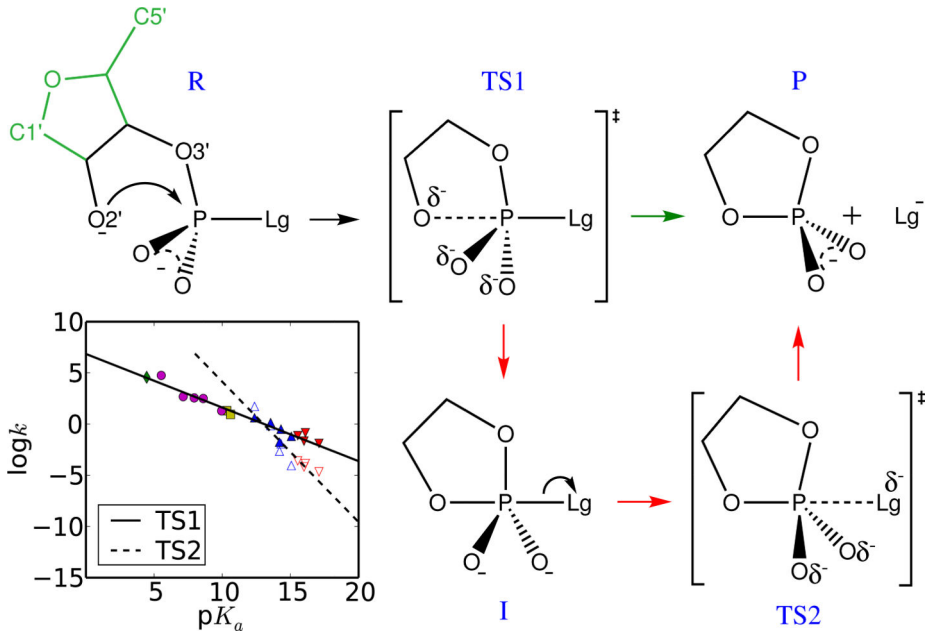


Figure 5.

For Table of Contents Only. Depending on the nature of the leaving group, reactions may proceed via a stepwise mechanism or through a single early TS1. Brønsted correlations can be used to distinguish these two mechanisms.

Table 1

Free energies (kcal/mol) of transition states ($\Delta G^{\neq}$), intermediates (ΔG_i) and products (ΔG) relative to reactants in RNA transesterification model reactions (Scheme 1) with different leaving groups as well as their experimental pK_as in water at 25°C.

	UFF				UAKS				Expt. pK_a
Leaving Group	$\Delta G_1^{\neq}$	ΔG_i	$\Delta G_2^{\neq}$	ΔG	$\Delta G_1^{\neq}$	ΔG_i	$\Delta G_2^{\neq}$	ΔG	
$(CH_3)_2CHO^-$	20.26	20.25	25.96	−1.56	20.03	17.66	23.82	−10.95	17.1
$CH_3CH_2CH_2O^-$	18.94	19.11	24.26	−1.40	18.61	15.61	22.73	−9.40	16.1
$CH_3CH_2O^-$	20.25	20.19	25.84	−0.04	19.73	16.80	23.17	−7.55	16
CH_3O^-	18.96	19.07	24.43	−0.59	19.00	15.94	22.36	−6.51	15.54
$HOCH_2CH_2O^-$	18.25	16.61	19.64	−6.44	19.10	17.01	22.99	−12.68	15.07
$ClCH_2CH_2O^-$	17.23	16.03	18.55	−6.94	18.13	15.44	20.08	−19.23	14.31
$FCH_2CH_2O^-$	17.72	16.52	20.11	−5.52	19.88	5.67	21.08	−17.44	14.2
CHCCH2O$^-$	16.85	15.27	17.75	−7.40	17.27	13.85	17.66	−16.66	13.55
$CF_3CH_2O^-$	15.87	12.49	13.59	−11.28	16.62	11.91	15.11	−21.17	12.37
$CH_3CH_2S^-$	14.42	···	···	−33.36	16.17	···	···	−30.04	10.61
CH_3S^-	14.94	···	···	−32.68	15.67	···	···	−29.45	10.33
PhO^-	14.76	···	···	−23.07	15.70	···	···	−23.59	9.95
3-CNPhO$^-$	13.87	···	···	−26.92	14.05	···	···	−28.18	8.61
4-CNPhO$^-$	12.62	···	···	−30.50	13.96	···	···	−29.78	7.95
p-NO_2PhO$^-$	11.99	···	···	−32.69	13.80	···	···	−30.65	7.14
2,3,5,6-F_4PhO$^-$	10.96	···	···	−36.10	10.97	···	···	−37.23	5.53
CH_3COO^-	11.91	···	···	−32.87	11.22	···	···	−34.95	4.46

$\Delta G_1^{\neq}$ 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),[71] except those of ethylene glycol and 2,3,5,6-tetrafluorophenol, which are taken from the CRC Handbook[72] and Bourne *et al.*,[73] respectively. The pK_a values 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.[58,74]

Table 2

Brønsted coefficients, β_{lg}, from experiments and calculation. The experimental coefficients, β_{lg}^1 and β_{lg}^2 were measured for transesterification reactions with aryl and alkyl leaving groups, respectively. The calculated coefficients, β_{lg}^1 and β_{lg}^2 were determined from theoretical rates derived from early and late transition states (TS1 and TS2), respectively, of the series of model reactions described in Scheme 1-2.

	Model System	β_{lg}^1	β_{lg}^2	β_{eq}	Temp (°C)	Ref./Model
Expt.	Uridine 3′-phosphate	−0.54	−1.28	···	25	33,56
	Uridine 3′-phosphate	−0.52	−1.34	···	25	55
	Uridine 3′-phosphate	···	−1.10	···	65	75
	2-hydroxypropyl phosphate[a]	−0.52	−1.09	···	80	58,55
	Guanosine 3′-phosphate	−0.38	···	···	35	76
	Phosphono monoanion	···	···	−1.74	25	57
Calc.	2-hydroxyethyl phosphate	−0.54	−2.04	−2.44	25	UFF
	2-hydroxyethyl phosphate	−0.54	−2.03	−2.44	80	UFF
	2-hydroxyethyl phosphate	−0.52	−1.37	−1.74	25	UAKS
	2-hydroxyethyl phosphate	−0.52	−1.32	−1.74	80	UAKS

[a]Linear Brønsted correlation coefficients (β_{lg}^1 and β_{lg}^2) were calculated from limited data[58] using adjusted pK_a values,[55] and may not be statistically reliable.

Table 3

Calculated and estimated experimental activation free energies $\Delta G^{\neq}$ (kcal/mol) for RNA transesterification model reactions with different leaving groups (Scheme 1). Experimental estimates were obtained from reaction rate constants using the Eyring equation.[77]

	Calc. $\Delta G^{\neq}$				Expt.[a] $\Delta G^{\neq}$			
	UFF		UAKS		PRPpOR	UpOR	UpOR	GpOR
Leaving Group	**25°C**	**80°C**	**25°C**	**80°C**	**80°C**	**25°C**	**65°C**	**35°C**
$(CH_3)_2CHO^-$	25.96	26.36	23.82	24.15	30.76	26.01	-	-
$CH_3CH_2CH_2O^-$	24.26	24.52	22.73	23.00	-	-	-	-
$CH_3CH_2O^-$	25.84	26.37	23.17	23.56	28.55	23.37	25.30	-
CH_3O^-	24.43	24.75	22.36	22.67	27.09	-	-	-
$HOCH_2CH_2O^-$	19.64	20.10	22.99	23.52	26.66	-	-	-
$ClCH_2CH_2O^-$	18.55	18.92	20.08	20.67	-	20.64	-	-
$FCH_2CH_2O^-$	20.11	20.51	21.08	21.75	-	-	-	-
CHCCH2O$^-$	17.75	18.23	17.66	18.11	-	-	-	-
$CF_3CH_2O^-$	15.87	16.37	16.62	17.14	-	-	-	-
$CH_3CH_2S^-$	14.42	14.64	16.17	16.46	-	-	-	-
CH_3S^-	14.94	15.32	15.67	15.94	-	-	-	-
PhO^-	14.76	15.14	15.70	16.24	22.62	15.98	-	24.22
3-CNPhO$^-$	13.87	14.42	14.05	14.53	-	-	-	-
4-CNPhO$^-$	12.62	13.03	13.96	14.55	-	-	-	-
p-NO_2PhO$^-$	11.99	12.41	13.80	14.50	20.12	13.46	-	22.81
2,3,5,6-F_4PhO$^-$	10.96	11.38	10.97	11.39	-	-		-
CH_3COO^-	11.91	12.31	11.22	11.58	-	-	-	-

[a]Experimental model systems are for the cyclization reaction of 2-hydroxypropyl phosphate (PRP-pOR),[58] uridine 3′-phosphate (UpOR)[33,56,75] and guanosine 3′-phosphate (GpOR).[76]

Table 4

All the linear relationships discovered in RNA transesterification model reactions.

Model	y	x	m	b	r
UFF	logk^1	pK_a	−0.54	7.50	−0.98
	logk^2	pK_a	−2.04	27.77	−0.96
	logK_{eq}	pK_a	−2.44	41.01	−0.93
	R^1_{P-Nu}	logk^1	−0.0190	2.67	−0.98
	R^1_{P-Nu}	pK_a	−0.0142	2.53	−0.98
	R^1_{P-Lg}	logk^1	−0.0096	1.90	−0.97
	R^1_{P-Lg}	pK_a	−0.0073	1.83	−0.98
	N^1_{P-Nu}	logk^1	0.0152	−0.82	0.97
	N^1_{P-Nu}	pK_a	0.0114	−0.71	0.97
	N^1_{P-Lg}	logk^1	0.0109	−0.30	0.98
	N^1_{P-Lg}	pK_a	0.0083	−0.22	0.99
UAKS	logk^1	pK_a	−0.52	6.85	−0.98
	logk^2	pK_a	−1.37	17.84	−0.93
	logK_{eq}	pK_a	−1.74	36.05	−0.95
	R^1_{P-Nu}	logk^1	−0.0216	2.66	−0.95
	R^1_{P-Nu}	pK_a	−0.0163	2.50	−0.98
	R^1_{P-Lg}	logk^1	−0.0090	1.89	−0.97
	R^1_{P-Lg}	pK_a	−0.0067	1.82	−0.98
	N^1_{P-Nu}	logk^1	0.0175	−0.81	0.95
	N^1_{P-Nu}	pK_a	0.0131	−0.68	0.97
	N^1_{P-Lg}	logk^1	0.0107	−0.29	0.97
	N^1_{P-Lg}	pK_a	0.0080	−0.21	0.99

These linear relationships are characterized by a linear equation, y=mx+b, together with the corresponding correlation coefficients, *r*. The unit of bond length R is Angstrom (Å).