Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 25 (2006) 423–433
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

QCRNA 1.0: A database of quantum calculations for RNA catalysis

Timothy J. Giese [a], Brent A. Gregersen [a], Yun Liu [a], Kwangho Nam [a], Evelyn Mayaan [a], Adam Moser [a], Kevin Range [a,1], Olalla Nieto Faza [a,b], Carlos Silva Lopez [a,b], Angel Rodriguez de Lera [b], Gijs Schaftenaar [e], Xabier Lopez [d], Tai-Sung Lee [a,f], George Karypis [c], Darrin M. York [a,*]

[a] *Department of Chemistry, University of Minnesota, 207 Pleasant St. SE, Minneapolis, MN 55455-0431, USA*
[b] *Departamento de Quimica Organica, Universidade de Vigo, Lagoas Marcosende, 36200 Vigo, Galicia, Spain*
[c] *Computer Science and Engineering Department, 4-192 EE/CS Building, 200 Union Street S.E., Minneapolis, MN 55455, USA*
[d] *Kimika Fakultatea, Euskal Herriko Unibertsitatea, P.K. 1072, 20080 Donostia, Euskadi, Spain*
[e] *Radboud University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands*
[f] *Consortium for Bioinformatics and Computational Biology, University of Minnesota, USA*

Received 20 October 2005; received in revised form 21 February 2006; accepted 25 February 2006
Available online 3 April 2006

Abstract

This work outlines a new on-line database of quantum calculations for RNA catalysis (*QCRNA*) available via the worldwide web at http://theory.chem.umn.edu/QCRNA. The database contains high-level density functional calculations for a large range of molecules, complexes and chemical mechanisms important to phosphoryl transfer reactions and RNA catalysis. Calculations are performed using a strict, consistent protocol such that a wealth of cross-comparisons can be made to elucidate meaningful trends in biological phosphate reactivity. Currently, around 2000 molecules have been collected in varying charge states in the gas phase and in solution. Solvation was treated with both the PCM and COSMO continuum solvation models. The data can be used to study important trends in reactivity of biological phosphates, or used as benchmark data for the design of new semiempirical quantum models for hybrid quantum mechanical/molecular mechanical simulations.

Keywords: Database; RNA catalysis; Reaction mechanism; Potential energy surfaces; Density functional theory

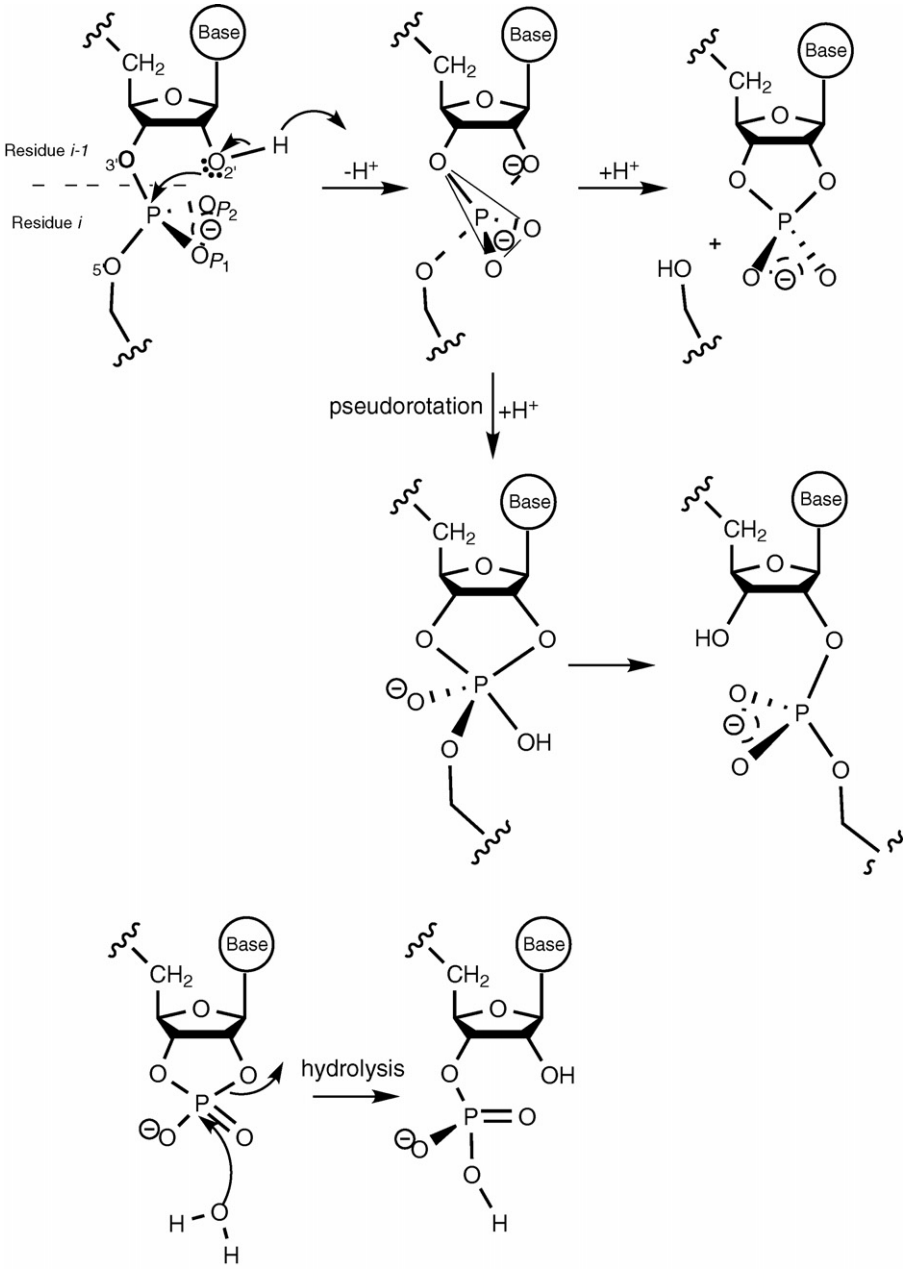
1. Introduction

An understanding of the molecular mechanisms of RNA catalysis is important from a fundamental biological perspective, and also for the design of new medical therapies that target genetic disorders [1–3], as well as the development of new biotechnology [4–11]. Much of what is known about phosphoryl transfer reactions have been derived from a wide spectrum of experimental data for reactions in solution ranging from the study of kinetics, pH and ion dependence, isotope and thio effects, along with information gained from spectroscopic techniques [12–15]. The understanding of the mechanisms of RNA catalysis [16,17] has been greatly facilitated by the availability of structural information derived from X-ray crystallography and related mechanistic studies of small prototype ribozymes such as the hammerhead [18,19], hairpin [20,21] and hepatitis delta virus [22,23] ribozymes that catalyze a cleavage transesterification reaction (Scheme 1) several orders of magnitude faster than non-enzymatic RNA molecules. Although no single experiment is able to conclusively provide a detailed mechanistic picture, a general consensus of the key factors that regulate biological phosphate reactivity and influence mechanism has emerged through analysis of the large accumulated body of experimental data.

Theoretical models have found widespread application by providing aid to the interpretation and refinement of experimental data. In principle, molecular simulations have the potential to unlock the catalytic *modus operandi* and unveil the mechanisms of allosteric control of catalytic RNA molecules. In order for modern theoretical chemistry methods to reliably

* Corresponding author. Tel.: +1 612 624 8042; fax: +1 610 626 7541.
E-mail address: york@chem.umn.edu (D.M. York).
[1] Present address: Department of Chemistry, Lock Haven University of Pennsylvania, Lock Haven, PA 17745, USA.

1093-3263/$ – see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2006.02.011



Scheme 1. Schematic illustrating typical transesterification (top), migration (middle) and hydrolysis (bottom) reactions that occur in RNA.

address large-scale problems of RNA catalysis, highly accurate and sufficiently fast quantum models are required that can be used with linear-scaling electronic structure methods [24,25] and hybrid quantum mechanical/molecular mechanical (QM/MM) simulations [26–28]. One approach is to develop so-called semiempirical quantum models that are specifically parameterized to obtain accurate results over a relevant range of chemical processes. In the case of RNA catalysis, the quantitative predictive capability of the semiempirical quantum models [29–32] must encompass a broad range of biological phosphate chemistry, and hence need to be parameterized against a large dataset of model compounds, complexes and phosphoryl transfer reactions calculated with high-level electronic structure methods.

Over the past two decades, biological phosphates and their reactions have been studied fairly extensively [13] with a variety of quantum chemical methods, both in the gas phase [33–37] and in solution [38–46], including studies of the nature of associative and dissociative paths [47–49]. However, as a dataset for training new semiempirical quantum models for RNA catalysis, available information collected from the literature is grossly inadequate in that:

(1) The literature data is incomplete in terms of the availability of detailed electronic, structural, and thermodynamic information.
(2) The literature data is inconsistent in that disparate theoretical/computational protocols have been employed

by different groups that severely limit meaningful quantitative cross-comparison.

(3) The range of biological phosphate model reactions found in the literature data is not sufficiently extensive or systematic to capture many of the phosphate reactivity trends important in RNA catalysis.

The purpose of this paper is to introduce a new on-line database of quantum calculations for RNA catalysis (*QCRNA*) [50]. The *QCRNA* database contains an extensive set of calculations of molecules, complexes, and chemical mechanisms involving a wide range of phosphoryl transfer reactions relevant to the study of RNA catalysis. The database was constructed using a strict computational protocol and level of theory that allows for a robust set of comparisons to be made and internal consistency to be maintained. This database can be used to identify key factors that influence biological phosphate reactivity [51–55], derive linear free energy relations [56–59] and train next-generation semiempirical quantum models [60–62] for linear-scaling electronic structure calculations [63–65] and hybrid QM/MM simulations [66–68] of RNA catalysis. The use of databases in quantum chemistry has gained considerable interest in recent years, and has been the focus of considerable effort and progress [69–71], although none to date has the same broad application to RNA catalysis as that of the present work.

2. Methods

The *QCRNA* database utilizes a strict computational protocol to determine optimized molecular structures, electronic structure properties and thermodynamic quantities including estimates of the solvation free energy and solvent-polarization using continuum models. Checks on both the wave function and nature of the stationary points are performed to insure reliability of the calculations. All calculations were performed using the GAUSSIAN 03 [72] software package. A summary of the *QCRNA* protocol is as follows:

- B3LYP/6-31++G (d,p) geometry optimization to a stationary point on the adiabatic potential energy surface;
- B3LYP/6-31++G (d,p) wave function stability check;
- B3LYP/6-31++G (d,p) vibrational frequency analysis (including enthalpic and free energy corrections to the energy at 298.15 K) and polarizability calculation;
- B3LYP/6-311++G (3df,2p) gas-phase single point energy refinement, evaluation of multipole moments, natural bond orbital (NBO) analysis, and ChelpG atomic charges;
- B3LYP/6-311++G (3df,2p) solution-phase single point energy refinement with COSMO solvation, evaluation of multipole moments, natural bond orbital (NBO) analysis, and ChelpG atomic charges;
- B3LYP/6-311++G (3df,2p) solution-phase single point energy refinement with PCM solvation, evaluation of multipole moments, natural bond orbital (NBO) analysis, and ChelpG atomic charges.

Details of the calculations in the gas phase and in solution are provided in Sections 2.1 and 2.2, respectively.

2.1. Gas-phase calculations

All structures were optimized in the gas phase with Kohn–Sham density functional theory (DFT) methods using the hybrid exchange functional of Becke [73,74] and the Lee, Yang, and Parr correlation functional [75] (B3LYP). Integrals involving the exchange-correlation potential used the default numerical integration mesh with a maximum of 75 radial shells and 302 angular quadrature points per shell pruned to approximately 7000 points per atom [76].

Geometry optimizations were performed in redundant internal coordinates with default convergence criteria [77] using the 6-31++G (d,p) basis set and the stability of the restricted closed shell Kohn–Sham determinant for each final structure were verified [78,79]. Frequency calculations at the optimized geometries were performed to establish the nature of all stationary points and used to calculate thermodynamic quantities. Electronic energies and other properties of the density, such as moments of the density and natural bond order (NBO) analysis [80] were further refined via single point calculations at the optimized geometries using the 6-311++G (3df,2p) basis set and the B3LYP hybrid density functional. All single point calculations were run with convergence criteria on the SCF wave function tightened to 10^{-8} au to ensure high precision for properties sensitive to the use of diffuse basis functions [76].

2.2. Solvation calculations

Solvation effects were treated by single point calculations based on the gas-phase optimized structures using the polarizable continuum model (PCM) [81–83] and a variation of the conductor-like screening model (COSMO) [84] as implemented in GAUSSIAN 03. The solvation free energy, ΔG_{sol}, is defined as:

$$\Delta G_{sol} = G_{aq} - G \quad (1)$$

where G and G_{aq} are the molecular free energies in the gas phase and in solution, respectively. In the present work, the approximation is made that the gas-phase geometry, entropy, and thermal corrections to the enthalpy do not change upon solvation. Within these approximations, the solvation energy is given by:

$$\Delta G_{sol} = (E[\psi_{sol}] + E_{sol}[\rho_{sol}]) - E[\psi_{gas}] \quad (2)$$

where $E[\psi_{gas}]$ and $E[\psi_{sol}]$ are the Kohn–Sham energy functionals that take as arguments the Kohn–Sham single-determinant wave function optimized in the gas phase (ψ_{gas}) and in solution (ψ_{sol}), and $E_{sol}[\rho_{sol}]$ is the solvation energy that takes as argument the polarized electron density in solution $\rho_{sol}(\mathbf{r})$ (which can be derived from ψ_{sol}).

In the PCM and COSMO models, the solvation energy functional, $E_{sol}[\rho_{sol}]$, can be written

$$E_{sol}[\rho] = \frac{1}{2}\left[\int \rho(\mathbf{r})\nu_{RF}(\mathbf{r})d^3r - \sum_{\alpha} Z_\alpha \nu_{RF}(\mathbf{R}_\alpha)\right] + G_{disp\text{-}rep} + G_{cav} \quad (3)$$

where $\nu_{RF}(\mathbf{r})$ is the solvent reaction-field potential, Z_α is the nuclear charge of atom α located at position $\mathbf{R}_\alpha$. $G_{disp\text{-}rep}$ and G_{cav} represent the dispersion–repulsion and cavitation contributions, respectively [82].

2.3. Thermodynamic quantities

The *QCRNA* database provides several thermodynamic quantities other than the internal electronic energy and are summarized below:

$$G = H - TS \quad (4)$$

$$H = U + RT \quad (5)$$

$$U = E_0 + E_{vib} + E_{rot} + E_{trans} \quad (6)$$

$$E_0 = (E_{elec} + E_{NN}) + E_{ZPV} = E + E_{ZPV} \quad (7)$$

where G, U, H, S, and T are the Gibbs free energy, internal energy, enthalpy, entropy, and temperature, respectively, R the universal gas constant, and E_{elec}, E_{NN}, E_{ZPV}, E_{vib}, E_{rot}, and E_{trans} are the electronic energy, nuclear–nuclear repulsion energy, zero-point vibrational energy, thermal vibrational energy correction, rotational, and translational energy components, respectively, and E is the adiabatic energy (i.e., $E = E_{elec} + E_{NN}$). The expression for the enthalpy (Eq. (5)) assumes the ideal gas law for 1 mol of particles. The internal energy and entropy were derived from standard statistical mechanical expressions for separable vibrational, rotational and translational contributions within the harmonic oscillator, rigid rotor, ideal gas/particle-in-a-box models in the canonical ensemble [85], and have been described in detail elsewhere [51]. Thermodynamic properties were calculated using unscaled vibrational frequencies and a standard state in the gas phase of 1 mol of particles at $T = 298.15$ K and 1 atm pressure.

The free energy in solution was calculated as a solvation free energy correction to the gas-phase free energy as:

$$G_{aq} = G + \Delta G_{sol} \quad (8)$$

where G and G_{aq} are the Gibbs free energy values in the gas phase and aqueous phase, respectively, and ΔG_{sol} is the solvation free energy (Eq. (1)) [81,86]. The standard state in solution was taken to be at 298.15 K and 1 M concentration.

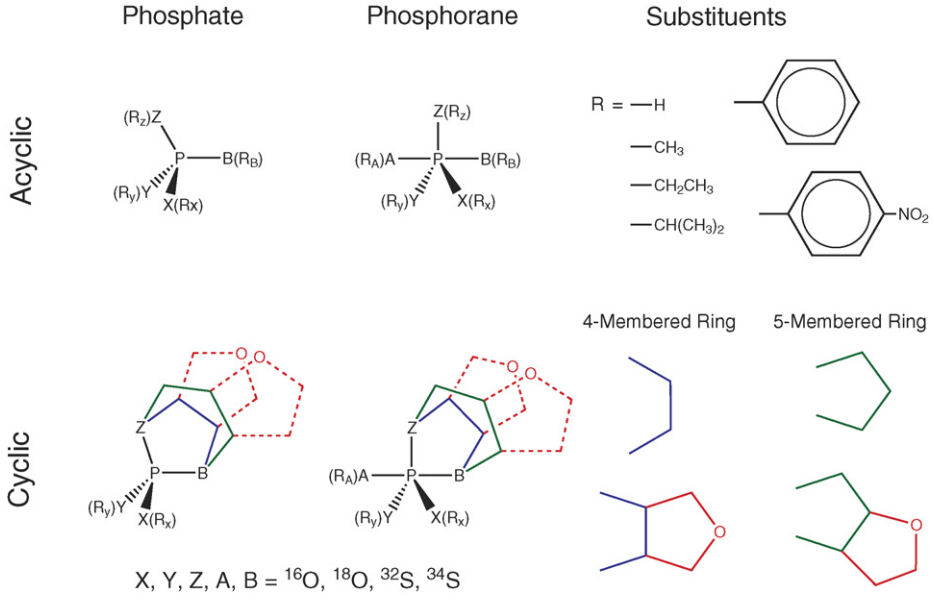
3. Results and discussion

The database contains three large bodies of information: molecule/complex (Section 3.2), chemical mechanism (Section 3.3), and potential energy surface (Section 3.4). Each of these bodies of data is individually searchable (Section 3.5), and can be manipulated with comparison tables (Section 3.6) and graphical interfaces (Section 3.7).

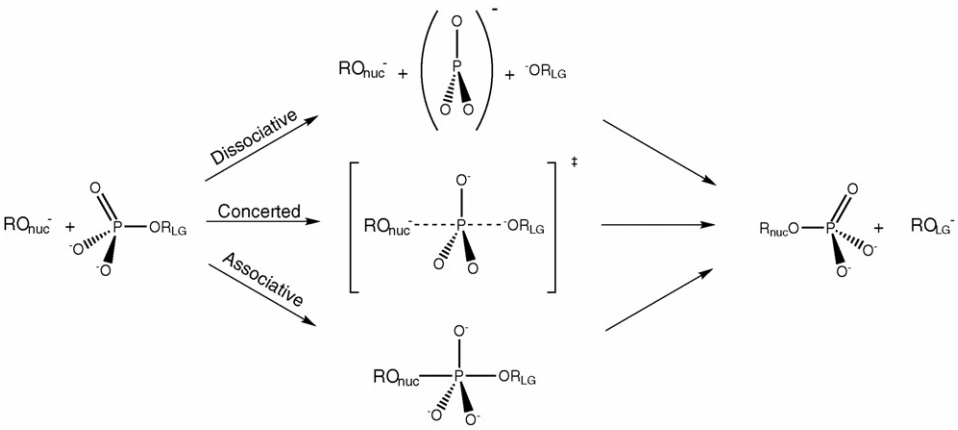
3.1. Overview

The main types of information that can be found in the database include:

- molecular properties;
- relative conformational energies;
- hydrogen bond energies;
- metal ion binding energies;
- proton affinities/gas-phase basicities;



Scheme 2. General schematic showing typical types of acyclic and cyclic phosphates and phosphoranes important to the study of phosphoryl transfer in RNA systems and in small model systems studied with experiment, including isotope effects and thio effects.



Scheme 3. General schematic showing idealized associative, concerted and dissociative mechanistic extremes for a typical phosphate monoester reaction.

- tautomerization energies;
- chemical mechanisms.

The molecular properties include the geometrical structure (either of a minimum or transition state), along with the host of electronic structure, thermodynamic, and solvation properties mentioned previously in Section 2. A main focus of the database involves the last item, *chemical mechanisms*. Phosphoryl transfer mechanisms relevant to RNA catalysis include not only reaction models of the native ribozyme systems (Scheme 1), but also of a host of other small model reactions with enhanced leaving groups and chemical modifications that have been well characterized experimentally. Of the mechanisms relevant to RNA catalysis contained in the database are those that involve:

- acyclic and cyclic phosphates;
- linear free energy relations that involve different nucleophiles and leaving groups;
- phosphate mono-, di- and triester systems;
- different protonation and charge states;

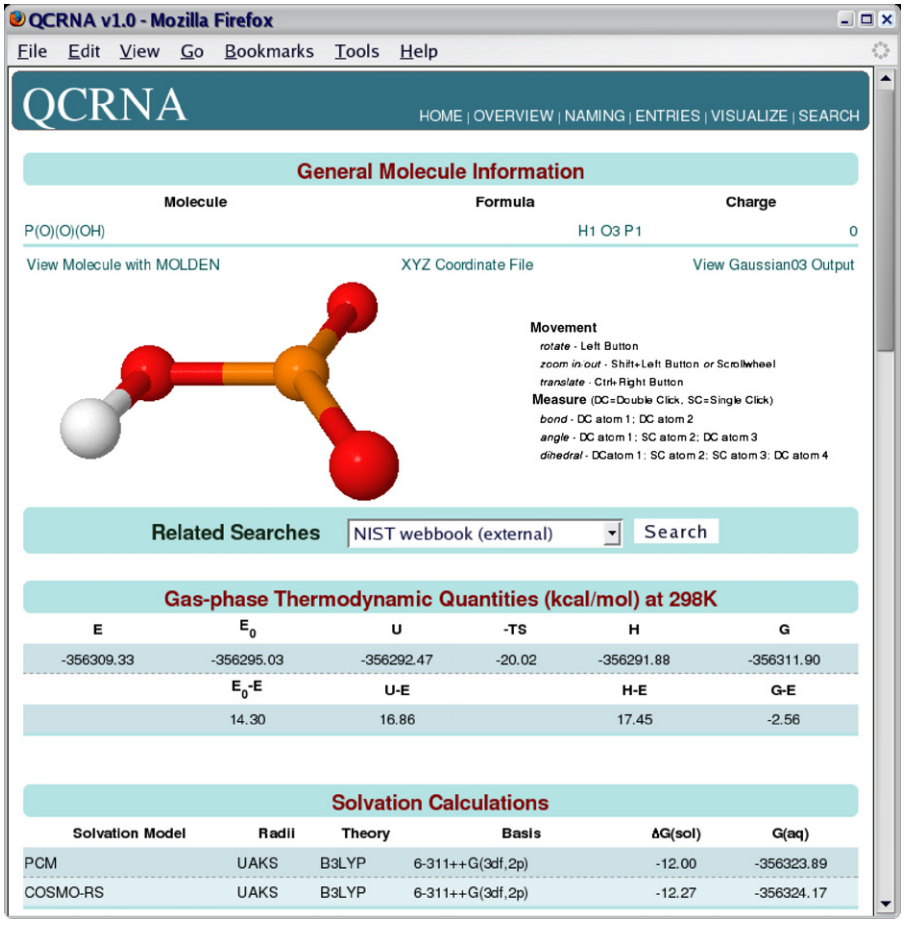


Fig. 1. A portion of the *QCRNA* database page displaying the gas and solvation energetics and thermochemistry data for protonated metaphosphate, P(O)(O)(OH).

- experimentally relevant thio effects;
- phosphorane pseudorotation reactions;
- metal-catalyzed reactions.

A general scheme of some of the acyclic and cyclic phosphates and phosphoranes contained in the database are illustrated in Scheme 2. The nature of the phosphates and phosphoranes involved in the reaction effect the chemical mechanism, which can proceed via a predominantly associative, concerted or dissociative pathways, as illustrated in Scheme 3 for an idealized phosphate monoester reaction.

3.2. Molecule/complex information

At the time of this writing, there are approximately 2000 molecules and complexes, more than half of which contain phosphorus, in the *QCRNA* database. Each molecule/complex contains the following energetic and thermodynamic information: E, E_0, $-TS$, H, G in the gas phase (Section 2.3), G_{aq} and ΔG_{sol} for both COSMO and PCM implicit water solvent (Section 2.2). Electronic properties are calculated in the gas phase and in solution treated with both the COSMO and PCM implicit solvent models. These properties include: the dipole moment components μ_x, μ_y, μ_z and magnitude $|\mu|$, eigenvalues of the quadrupole moment tensor Q_{xx}, Q_{yy}, Q_{zz}, the molecular orbital energies of the highest occupied and lowest unoccupied molecular orbitals HOMO and LUMO, respectively, the Koopman ionization potential IP, and related vertical excitation energy, Vertical E^* (under Koopman's approximation, i.e., Vertical $E^* = \epsilon_{LUMO} - \epsilon_{HOMO}$) In many cases, gas-phase and solution polarizability data is also available. Full structural and vibrational frequency information is also available in raw form or through the graphical interfaces (Section 3.7). An example portion of a screen-shot containing some of the molecule information is shown in Fig. 1.

3.3. Chemical mechanism information

There are currently over 250 chemical mechanisms in the *QCRNA* database. A chemical mechanism is a particular pathway leading from reactant to product. For each reaction mechanism, the structures of each stationary point along the reaction path have been calculated and collected. If a particular reaction contains more than one mechanistic path, these also

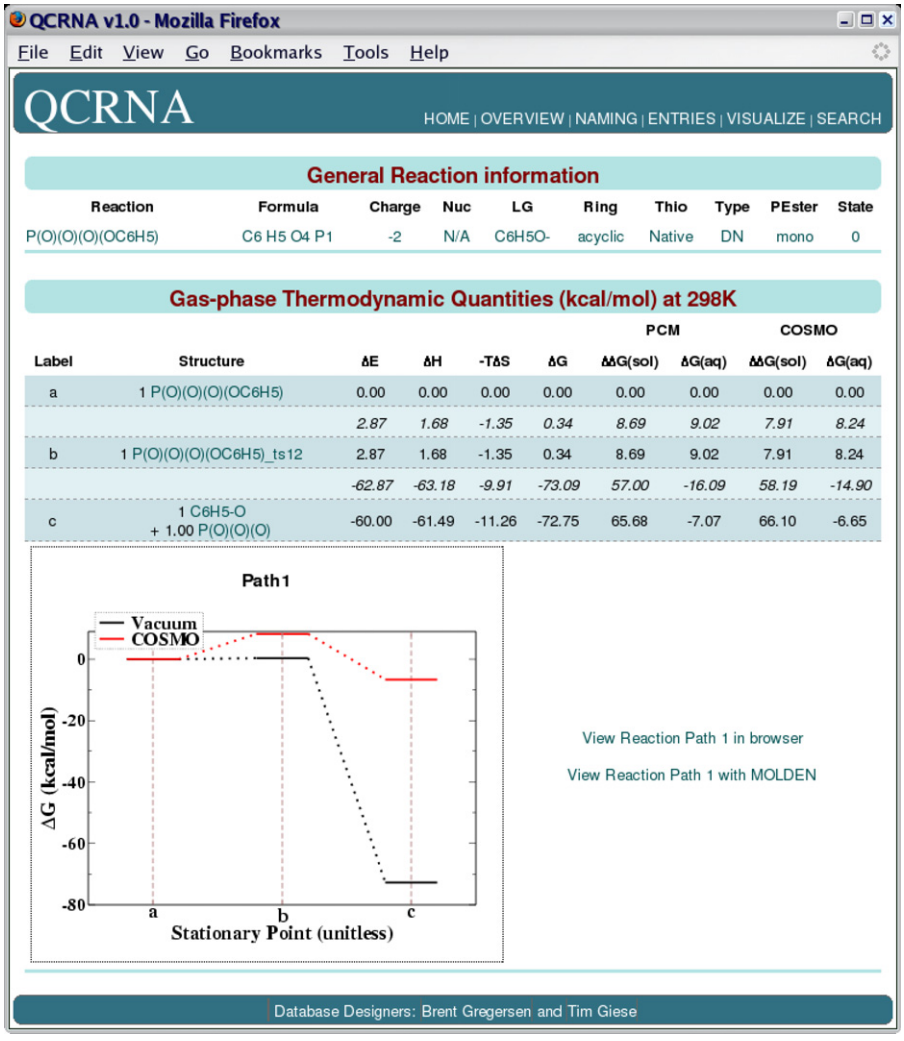


Fig. 2. Mechanism table for the dissociative path of phenyl phosphate dephosphorylation: $PO_3^- \cdots C_6H_5O^-$.

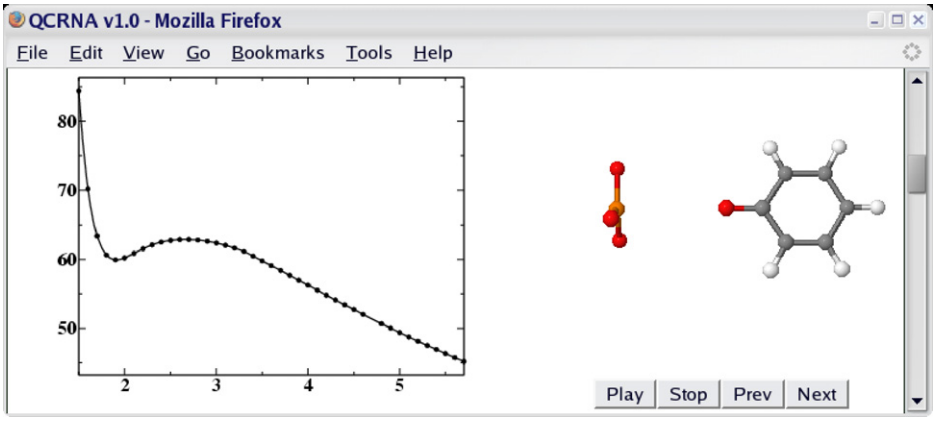


Fig. 3. Potential energy surface and animation for the dissociative path of phenyl phosphate dephosphorylation: $PO_3^- \cdots C_6H_5O^-$.

are collected. The same information described for the molecule/complex are available for each stationary point along a particular path via a link, and hence reaction energetics can be compared in the gas phase and in solution modeled by either COSMO or PCM.

This hierarchy of information lends to the construction of a mechanism table as a collection of path tables (Fig. 2), which contain a row for each step in the path and columns: ΔE, ΔH, $-T\Delta S$, ΔG, $\Delta\Delta G_{sol}$, and ΔG_{aq}, which are the relative differences in energy, enthalpy, entropy, Gibb's free energy, change in solvation free energy, and change in free energy in solution, respectively, between a step and a reference step.

3.4. Potential energy surface information

There are approximately 50 potential energy surfaces (PESs) currently in the *QCRNA* database. Each system contains *at least* a single potential energy surface; however, many

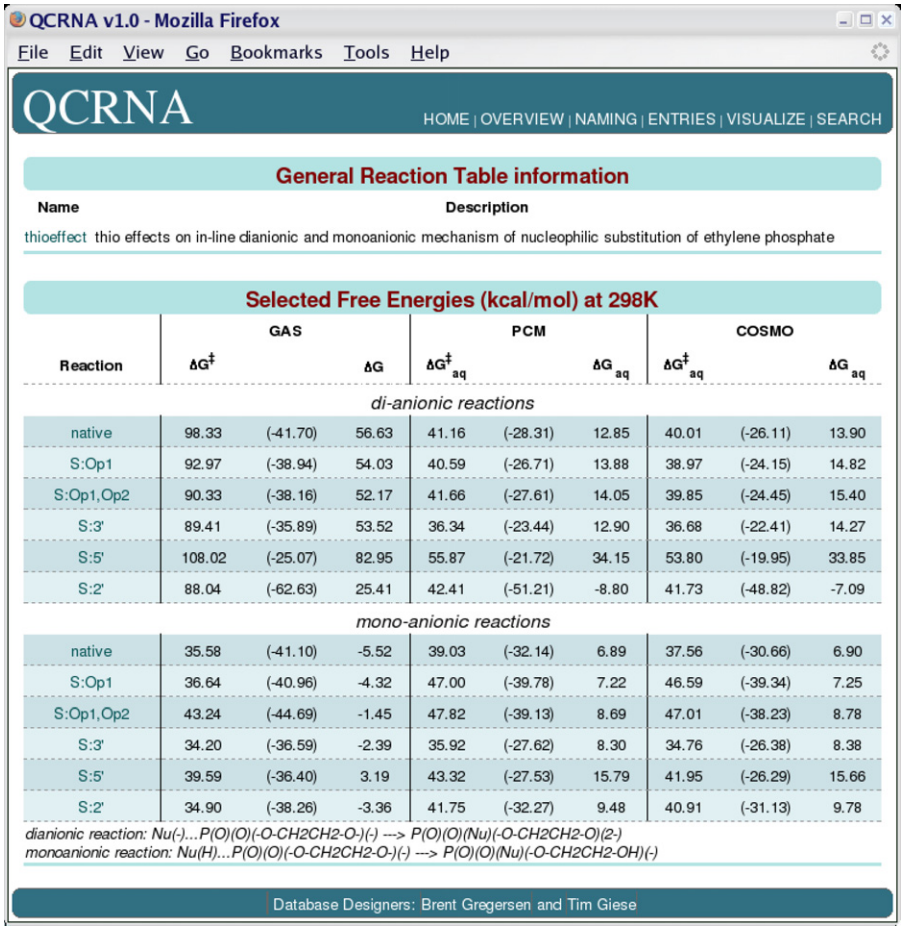


Fig. 4. A "reaction comparison table" comparing thio substitution effects on in-line dianionic and monoanionic mechanism of nucleophilic substitution of ethylene phosphate in the gas phase and in implicit solvent.

special cases exist that contain multiple potential energy surfaces corresponding to different basis sets and levels of theory, or that contain non-energetic surfaces, e.g., dipole and polarizability surfaces. Figures of the potential energy surface, interactive animations of the molecular coordinates (Fig. 3) and *XYZ* coordinate files for each point along the potential energy surface are available for each potential energy surface. A complete description of each system and the methods employed in the creation of the potential energy surfaces are provided for each system. Increasing the PES data for phosphoryl transfer mechanisms for use in developing new semiempirical quantum models for linear-scaling electronic structure and combined QM/MM applications is an area of active effort.

3.5. Search utilities

The size and complexity of the diverse set of phosphoryl transfer molecules, reactions and PESs makes it critical to have flexible search tools in order to find and access important data. The molecule/complex, chemical mechanism, and potential energy surface portions of the database are searchable by name, name substring, charge, chemical formula and by simple regular expressions. The simple regular expressions refer to the use of quantifying wildcards, e.g., an asterisk (*) is equivalent to ''zero-or-more'' and a plus (+) is equivalent to ''one-or-more''. For example, *C+H+* would return a list of all hydrocarbons in the database, whereas *H*O+P* would return all molecules which contained zero-or-more H atoms, one-or-more O atoms and exactly 1 P atom. Mechanisms contain the additional ability to search by nucleophile or leaving group, thio substitution, and other reaction descriptor indexes.

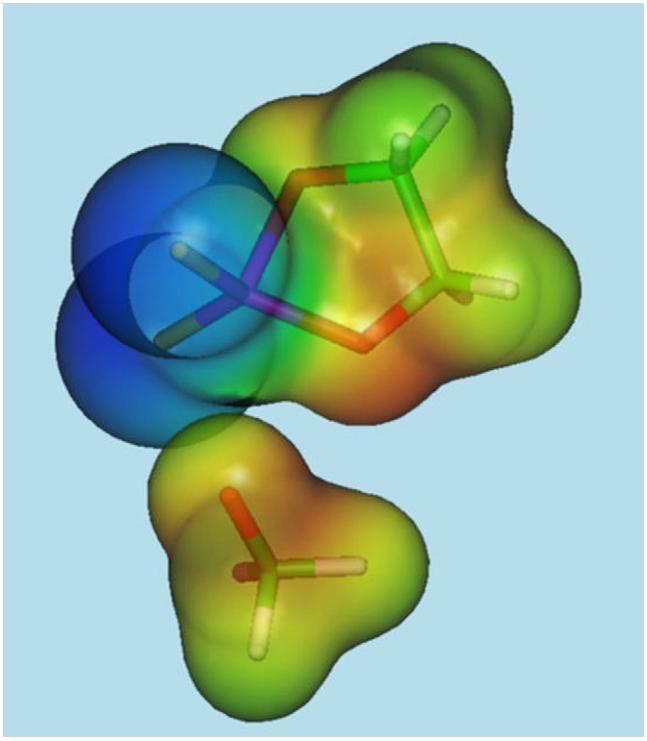


Fig. 5. MOLDEN picture of the electrostatic potential of the transition state for in-line dianionic methanolysis of dithio-substituted ethylene phosphate.

3.6. Comparison tables

The *QCRNA* database contains a set of tables comparing reactions of significant interest to phosphoryl transfer and RNA catalysis. The relational design of the database allows the user to quickly traverse from any of these tables to the mechanism information tables (Section 3.3), from which one can reach the information on individual molecules (Section 3.2). These tables include comparisons of thio effects on in-line dianionic and monoanionic mechanism of nucleophilic substitution of ethylene phosphate (Fig. 4), the effect of nucleophile on in-line dianionic and monoanionic mechanism of nucleophilic substitution of ethylene phosphate with sugar ring, the effect of nucleophile on in-line dianionic mechanism of nucleophilic substitution of dimethyl phosphate, the gas-phase proton affinity of common leaving groups in transphosphorylation reactions, the degree of esterification on pseudorotation of cyclic and acyclic phosphates, the thermodynamics of nucleic acid base-pair interactions and Mg^{2+} binding energies on biologically relevant phosphates. Reactions not contained in these tables can be constructed with a custom reaction builder that allows the user to construct individual reactions (not necessarily mechanisms) from molecules/complexes already in the database and compare relative thermodynamic values.

3.7. Graphical interfaces

Geometrical information for any molecule/complex can be retrieved by a link that returns Cartesian coordinates in *XYZ* format.[1] The *XYZ* file can be converted into other file formats, including internal coordinates, with the use of the OPEN BABEL software program [87]. A screen-shot of the molecule information page for protonated metaphosphate, PO_3H, is shown in Fig. 1.

Molecular structures can be viewed with the JMOL [88,89] or MOLDEN [90,91] programs as viewers for chemical MIME types. If the web browser is JAVA-enabled, then the JMOL software will automatically load as a web applet. Both programs allow the structure to be manipulated, i.e., rotated, scaled, and translated, and allow for measurement of internal coordinates, e.g., bond lengths, angles, and dihedral angles. Similarly, animations of the vibrational frequencies are available and can be viewed with either program. Using the provided output files, MOLDEN can also be used to construct electrostatic potential maps (Fig. 5).

[1] The *XYZ* format contains a 2-line header with the first line an integer number of atoms, *N*, and the second a comment line. The following *N* lines contain atom information, separated by white space: $atom_i$, X_i, Y_i, Z_i, for atom *i*, where ''atom'' is the atomic symbol, and ''*X*'', ''*Y*'', and ''*Z*'' are Cartesian coordinates.

4. Conclusion

This paper introduces the new on-line *QCRNA* database [50] used for collecting and analyzing high-level quantum chemical data related to phosphoryl transfer and RNA catalysis. The database captures a broad range of biological phosphate chemistry at a consistent level of theory that allows extensive cross-comparisons to be made. Calculations are performed in the gas phase and with solvation treated with the COSMO and PCM solvation methods. The data can be used to study important trends in reactivity of biological phosphates, or used as benchmark data for the design of new semiempirical quantum models for linear-scaling and hybrid QM/MM simulations of RNA catalysis. Future work will involve further extension of the database, enhancement of graphical tools, and implementation of methods that enable more powerful chemical classification, comparative modeling and data mining capabilities [92,93].

Acknowledgements

D.Y. is grateful for financial support provided by the National Institutes of Health (Grant GM62248), and the Army High Performance Computing Research Center (AHPCRC) under the auspices of the Department of the Army, Army Research Laboratory (ARL) under Cooperative Agreement number DAAD19-01-2-0014. O.N.F. and X.L. thank the Minnesota Supercomputing Institute for funding through Research Scholar awards, and T.-S.L. thanks the University of Minnesota for partial funding as part of the Consortium for Bioinformatics and Computational Biology initiative. D.Y., G.K., and T.-S.L. thank the IBM-Rochester Life Sciences Group for graduate funding (Y.L.) for this project. Computational resources were provided by the Minnesota Super-computing Institute.

References

[1] M.D. Hughes, M. Hussain, Q. Nawaz, P. Sayyed, S. Akhtar, The cellular delivery of antisense oligonucleotides and ribozymes, DDT 6 (2001) 313–315.
[2] N. Usman, L. Beigelman, J.A. McSwiggen, Hammerhead ribozyme engineering, Curr. Opin. Struct. Biol. 1 (1996) 527–533.
[3] E. Uhlmann, A. Peyman, Antisense oligonucleotides: a new therapeutic principle, Chem. Rev. 90 (4) (1990) 543–584.
[4] S.K. Silverman, Rube Goldberg goes (ribo)nuclear? Molecular switches and sensors made from RNA, RNA 9 (2003) 377–383.
[5] P.T. Sekella, D. Rueda, N.G. Walter, A biosensor for theophylline based on fluorescence detection of ligand-induced hammerhead ribozyme cleavage, RNA 8 (10) (2002) 1242–1252.
[6] G.A. Soukup, R.R. Breaker, Nucleic acid molecular switches, Trends Biotechnol. 17 (1999) 469–476.
[7] E. Puerta-Fernández, C. Romero-López, A. Barroso-delJesus, A. Berzal-Herranz, Ribozymes: recent advances in the development of RNA tools, FEMS Microbiol. Rev. 27 (1) (2003) 75–97.
[8] T.R. Cech, Ribozyme engineering, Curr. Opin. Struct. Biol. 2 (1992) 605–609.
[9] R.R. Breaker, In vitro selection of catalytic polynucleotides, Chem. Rev. 97 (1997) 371–390.
[10] R.R. Breaker, Engineered allosteric ribozymes as biosensor components, Curr. Opin. Biotech. 13 (2002) 31–39.
[11] A. Lescoute, E. Westhof, Riboswitch structures: purine ligands replace tertiary contacts, Chem. Biol. 12 (1) (2005) 10–13.
[12] D.M. Perreault, E.V. Anslyn, Unifying the current data on the mechanism of cleavage-transesterification of RNA, Angew. Chem. Int. Ed. 36 (1997) 432–450.
[13] D.-M. Zhou, K. Taira, The hydrolysis of RNA: from theoretical calculations to the hammerhead ribozyme-mediated cleavage of RNA, Chem. Rev. 98 (1998) 991–1026.
[14] M. Oivanen, S. Kuusela, H. Lönnberg, Kinetics mechanisms for the cleavage and isomerization of the phosphodiester bonds of RNA by Brønsted acids and bases, Chem. Rev. 98 (1998) 961–990.
[15] A.C. Hengge, Isotope effects in the study of phosphoryl and sulfuryl transfer reactions, Acc. Chem. Res. 35 (2002) 105–112.
[16] W.G. Scott, RNA catalysis, Curr. Opin. Struct. Biol. 8 (6) (1998) 720–726.
[17] Y. Takagi, Y. Ikeda, K. Taira, Ribozyme mechanisms, Top. Curr. Chem. 232 (2004) 213–251.
[18] W.G. Scott, J.B. Murray, J.R.P. Arnold, B.L. Stoddard, A. Klug, Capturing the structure of a catalytic RNA intermediate: the Hammerhead ribozyme, Science 274 (1996) 2065–2069.
[19] W.G. Scott, Biophysical and biochemical investigations of RNA catalysis in the hammerhead ribozyme, Q. Rev. Biophys. 32 (1999) 241–294.
[20] N.G. Walter, J.M. Burke, The hairpin ribozyme: structure, assembly and catalysis, Curr. Opin. Chem. Biol. 2 (1998) 24–30.
[21] P.B. Rupert, A.P. Massey, S.T. Sigurdsson, A.R. Ferré-D'Amaré, Transition state stabilization by a catalytic RNA, Science 298 (2002) 1421–1424.
[22] A. Ke, K. Zhou, F. Ding, J.H.D. Cate, J.A. Doudna, A conformational switch controls hepatitis delta virus ribozyme catalysis, Nature 429 (2004) 201–205.
[23] S. Das, J. Piccirilli, General acid catalysis by the hepatitis delta virus ribozyme, Nat. Chem. Biol. 1 (1) (2005) 45–52.
[24] S. Goedecker, Linear scaling electronic structure methods, Rev. Mod. Phys. 71 (1999) 1085–1123.
[25] W. Yang, J.M. Pérez-Jordá, Linear scaling methods for electronic structure calculation, in: P.R. von Schleyer (Ed.), Encyclopedia of Computational Chemistry, John Wiley and Sons, New York, 1998, pp. 1496–1513.
[26] A. Warshel, M. Levitt, Theoretic studies of enzymatic reactions: dielectric electrostatic and steric stabilization in the reaction of lysozyme, J. Mol. Biol. 103 (1976) 227–249.
[27] R.A. Friesner, M.D. Beachy, Quantum mechanical calculations on biological systems, Curr. Opin. Struct. Biol. 8 (1998) 257–262.
[28] M. Garcia-Viloca, J. Gao, M. Karplus, D.G. Truhlar, How enzymes work: analysis by modern rate theory and computer simulations, Science 303 (2004) 186–195.
[29] W. Thiel, Perspectives on Semiempirical Molecular Orbital Theory, in: I. Prigogine, S.A. Rice (Eds.), Adv. Chem. Phys., vol. 93, John Wiley and Sons, New York, 1996, pp. 703–757.
[30] T. Clark, Quo vadis semiempirical MO-theory? J. Mol. Struct. (Theochem.) 530 (2000) 1–10.
[31] P. Winget, C. Selçuki, A. Horn, B. Martin, T. Clark, Towards a ''next generation'' neglect of diatomic differential overlap based semiemperical molecular orbital technique, Theor. Chem. Acc. 110 (4) (2003) 254–266.
[32] W. Thiel, Semiempirical theories, in: S. Wilson (Ed.), Handbook of Molecular Physics and Quantum Chemistry, vol. 2, John Wiley and Sons, Chicester, 2003, pp. 487–502.
[33] C. Lim, M. Karplus, Nonexistence of dianionic pentacovalent intermediates in an ab initio study of the base-catalysed hydrolysis of ethylene phosphate, J. Am. Chem. Soc. 112 (1990) 5872–5873.
[34] T. Uchimaru, K. Tanabe, S. Nishikawa, K. Taira, Ab initio studies of a marginally stable intermediate in the base-catalyzed methanolysis of dimethyl phosphate and nonexistence of the stereoelectronically unfavorable transition state, J. Am. Chem. Soc. 113 (1991) 4351–4353.
[35] C. Lim, P. Tole, Concerted hydroxyl ion attack and pseudorotation in the base-catalyzed hydrolysis of methyl ethylene phosphate, J. Phys. Chem. 96 (1992) 5217–5219.
[36] J.M. Mercero, P. Barrett, C.W. Lam, J.E. Fowler, J.M. Ugalde, L.G. Pedersen, Quantum mechanical calculations on phosphate hydrolysis reactions, J. Comput. Chem. 21 (2000) 43–51.

[37] G.M. Arantes, B. Chaimovich, Thiolysis, Alcoholyses of phosphate tri- and monoesters with alkyl and aryl leaving groups. An ab initio study in the gas phase, J. Phys. Chem. A 109 (2005) 5625–5635.
[38] A. Dejaegere, C. Lim, M. Karplus, Dianionic pentacoordinate species in the base-catalyzed hydrolysis of ethylene and dimethyl phosphate, J. Am. Chem. Soc. 113 (1991) 4353–4355.
[39] A. Dejaegere, M. Karplus, Hydrolysis rate difference between cyclic and acyclic phosphate esters: solvation versus strain, J. Am. Chem. Soc. 115 (12) (1993) 5316–5317.
[40] P. Tole, C. Lim, New insights into the base-catalyzed hydrolysis of methyl ethylene phosphate, J. Phys. Chem. 97 (1993) 6212–6219.
[41] P. Tole, C. Lim, The significance of electrostatic effects in phospho-ester hydrolysis, J. Am. Chem. Soc. 116 (9) (1994) 3922–3931.
[42] N. Chang, C. Lim, An ab initio study of nucleophilic attack of trimethyl phosphate: factors influencing site reactivity, J. Phys. Chem. A 101 (1997) 8706–8713.
[43] X. Lopez, A. Dejaegere, M. Karplus, Solvent effects on the reaction coordinate of the hydrolysis of phosphates and sulfates: application of Hammond and anti-Hammond postulates to understand hydrolysis in solution, J. Am. Chem. Soc. 123 (2001) 11755–11763.
[44] X. Lopez, M. Schaefer, A. Dejaegere, M. Karplus, Theoretical evaluation of pK_a in phosphoranes: implications for phosphate ester hydrolysis, J. Am. Chem. Soc. 124 (18) (2002) 5010–5018.
[45] X. Chen, C.-G. Zhan, Theoretical determination of activation free energies for alkaline hydrolysis of cyclic and acyclic phosphodiesters in aqueous solution, J. Phys. Chem. A 108 (2004) 6407–6413.
[46] D. Xu, H. Guo, Y. Liu, D.M. York, Theoretical studies of dissociative phosphoryl transfer in interconversion of phosphoenolpyruvate to phosphonopyruvate: solvent effects, thio effects, and implications for enzymatic reactions, J. Phys. Chem. B 109 (2005) 13827–13834.
[47] J. Florián, A. Warshel, A fundamental assumption about OH-attack in phosphate ester hydrolysis is not fully justified, J. Am. Chem. Soc. 119 (1997) 5473–5474.
[48] J. Florián, A. Warshel, Phosphate ester hydrolysis in aqueous solution: associative versus dissociative mechanisms, J. Phys. Chem. B 102 (1998) 719–734.
[49] C.-H. Hu, T. Brinck, Theoretical studies of the hydrolysis of the methyl phosphate anion, J. Phys. Chem. A 103 (1999) 5379–5386.
[50] Qcrna http://theory.chem.umn.edu/Database/QCRNA.
[51] K. Range, M.J. McGrath, X. Lopez, D.M. York, The Structure, Stability of biological metaphosphate, phosphate, and phosphorane compounds in the gas phase and in solution, J. Am. Chem. Soc. 126 (2004) 1654–1665.
[52] E. Mayaan, K. Range, D.M. York, Structure and binding of Mg(II) ions and di-metal bridge complexes with biological phosphates and phosphoranes, J. Biol. Inorg. Chem. 9 (7) (2004) 807–817.
[53] C.S. López, O.N. Faza, B.A. Gregersen, X. Lopez, A.R. de Lera, D.M. York, Pseudorotation of natural and chemically modified biological phosphoranes: implications for RNA catalysis, Chem. Phys. Chem. 5 (2004) 1045–1049.
[54] C.S. López, O.N. Faza, A.R. de Lera, D.M. York, Pseudorotation barriers of biological oxyphosphoranes: a challenge for simulations of ribozyme catalysis, Chem. Eur. J. 11 (2005) 2081–2093.
[55] Y. Liu, X. Lopez, D.M. York, Kinetic isotope effects on thio-substituted biological phosphoryl transfer reactions from density-functional theory, Chem. Commun. 31 (2005) 3909–3911.
[56] S.A. Khan, A.J. Kirby, The reactivity of phosphate esters. multiple structure–reactivity correlations for the reactions of triesters with nucleophiles, J. Chem. Soc. B (1970) 1172–1182.
[57] W.P. Jencks, A primer for the Bema Hapothle. An empirical approach to the characterization of changing transition-state structures, Chem. Rev. 85 (6) (1985) 511–527.
[58] A. Warshel, T. Schweins, M. Fothergill, Linear free energy relationships in enzymes. Theoretical analysis of the reaction of tyrosyl-tRNA synthetase, J. Am. Chem. Soc. 116 (1994) 8437–8442.
[59] J. Florián, J. Åqvist, A. Warshel, On the reactivity of phosphate monoester dianions in aqueous solution: Brønsted linear free-energy relationships do not have an unique mechanistic interpretation, J. Am. Chem. Soc. 120 (1998) 11524–11525.
[60] X. Lopez, D.M. York, Parameterization of semiempirical methods to treat nucleophilic attacks to biological phosphates: AM1/d parameters for phosphorus, Theor. Chem. Acc. 109 (2003) 149–159.
[61] T.J. Giese, E.C. Sherer, C.J. Cramer, D.M. York, A semiempirical quantum model for hydrogen-bonded nucleic acid base pairs, J. Chem. Theory Comput. 1 (6) (2005) 1275–1285.
[62] B.A. Gregersen, T.J. Giese, Y. Liu, E. Mayaan, K. Nam, K. Range, D.M. York, Simulations of phosphoryl transfer reactions using multi-scale quantum models, in: K.J. Naidoo, M. Hann, J. Gao, M. Field, J. Brady (Eds.), Modelling Molecular Structure and Reactivity in Biological Systems, Royal Society Special Volume: WATOC 2005, Elsevier Science, Amsterdam, 2006 (Chapter 1).
[63] J. Khandogin, D.M. York, Quantum mechanical characterization of nucleic acids in solution: a linear-scaling study of charge fluctuations in DNA and RNA, J. Phys. Chem. B 106 (2002) 7693–7703.
[64] J. Khandogin, K. Musier-Forsyth, D.M. York, Insights into the regioselectivity and RNA-binding affinity of HIV-1 nucleocapsid protein from linear-scaling quantum methods, J. Mol. Biol. 330 (2003) 993–1004.
[65] J. Khandogin, D.M. York, Quantum descriptors for biological macromolecules from linear-scaling electronic structure methods, Proteins 56 (2004) 724–737.
[66] B.A. Gregersen, X. Lopez, D.M. York, Hybrid QM/MM study of thio effects in transphosphorylation reactions, J. Am. Chem. Soc. 125 (2003) 7178–7179.
[67] B.A. Gregersen, X. Lopez, D.M. York, Hybrid QM/MM study of thio effects in transphosphorylation reactions: the role of solvation, J. Am. Chem. Soc. 126 (2004) 7504–7513.
[68] K. Nam, J. Gao, D.M. York, An efficient linear-scaling Ewald method for long-range electrostatic interactions in combined QM/MM calculations, J. Chem. Theory Comput. 1 (1) (2005) 2–13.
[69] B. Beck, A. Horn, J.E. Carpenter, T. Clark, Enhanced 3D-databases: a fully electrostatic database of AM1-optimized structures, J. Chem. Inf. Comput. Sci. 38 (1998) 1214–1217.
[70] Y.-P. Liu, K. Kim, B.J. Berne, R.A. Friesner, S.W. Rick, Constructing ab initio force fields for molecular dynamics simulations, J. Chem. Phys. 108 (1998) 4739–4755.
[71] Y. Zhao, D.G. Truhlar, Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions, J. Phys. Chem. A 109 (2005) 5656–5667.
[72] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, 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, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M. W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.
[73] A.D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, Phys. Rev. A. 38 (1988) 3098–3100.
[74] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (7) (1993) 5648–5652.
[75] C. Lee, W. Yang, R.G. Parr, Development of the Colle-Savetti correlation energy formula into a functional of the electron density, Phys. Rev. B. 37 (1988) 785–789.
[76] Æ. Frisch, M.J. Frisch, Gaussian 98 User's Reference, second ed., Gaussian, Inc., Pittsburgh, PA, 1999.
[77] C. Peng, P.Y. Ayala, H.B. Schlegel, M.J. Frisch, Using redundant internal coordinates to optimize equilibrium geometries and transition states, J. Comput. Chem. 17 (1996) 49–56.

[78] R. Bauernschmitt, R. Ahlrichs, Stability analysis for solutions of the closed shell Kohn–Sham equation, J. Chem. Phys. 104 (1996) 9047–9052.
[79] R. Seeger, J.A. Pople, Self-consistent molecular orbital methods. XVIII. Constraints and stability in Hartree-Fock theory, J. Chem. Phys. 66 (7) (1977) 3045–3050.
[80] A.E. Reed, R.B. Weinstock, F. Weinhold, Natural population analysis, J. Chem. Phys. 83 (1985) 735–746.
[81] J. Tomasi, M. Persico, Molecular interaction in solution: an overview of methods based on continuous distributions of the solvent, Chem. Rev. 94 (1994) 2027–2094.
[82] M. Cossi, V. Barone, R. Cammi, J. Tomasi, Ab initio study of solvated molecules: a new implementation of the polarizable continuum model, Chem. Phys. Lett. 255 (1996) 327–335.
[83] T. Mineva, N. Russo, E. Sicilia, Solvation effects on reaction profiles by the polarizable continuum model coupled with the Gaussian density functional method, J. Comput. Chem. 19 (3) (1998) 290–299.
[84] V. Barone, M. Cossi, Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model, J. Phys. Chem. A 102 (1998) 1995–2001.
[85] C.J. Cramer, Essentials of Computational Chemistry: Theories and Models, second ed., John Wiley & Sons, Chichester, England, 2002.
[86] C.J. Cramer, D.G. Truhlar, Implicit solvation models: equilibria, structure, spectra, and dynamics, Chem. Rev. 99 (8) (1999) 2161–2200.
[87] M. Banck, F. Bresciani, J. Bréfort, A. Clark, J. Corkery, V. Favre-Nicolin, F. Fontaine, M. Gillies, R. Gillilan, B. Goldman, T. Hassinen, B. Herger, G. Hutchison, S. Kebekus, E. Kruus, E. Leitl, D. Mathog, C. Morley, P. Murray-Rust, A. Nicholls, S. Patchkovskii, S. Reith, L. Richard, R. Sayle, A. Shah, M. Stahl, B. Tolbert, P. Walters, P. Wolinski, J. Wegner, Open Babel, http://openbabel.sourceforge.net/ (2005).
[88] Jmol, http://www.jmol.org.
[89] E. Willighagen, M. Howard, Jmol as 3D Viewer for CDK, CDK News 2 (1) (2005) 17–20.
[90] G. Schaftenaar, J. H. Noordik, Molden, http://www.cmbi.ru.nl/molden/molden.html.
[91] G. Schaftenaar, J.H. Noordik, Molden: a pre- and post-processing program for molecular and electronic structures, J. Comput. Aided Mol. Des. 14 (2) (2000) 123–134.
[92] M. Deshpande, M. Kuramochi, N. Wale, G. Karypis, Frequent substructure-based approaches for classifying chemical compounds, IEEE Trans. Knowl. Data Eng. 20 (2005) 1–15.
[93] J. Wang, G. Karypis, HARMONY: efficiently mining the best rules for classification SIAM 2005 Data Mining Conference, 2005, p. 1.