J|A|C|S
ARTICLES
Published on Web 01/24/2004

The Structure and Stability of Biological Metaphosphate, Phosphate, and Phosphorane Compounds in the Gas Phase and in Solution

Kevin Range,[†] Matthew J. McGrath,[†] Xabier Lopez,[‡] and Darrin M. York*[,†]

Contribution from the Department of Chemistry, University of Minnesota, 207 Pleasant Street SE, Minneapolis, Minnesota 55455-0431, and Kimika Fakultatea, Euskal Herriko Unibertsitatea, P.K. 1072, 20080 Donostia, Euskadi, Spain

Received April 14, 2003; E-mail: york@chem.umn.edu

Abstract: Density functional calculations of a series of metaphosphates, acyclic and cyclic phosphates and phosphoranes relevant to RNA catalysis are presented. Solvent effects calculated with three well-established solvation models are analyzed and compared. The structure and stability of the compounds are characterized in terms of thermodynamic quantities for isomerization and ligand substitution reactions, gas-phase proton affinities, and microscopic solution pK_a values. The large dataset of compounds allows the estimation of bond energies to determine the relative strengths of axial and equatorial P−O phosphorane single bonds and P−O single and double bonds in metaphosphates and phosphates. The relative apicophilicty of hydroxyl and methoxy ligands in phosphoranes are characterized. The results presented here provide quantitative insight into RNA catalysis and serve as a first step toward the construction of a high-level quantum database for development of new semiempirical Hamiltonian models for biological reactions

1. Introduction

The study of biological phosphate chemistry is of key importance to understanding many cellular processes such as the hydrolysis of phosphates involved in transcription, cell signaling, and respiration.[1−3] Of particular interest is the study of the molecular mechanisms whereby RNA can catalyze fairly complicated reactions.[1−7] The unraveling of the details of the action of RNA enzymes, or "ribozymes", would allow deeper understanding of biological processes and provide valuable insight for the design of therapeutics that target viral or genetic disease[8] and new biotechnology such as RNA chips[9] or allosteric molecular switches in nanodevices.[10]

Quantum chemical methods have provided considerable insight into the behavior of model phosphates. Seminal work has been done by Florian, Warshel, and co-workers,[11−14] the group of Karplus and co-workers,[15−21] Lim et al.,[22−28] Zhou, Taira, and co-workers[3,29−33] Wladkowski and collaborators,[34,35] and many others.[36−41] The study of small phosphate models is

[†] University of Minnesota.
[‡] Euskal Herriko Unibertsitatea.

(1) Perreault, D. M.; Anslyn, E. V. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 432−450.
(2) Oivanen, M.; Kuusela, S.; Lönnberg, H. *Chem. Rev.* **1998**, *98*, 961−990.
(3) Zhou, D.-M.; Taira, K. *Chem. Rev.* **1998**, *98*, 991−1026.
(4) Bashkin, J. K. *Chem. Rev.* **1998**, *98*, 415−416.
(5) Kuimelis, R. G. *Chem. Rev.* **1998**, *98*, 1027−1044.
(6) Carola, C.; Eckstein, F. *Curr. Opin. Chem. Biol.* **1999**, *3*, 274−283.
(7) Scott, W. G. *Curr. Opin. Struct. Biol.* **2000**, *3*, 705−709.
(8) Novina, C. D.; Murray, M. F.; Dykxhoorn, D. M.; Beresford, P. J.; Riess, J.; Lee, S.-K.; Collman, R. G.; Lieberman, J.; Shankar, P.; Sharp, P. A. *Nat. Med.* **2002**, *8*, 681−686.
(9) Yeakley, J. M.; Fan, J.-B.; Doucet, D.; Luo, L.; Wickham, E.; Ye, Z.; Chee, M. S.; Fu, X.-D. *Nat. Biotechnol.* **2002**, *20*, 353−358.
(10) Soukup, G. A.; Breaker, R. R. *Trends Biotechnol.* **1999**, *17*, 469−476.
(11) Florián, J.; Warshel, A. *J. Am. Chem. Soc.* **1997**, *119*, 5473−5474.
(12) Florián, J.; ×b3Trajbl, M.; Warshel, A. *J. Am. Chem. Soc.* **1998**, *120*, 7959−7966.
(13) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 719−734.
(14) Åqvist, J.; Kolmodin, K.; Florian, J.; Warshel, A. *Chem. Biol.* **1999**, *6*, R71−R80.
(15) Dejaegere, A.; Lim, C.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 4353−4355.
(16) Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **1993**, *115*, 5316−5317.
(17) Dejaegere, A.; Liang, X. L.; Karplus, M. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 1763−1767.
(18) Lopez, X.; Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **1999**, *121*, 5548−5558.
(19) Lopez, X.; Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **2001**, *123*, 11755−11763.
(20) Lopez, X.; Schaefer, M.; Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **2002**, *124*, 5010−5018.
(21) Lopez, X.; York, D. M.; Dejaegere, A.; Karplus, M. *Int. J. Quantum Chem.* **2002**, *86*, 10−26.
(22) Lim, C.; Karplus, M. *J. Am. Chem. Soc.* **1990**, *112*, 5872−5873.
(23) Lim, C.; Tole, P. *J. Am. Chem. Soc.* **1992**, *114*, 7245−7252.
(24) Lim, C.; Tole, P. *J. Phys. Chem.* **1992**, *96*, 5217−5219.
(25) Tole, P.; Lim, C. *J. Phys. Chem.* **1993**, *97*, 6212−6219.
(26) Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **1998**, *1120*, 4450−4458.
(27) Chang, N.; Lim, C. *J. Phys. Chem. A* **1997**, *101*, 8706−8713.
(28) Chang, N.; Lim, C. *J. Am. Chem. Soc.* **1998**, *120*, 2156−2167.
(29) Uchimaru, T.; Tanabe, K.; Nishikawa, S.; Taira, K. *J. Am. Chem. Soc.* **1991**, *113*, 4351−4353.
(30) Taira, K.; Uchimaru, T.; Storer, J. W.; Yliniemela, A.; Uebayasi, M.; Tanabe, K. *J. Org. Chem.* **1993**, *58*, 3009−3017.
(31) Yliniemela, A.; Uchimaru, T.; Kazutoshi, T.; Taira, K. *J. Am. Chem. Soc.* **1993**, *115*, 3032−3033.
(32) Uchimaru, T.; Tsuzuki, S.; Storer, J. W.; Tanabe, K.; Taira, K. *J. Org. Chem.* **1994**, *59*, 1835−1843.
(33) Uchimaru, T.; Stec, W. J.; Tsuzuki, S.; Hirose, T.; Tanabe, K.; Taira, K. *Chem. Phys. Lett.* **1996**, *263*, 691−696.
(34) Wladkowski, B. D.; Krauss, M.; Stevens, W. J. *J. Am. Chem. Soc.* **1995**, *117*, 10537−10545.
(35) Jones, G. A.; Carpenter, B. K.; Paddon-Row, M. N. *J. Am. Chem. Soc.* **1998**, *120*, 5488−5498.
(36) Jayaram, B.; Mezei, M.; Beveridge, D. *J. Comput. Chem.* **1987**, *8*, 917−942.
(37) Jayaram, B.; Mezei, M.; Beveridge, D. *J. Am. Chem. Soc.* **1988**, *110*, 1691−1694.

10.1021/ja0356277 CCC: $27.50

an important first step toward a detailed understanding of RNA catalysis at the atomic level.

Application of theoretical methods to RNA catalysis is extremely challenging. The polyanionic nature of RNA amplifies the importance of rigorous treatment of electrostatic interactions, polarization and other quantum many-body effects, interaction with solvent, monovalent, and divalent ions. Moreover, RNA is considerably flexible, and consequently, simulations typically require long equilibration and sampling times relative to those for proteins. These factors combine to make the accurate theoretical study of RNA catalysis particularly challenging and largely beyond the reach of many conventional techniques and models.

A promising approach is to use hybrid quantum mechanical/molecular mechanical (QM/MM) methods where a small part of the system is treated quantum mechanically while the remainder of the system is modeled by a classical empirical force field.[42] The quantum mechanical method should be sufficiently fast such that simulations can be performed on meaningful time scales for the properties being studied. Consequently, the design of new quantum models that are both highly accurate for the specific chemical problem of interest *and* sufficiently efficient for long-time simulations are of tremendous value.

In the present contribution, quantum results for a large dataset of biological metaphosphate, acyclic and cyclic phosphate and phosphorane compounds are presented. These molecules are relevant for the study of phosphate hydrolysis reactions in solution and in ribozymes. The dataset is analyzed in terms of molecular structures, thermodynamic quantities for isomerization and ligand substitution reactions, solvation effects, proton affinities and microscopic pK_a values, and bond energies. The results presented here provide valuable insight into phosphate hydrolysis mechanisms and a first step in the construction of a quantum database for RNA catalysis from which new quantum models, capable of being applied efficiently in molecular simulations, can be derived.

2. Methods

2.1. Gas-Phase Calculations. All the structures were optimized in the gas phase with Kohn−Sham density functional theory (DFT) methods using the hybrid exchange functional of Becke[43,44] and the Lee, Yang, and Parr correlation functional[45] (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.[46] Geometry optimizations were done in redundant internal coordinates with default convergence criteria,[47] and stability conditions of the restricted closed shell Kohn−Sham determinant for each final structure were verified.[48,49] Frequency calculations were performed to establish the nature of all stationary points and to allow evaluation of thermodynamic quantities such as the zero-point vibrational energy, and thermal vibrational contributions to the enthalpy, entropy and Gibbs free energy.

The geometry optimization and frequency calculations were performed using the 6-31++G(d,p) basis set. This basis set is similar to (slightly larger than) that used for geometries and frequencies in the G2 method,[50,51] which typically yield atomization energies, ionization energies, and relative energies within 1 kcal/mol of experimental values.

Electronic energies and other properties of the density, such as dipole moments, were obtained via single-point calculations at the optimized geometries using the 6-311++G(3df,2p) basis set and the B3LYP hybrid density functional. This protocol for obtaining the geometry and energy is designated by the abbreviated notation B3LYP/6-311++G(3df,2p)//B3LYP/6-31++G(d,p). Single-point calculations were run with "tight" convergence criteria[46] to ensure high precision for properties sensitive to the use of diffuse basis functions. All electronic structure calculations were performed with the GAUSSIAN98 suite of programs.[52]

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)[53−55] and a variation of the conductor-like screening model (COSMO)[56] as implemented in GAUSSIAN98 and the SM5.42R solvation model[57,58] as implemented in MN-GSM.[59]

The solvation free energy, ΔG_{sol}, is defined as

$$\Delta G_{sol} = G_{aq} - G_{gas} \tag{1}$$

where G_{gas} and G_{sol} 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. The practical reason for introducing this approximation resides in the difficulty and considerable computational cost associated with obtaining stationary points and Hessians with the boundary element solvation methods. Within these approximations, the solvation energy is given by

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

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

(38) Hu, C.-H.; Brinck, T. *J. Phys. Chem. A* **1999**, *103*, 5379−5386.
(39) Mercero, J. M.; Barrett, P.; Lam, C. W.; Fowler, J. E.; Ugalde, J. M.; Pedersen, L. G. *J. Comput. Chem.* **2000**, *21*, 43−51.
(40) Wilkie, J.; Gani, D. *J. Chem. Soc., Perkin Trans. 2* **1996**, *2*, 783−787.
(41) Saint-Martin, H.; Ruiz-Vicent, L. E.; Ramırez-Solıs, A.; Ortega-Blake, I. *J. Am. Chem. Soc.* **1996**, *118*, 12167−12173.
(42) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.
(43) Becke, A. D. *Phys. Rev. A.* **1988**, *38*, 3098−3100.
(44) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(45) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, *37*, 785−789.
(46) Frisch, Æ.; Frisch, M. J. *Gaussian 98 User's Reference*, 2nd ed.; Gaussian, Inc.: Pittsburgh, PA, 1999.
(47) Peng, C.; Ayala, P. Y.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **1996**, *17*, 49−56.
(48) Bauernschmitt, R.; Ahlrichs, R. *J. Chem. Phys.* **1996**, *104*, 9047−9052.
(49) Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1977**, *66*, 3045−3050.
(50) Curtiss, L. A.; Jones, C.; Trucks, G. W.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1990**, *93*, 2537−2545.
(51) Curtiss, L. A.; Carpenter, J. E.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1992**, *96*, 9030−9034.
(52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.
(53) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027−2094.
(54) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett.* **1996**, *255*, 327−335.
(55) Mineva, T.; Russo, N.; Sicilia, E. *J. Comput. Chem.* **1998**, *19*, 290−299.
(56) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995−2001.
(57) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chim. Acta* **1999**, *103*, 9−63.
(58) Xidos, J. D.; Li, J.; Thompson, J. D.; Hawkins, G. D.; Winget, P. D.; Zhu, T.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G.; Frisch, M. J. *MN-GSM: A Module Incorporating the SM5.42 Solvation Models, the CM2 Charge Model, and Löwdin Population Analysis in the Gaussian98 Program*, 1.8 ed.; University of Minnesota: Minneapolis, MN, 2002.
(59) Xidos, J. D.; Li, J.; Thompson, J. D.; Hawkins, G. D.; Winget, P. D.; Zhu, T.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G.; Frisch, M. J. *MN-GSM*, version 1.8; University of Minnesota: Minneapolis, MN, 2001.

density in solution $\rho_{sol}(\mathbf{r})$ (which can be derived from ψ_{sol}). The Kohn−Sham energy functional $E[\psi]$ is given by

$$E[\psi] = T_S[\psi] + J[\rho] + E_{XC}[\psi] + \int \rho(\mathbf{r}) v_0(\mathbf{r})\, d^3r + E_{NN} = E_{elec}[\psi] + E_{NN} \quad (3)$$

where $T_S[\psi]$ is the noninteracting kinetic energy functional, $J[\rho]$ is the classical electrostatic energy, $E_{XC}[\psi]$ is the exchange-correlation energy, and E_{NN} is the nuclear−nuclear repulsion energy. The functional in eq 3 depends parametrically on the nuclear coordinates and charges under the classical Born−Oppenheimer approximation through the external nuclear potential $v_0(\mathbf{r})$, which under the fixed geometry approximation used here is the same in the gas phase and in solution. It remains to briefly describe the nature of the various models for $E_{sol}[\rho_{sol}]$ employed in the present work.

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

$$E_{sol}[\rho] = \frac{1}{2}\left[\int \rho(\mathbf{r}) \cdot v_{RF}(\mathbf{r})\, d^3r - \sum_\alpha Z_\alpha \cdot v_{RF}(\mathbf{R}_\alpha)\right] + G_{disp-repul} + G_{cav} \quad (4)$$

where $v_{RF}(\mathbf{r})$ is the solvent reaction-field potential, Z_α is the nuclear charge of atom α located at position $\mathbf{R}_\alpha$. The factor of $^1/_2$ in eq 4 results from the linear-response nature of the dielectric models, and the $G_{dis-repul}$ and G_{cav} represent the dispersion−repulsion and cavitation contributions, respectively.[54]

The cavitation term is computed using an expression obtained from scaled particle theory[60] with a cavity constructed from the UAHF radii.[61] The dispersion−repulsion term is calculated according to the prescription described by Floris et al.[62] with a solvent-accessible surface that is constructed from the UAHF radii plus a solvent probe radius of 1.385 Å.

The difference between the PCM and COSMO methods used here resides in the way in which the solvent reaction field potential $v_{RF}(\mathbf{r})$ is generated (see refs 54 and 56 for details). In the case of the PCM model, a cavity of unit dielectric is surrounded by a linear isotropic polarizable continuum of dielectric constant ϵ, the reaction field potential for which is solved numerically using a boundary element method.[53,61,63,64] In the conductor-like screening model,[65] a similar dielectric problem that involves a surrounding conductor ($\epsilon = \infty$) is solved, in the present case using a variation of the PCM method,[56] and the resulting reaction-field potential is corrected approximately for finite external dielectric ϵ.[65,66] Both of the methods used here do not account explicitly for the effect of "volume polarization" (referred to also as "outlying charge" or "charge penetration" effects).[67] However, both PCM and COSMO as implemented in GAUSSIAN98 normalize the polarization charge in accord with Gauss' law,[46] which is an implicit attempt to capture the main features of the volume polarization term. The parameters in both models have also been adjusted to reproduce experimental solvation free energies (and hence, to some extent, take into account higher-order effects implicitly). All PCM and COSMO calculations were carried out using the B3LYP/6-311++G(3df,2p) level of theory (the same level as the gas-phase single points) using the gas-phase optimized geometries.

In the SM5.42R model, E_{sol} is computed as the sum of two terms:

$$E_{sol} = G_{es} + G_{cds} \quad (5)$$

where G_{es} and G_{cds} represent the electrostatic and "cavity-dispersion−solvent-structure" contributions, respectively. The electrostatic term G_{es} in the SM5.42R model is calculated using a generalized Born expression[68] with the solute molecular electrostatic potential derived from the CM2 charge model.[57,69] The "cavity-dispersion−solvent-structure" term is given by:

$$G_{cds} = \sum_k A_k \sigma_k \quad (6)$$

where A_k and σ_k are the surface area and effective surface tension, respectively, for atom k and are described in detail elsewhere.[58] The SM5.42R model has been demonstrated to give accurate estimates for vapor pressures[70] and has recently been applied to biological molecules including nucleic acid bases.[71] All SM5.42R calculations were carried out using the B3LYP functional and the MIDI! basis set[72] at the optimized gas-phase geometries. The SM5.42R/B3LYP/MIDI! was parametrized using HF/MIDI! derived geometries; however, the SM5.42R model does not appear to be overly sensitive to small changes in geometry.[57]

2.3. Thermodynamic Quantities. For the energetic analysis, results are categorized into gas-phase and solvation contributions. In these sections, several thermodynamic contributions are distinguished (see section 2.1.). The breakdown of the key thermodynamics relations and energy components in the gas phase are summarized below:

$$G = H - T \cdot S \quad (7)$$

$$H = U + R \cdot T \quad (8)$$

$$U = E_0 + E_{vib} + E_{rot} + E_{trans} \quad (9)$$

$$E_0 = (E_{elec} + E_{NN}) + E_{ZPV} = E + E_{ZPV} \quad (10)$$

where G, U, H, S, and T are the Gibbs free energy, internal energy, enthalpy, entropy, and temperature, respectively, R is 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. The expression for the enthalpy (eq 8) assumes the ideal gas law for a mole 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.[73] The standard state is for a mole of particles at $T = 298$ K and 1 atm pressure ($V = R \cdot T/P$). All quantities above except E_0, E_{elec}, E_{NN}, and E_{ZPV} have explicit temperature dependence.

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

$$G_{aq} = G_{gas} + \Delta G_{sol} \quad (11)$$

where ΔG_{sol} is the solvation free energy (eq 2) described in 2.2.[53,68] Henceforth, unless otherwise stated, all thermodynamic quantities are assumed to be evaluated at 298 K and, if unsubscripted, correspond to gas-phase values (i.e., the explicit subscript "gas" will be dropped hereafter).

(60) Pierotti, R. A. *Chem. Rev.* **1976**, *76*, 717−726.
(61) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210−3221.
(62) Floris, F. M.; Tomasi, J.; Ahuir, J. L. P. *J. Comput. Chem.* **1991**, *12*, 784−791.
(63) Miertuš, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117−129.
(64) Cammi, R.; Tomasi, J. *J. Comput. Chem.* **1995**, *16*, 1449−1458.
(65) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, *2*, 799−805.
(66) York, D. M.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 11060−11079.
(67) Chipman, D. M. *J. Chem. Phys.* **1997**, *106*, 10194−10206.
(68) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2200.
(69) Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 1820−1831.
(70) Winget, P.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2000**, *104*, 4726−4734.
(71) Li, J.; Cramer, C. J.; Truhlar, D. G. *Biophys. Chem.* **1999**, *78*, 147−155.
(72) Easton, R. E.; Giesen, D. J.; Welch, A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chim. Acta* **1996**, *93*, 281−301.
(73) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*; John Wiley & Sons: Chichester, England, 2002.

2.4. Evaluation of Microscopic pK_a Values. The evaluation of pK_a values is of critical importance to the understanding of the mechanisms of transphosphorylation and hydrolysis in RNA catalysis. The absolute pK_a of a general acid AH is related to the standard Gibbs free energy of the ionization reaction

$$AH_{(aq)} \overset{K_a(AH)}{\rightleftharpoons} A^-_{(aq)} + H^+_{(aq)} \tag{12}$$

The pK_a is derived from the equilibrium constant as

$$\begin{aligned} pK_a(\mathrm{AH}) &= -\log K_a(\mathrm{AH}) \\ &= \frac{(G_{aq}(A^-) + Gaq(H^+)) - G_{aq}(AH)}{2.303RT} \end{aligned} \tag{13}$$

The evaluation of absolute pK_a values from quantum chemical calculations is a subject of intense interest—and also one that presents considerable challenges that have yet to be overcome. It is further complicated by some uncertainty regarding both the theoretical treatment and the experimental value for the solvation energy of the proton.[74,75]

Alternately, a *relative* pK_a (ΔpK_a) can be calculated with respect to the known reference molecule *BH* in an manner analogous to that for the absolute pK_a by considering the equilibrium

$$\mathrm{AH}_{(aq)} + \mathrm{B}^-_{(aq)} \overset{K_a(\mathrm{AH})/K_a(\mathrm{BH})}{\rightleftharpoons} \mathrm{A}^-_{(aq)} + \mathrm{BH}_{(aq)} \tag{14}$$

The relative pK_a of AH with respect to BH, ΔpK_a(AH/BH), can be calculated from

$$\begin{aligned} \Delta pK_a(\mathrm{AH/BH}) &= -\log K_a(\mathrm{AH}) + \log K_a(\mathrm{BH}) \\ &= \frac{(G_{aq}(\mathrm{A}^-) + G_{aq}(\mathrm{BH})) - (G_{aq}(\mathrm{B}^-) + G_{aq}(\mathrm{AH}))}{2.303RT} \end{aligned} \tag{15}$$

If the reference molecule BH has a known pK_a and is similar to the molecule AH for which a predicted pK_a is desired, substantial cancellation of errors may be obtained from calculation of the free energies in solution for both AH and BH to give

$$pK_a^{pred}(\mathrm{AH}) = pK_a^{known}(\mathrm{BH}) + \triangle pK_a^{calc}(\mathrm{AH/BH}) \tag{16}$$

This type of approach has been used extensively for the calculation of pK_a shifts of amino acid residues in proteins on the basis of molecular mechanical calculations[76−80] and also for the estimation of pK_a values for carboxylic acids[81] and other functional groups using a cluster-continuum approach,[82,83] the 2′ hydroxyl group in the hammerhead ribozyme[84] and small biologically relevant phosphoranes.[20] Calculation of a *macroscopic* pK_a requires the enumeration of an ensemble of possible protonation states and configurations.[76] A *microscopic* pK_a, on the other hand, considers only one particular protonation state and configuration, and hence neglects entropic effects due to multiple degenerate states and other protonation states that might have non-negligible occupation at 298 K.

In the present work, unless explicitly stated otherwise (such as in Table 3), predicted pK_a values are based on the *microscopic* pK_a calculated relative to dimethyl phosphate (DMPH) through eq 16. DMPH has an experimental *macroscopic* pK_a of 1.29;[85] since there are two possible protonation sites, the microscopic pK_a value is lower than 1.29. The microscopic pK_a values can be estimated from macroscopic ones according to the following formula[20]

$$\begin{aligned} pK_a^{Macro} &= -log\frac{m[A^-]_{micro}[H^+]}{n[AH]_{micro}} \\ &= pK_a^{micro} + \log\frac{n}{m} \end{aligned} \tag{17}$$

where the protonated species AH is assumed to have n indistinguishable microscopic states, and its unprotonated counterpart (A^-) to have m indistinguishable states. For DMPH, there are two possible protonation states ($n = 2$), depending on the phosphoryl oxygen being protonated, and only one for (DMP^-), so that the *microscopic* pK_a of DMPH is 0.99 (i.e., 1.29 − log(2/1)).

2.5. Bond Energy Model. Useful generalizations with regard to the stabilities of bonding interactions can be derived from a simple *bond energy model*. Consider the enthalpy of formation of a single (neutral) molecular structure "α" resulting from the binding of its constituent atoms in the gas phase

$$\sum_{l}^{atom\ types} n(\alpha)_l \cdot H^{atom}_{l,(g)} \xrightarrow{\Delta H_{bind}(\alpha)} H^{mol}(\alpha)_{(g)} \tag{18}$$

where $n(\alpha)_l$ is the number of atoms of type "l" in the molecular structure "α", $H^{atom}_{l,(g)}$ is the gas-phase enthalpy of the atom type l, $H^{mol}(\alpha)_{(g)}$ is the gas-phase enthalpy of molecule α, and $\Delta H_{bind}(\alpha)$ is the corresponding enthalpy of formation (binding) from constituent atoms.

The enthalpy of formation $\Delta H_{bind}(\alpha)$ is modeled as the sum of bond energies as

$$\Delta H_{bind}(\alpha) = -\sum_{i}^{bond\ types} B(\alpha)_i \cdot x_i \tag{19}$$

where $B(\alpha)_i$ is the number of bonds of type i in the molecule α and x_i is the *bond energy* of the bond type i. The bond energy model can provide useful information about the relative strengths of bonds given the use of a reliable fitting procedure and a sufficiently suitable dataset. A detailed account of the fitting procedure employed in the present work is provided elsewhere.[86]

3. Results

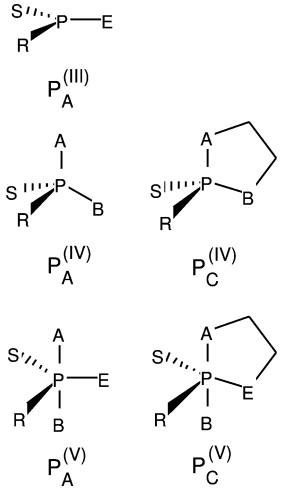
This section presents the results of quantum chemical calculations of a systematic series of (trivalent) metaphosphates and cyclic and acyclic (tetravalent) phosphate and (pentavalent) phosphorane molecules of biological significance. The first subsection introduces nomenclature conventions used in the tables and later discussions. The second, third, and fourth subsections make a brief presentation of the results for metaphosphates, phosphates, and phosphoranes, respectively.

3.1. Nomenclature and Abbreviations. The general forms of the metaphosphate, phosphate, and phosphorane compounds presented in this work are illustrated in Scheme 1. The following nomenclature is introduced to facilitate presentation and discussion:

• A neutral phosphorus compound is designated $P^{(Rn)}_{A/C}$−#, where *P* stands for phosphorus, the superscript *Rn* is a Roman

(74) Llano, J.; Eriksson, L. A. *J. Chem. Phys.* **2002**, *117*, 10193−10206.
(75) Chipman, D. M. *J. Phys. Chem. A* **2002**, *106*, 7413−7422.
(76) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219−10225.
(77) Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Chem.* **1990**, *19*, 301−332.
(78) Yang, A.-S.; Gunner, M. R.; Sampogna, R.; Sharp, K.; Honig, B. *Proteins* **1993**, *15*, 252−265.
(79) Gilson, M. K. *Curr. Opin. Struct. Biol.* **1995**, *5*, 216−223.
(80) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *Biochemistry* **1996**, *35*, 7819−7833.
(81) Toth, A. M.; Liptak, M. D.; Phillips, D. L.; Shields, G. C. *J. Chem. Phys.* **2001**, *114*, 4595−4606.
(82) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2001**, *105*, 7241−7247.
(83) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2002**, *106*, 7434−7439.
(84) Lyne, P. D.; Karplus, M. *J. Am. Chem. Soc.* **2000**, *122*, 166−167.
(85) Kumler, W.; Eiler, J. J. *J. Am. Chem. Soc.* **1943**, *65*, 2355−2361.
(86) York, D. M. Manuscript in preparation, 2003.

Scheme 1. Ligand Designations for Metaphosphate, Acyclic and Cyclic Phosphate, and Phosphorane Compounds



numeral (III, IV, or V) that indicates the number of covalent P−O bonds in the molecule, the subscript *A* or *C* indicates the "acyclic" or "cyclic" species shown in Scheme 1, and "#" is the number of the molecule listed in Table 1.

• A phosphorus monoanionic compound is designated $P_{A/C}^{(Rn)}$−#(X^-), where *Rn*, *A/C*, and # refer to the neutral compound from which the anion is derived, and the additional descriptor "*X*" identifies the specific deprotonation site (R, S, E, A, or B shown in Scheme 1).

The entire series of neutral molecules presented in this work are listed in Table 1. A comprehensive list of key thermodynamic quantities for neutral molecules in the gas phase and in solution is available in the Supporting Information.

3.2. Metaphosphates. Metaphosphates are important intermediates in so-called dissociative mechanisms for phosphate hydrolysis.[1,2,13] The inherent structure and stability of these species is therefore of fundamental interest. The neutral metaphosphate compounds discussed here are shown in Table 1, and indicated by the superscripted Roman numeral III.

The neutral metaphosphates $P_A^{(III)}$−**1** and $P_A^{(III)}$−**2** are both planar molecules that differ by the substitution of a hydroxyl group with a methoxy group. Both $P_A^{(III)}$−**1** and $P_A^{(III)}$−**2** have two P═O double bonds that range from 1.470 to 1.478 Å, respectively, and one P−O single bond of 1.603 and 1.591 Å, respectively. The solvation energy is smaller in magnitude by around 5 kcal/mol (4.26, 5.40 and 6.28 kcal/mol for SM5, COSMO and PCM models, respectively) when the hydroxyl group in $P_A^{(III)}$−**1** is replaced by a methoxy group in $P_A^{(III)}$−**2**. The pK_a shift relative to dimethyl phosphate (ΔpK_a) for $P_A^{(III)}$−**1** (Table 2) is predicted to be between −9.73 (PCM) and −12.43 (COSMO); hence, the anionic form is expected to be dominant in the aqueous phase around neutral pH. The anion PO_3^- is also planar, with P═O bonds having formal bond orders of 5/3 that are slightly elongated (1.509 Å) relative to the P═O double bonds of the neutral species. The solvation energy of the anion is −62.10 (PCM), −62.03 (COSMO), and −66.82 (SM5) kcal/mol and is similar to the calculated values for $H_2PO_4^-$ (Table 3).

3.3. Phosphates. Phosphate moieties form links between sugars that make up the sugar−phosphate backbone of DNA and RNA molecules. The understanding of both cyclic and acyclic phosphate compounds has relevance to biological phosphate hydrolysis mechanisms since they form typical reactants and products in the reactions. Here, a systematic study is made of a series of acyclic and cyclic monophosphates, the neutral forms of which are listed in Table 1 and indicated by the superscripted Roman numeral IV. Figure 1 shows the gas-phase optimized structures of selected acyclic and cyclic protonated phosphate compounds.

The H_3PO_4 molecule ($P_A^{(IV)}$−**1**) has a C_3 symmetry axis along the P═O double bond (bond length 1.482 Å) with the hydrogens making an O═P−O−H dihedral angle of 34.7° and oriented in the direction of doubly bonded oxygen. The structures $P_A^{(IV)}$−**2**, $P_A^{(IV)}$−**3**, and $P_A^{(IV)}$−**4** result from successive substitution of methyl groups at the hydrogen positions without significant change in orientation. As with H_3PO_4, the fully methylated structure ($P_A^{(IV)}$−**4**) also has a C_3 symmetry axis along the P═O double bond (Figure 1). There is a subtle overall lengthening of the P═O and P−O bond distances with number of methyl groups in the acyclic phosphates; for example, the P═O bond distances are 1.4835, 1.4837, 1.4849, and 1.4856 Å, for 0, 1, 2, and 3 methyl groups, respectively. The average O═P−O angles show a subtle decreasing trend with increasing methyl substitution. Molecules $P_A^{(IV)}$−**5** and $P_A^{(IV)}$−**6** are fully substituted with larger linear alkyl groups and have very similar structures in the gas phase. The protonated acyclic phosphates have slightly increased P═O double bond distances, decreased P−O single bond distances, and decreased O═P−O(C) angles relative to the protonated cyclic phosphates with the same number of carbons. The (C)O−P−O(C) angles of the acyclic phosphates (101.5−102.3°) are around 5.5° larger than the slightly strained ring of the cyclic phosphates (96.5−96.7°). The most notable difference between acyclic and cyclic protonated phosphates is the significantly increased dipole moments of the latter: the dipole moments for the cyclic phosphates $P_C^{(IV)}$−**1** and $P_C^{(IV)}$−**2** (4.1 and 3.9 D, respectively) are around a factor of 3.5 greater than the values of the corresponding acyclic phosphates $P_A^{(IV)}$−**3** and $P_A^{(IV)}$−**4** (1.2 and 1.1 D, respectively).

An interesting feature of the neutral phosphates involves the effect of methylation on the solvation energy. Figure 2 plots the solvation energy as a function of the number of carbon atoms for the 6 protonated phosphates. The acyclic molecules that have 0−3 carbons represent a series whereby hydroxyl groups are sequentially substituted by methoxy groups and the plot of the solvation energy versus number of carbons is nearly linear, especially with the SM5 solvation model (slope of −3 kcal/mol per carbon). For molecules with 3, 6, and 9 carbons, there are no hydroxyl groups present, and these molecules represent the series $OP(OC_nH_{2n+1})_3$ for $n = 1, 2, 3$, respectively. In the SM5 model, the plot over these three points is also nearly linear (slope of 0.4 kcal/mol per carbon, data not shown), whereas the PCM and COSMO models have a shallow minimum value for $n = 2$.

Comparison of the solvation energies of the protonated acyclic and cyclic phosphates shows that, for molecules with the same number of carbons ($P_A^{(IV)}$−**3**/$P_C^{(IV)}$−**1** and $P_A^{(IV)}$−**4**/$P_C^{(IV)}$−**2**), the cyclic structures have solvation energies 3.5−6 kcal/mol larger in magnitude than the corresponding acyclic structures. This results mainly from the increased dipole moments of the protonated cyclic phosphates noted earlier.

Table 1. Definition of Neutral Metaphosphates, Phosphates, and Phosphoranes

molecule	symbol	R	S	E	A	B
		Metaphosphates (Trivalent Phosphorus)				
P(O)(O)(OH)	$P_A^{(III)}$−**1**	O	O	OH	−	−
P(O)(O)(OCH_3)	$P_A^{(III)}$−**2**	O	O	OCH_3	−	−
		Acyclic Phosphates (Tetravalent Phosphorus)				
P(O)(OH)(OH)(OH)	$P_A^{(IV)}$−**1**	O	OH	−	OH	OH
P(O)(OH)(OH)(OCH_3)	$P_A^{(IV)}$−**2**	O	OH	−	OH	OCH_3
P(O)(OH)(OCH_3)(OCH_3)	$P_A^{(IV)}$−**3**	O	OH	−	OCH_3	OCH_3
P(O)(OCH_3)(OCH_3)(OCH_3)	$P_A^{(IV)}$−**4**	O	OCH_3	−	OCH_3	OCH_3
P(O)(OC_2H_5)(OC_2H_5)(OC_2H_5)	$P_A^{(IV)}$−**5**	O	OC_2H_5	−	OC_2H_5	OC_2H_5
P(O)(OC_3H_7)(OC_3H_7)(OC_3H_7)	$P_A^{(IV)}$−**6**	O	OC_3H_7	−	OC_3H_7	OC_3H_7
		Cyclic Phosphates (tetravalent phosphorus)				
P(O)(OH)(−O−CH_2CH_2−O−)	$P_C^{(IV)}$−**1**	O	OH	−	−O−CH_2CH_2−O-	
P(O)(OCH_3)(−O−CH_2CH_2−O−)	$P_C^{(IV)}$−**2**	O	OCH_3	−	−O−CH_2CH_2−O-	
		Acyclic Phosphoranes (Pentavalent Phosphorus)				
P(OH)(OH)(OH)(OH)(OH)	$P_A^{(V)}$−**1**	OH	OH	OH	OH	OH
P(OH)(OH)(OH)(OH)(OCH_3)	$P_A^{(V)}$−**2**	OH	OH	OH	OH	OCH_3
P(OH)(OH)(OCH_3)(OH)(OH)	$P_A^{(V)}$−**3**	OH	OH	OCH_3	OH	OH
P(OH)(OH)(OH)(OCH_3)(OCH_3)	$P_A^{(V)}$−**4**	OH	OH	OH	OCH_3	OCH_3
P(OH)(OH)(OCH_3)(OH)(OCH_3)	$P_A^{(V)}$−**5**	OH	OH	OCH_3	OH	OCH_3
P(OH)(OCH_3)(OCH_3)(OH)(OH)	$P_A^{(V)}$−**6**	OH	OCH_3	OCH_3	OH	OH
P(OH)(OH)(OCH_3)(OCH_3)(OCH_3)	$P_A^{(V)}$−**7**	OH	OH	OCH_3	OCH_3	OCH_3
P(OH)(OCH_3)(OCH_3)(OH)(OCH_3)	$P_A^{(V)}$−**8**	OH	OCH_3	OCH_3	OH	OCH_3
P(OCH_3)(OCH_3)(OCH_3)(OH)(OH)	$P_A^{(V)}$−**9**	OCH_3	OCH_3	OCH_3	OH	OH
P(OH)(OCH_3)(OCH_3)(OCH_3)(OCH_3)	$P_A^{(V)}$−**10**	OH	OCH_3	OCH_3	OCH_3	OCH_3
P(OCH_3)(OCH_3)(OCH_3)(OH)(OCH_3)	$P_A^{(V)}$−**11**	OCH_3	OCH_3	OCH_3	OH	OCH_3
P(OCH_3)(OCH_3)(OCH_3)(OCH_3)(OCH_3)	$P_A^{(V)}$−**12**	OCH_3	OCH_3	OCH_3	OCH_3	OCH_3
		Cyclic Phosphoranes (Pentavalent Phosphorus)				
P(OH)(OH)(−O−CH_2CH_2−O−)(OH)	$P_C^{(V)}$−**1**	OH	OH	-O−CH_2CH_2−O-		OH
P(OH)(OH)(−O−CH_2CH_2−O−)(OCH_3)	$P_C^{(V)}$−**2**	OH	OH	-O−CH_2CH_2−O-		OCH_3
P(OH)(OCH_3)(−O−CH_2CH_2−O−)(OH)	$P_C^{(V)}$−**3**	OH	OCH_3	-O−CH_2CH_2−O-		OH
P(OH)(OCH_3)(−O−CH_2CH_2−O−)(OCH_3)	$P_C^{(V)}$−**4**	OH	OCH_3	-O−CH_2CH_2−O-		OCH_3
P(OCH_3)(OCH_3)(−O−CH_2CH_2−O−)(OH)	$P_C^{(V)}$−**5**	OCH_3	OCH_3	-O−CH_2CH_2−O-		OH
P(OCH_3)(OCH_3)(−O−CH_2CH_2−O−)(OCH_3)	$P_C^{(V)}$−**6**	OCH_3	OCH_3	-O−CH_2CH_2−O-		OCH_3

In nucleic acids, the sugar−phosphate backbone contains a negatively charged, flexible phosphodiester linkage. The most common small molecule model for this group is dimethyl phosphate. Thermodynamic quantities for several stationary points on the potential energy surface corresponding to rotation about the α/ζ torsions (rotation about the P−O(C) single bonds) are summarized in Table 4. The conformational potential energy surface for dimethyl phosphate and other important phosphates, phosphonates, and phosphorothioates have been studied previously with ab initio methods in the gas phase and in solution.[12,89] The results are presented here for comparison and completeness. The g−g and g−t isomers are predicted to be separated by 0.74 kcal/mol in the gas phase, and 0.3−1.3 kcal/mol in the solution. The transition state TS1 is predicted to be slightly lower by 0.25 kcal/mol in the gas than TS3; however, in solution all solvation models predict that TS3 is lower by 0.61−1.12 kcal/mol.

3.4. Phosphoranes. Phosphorane molecules are important intermediates/transition states in biological phosphate transesterification and hydrolysis reactions. Here, a systematic study is made of a series of acyclic and cyclic phosphoranes, the neutral forms of which are listed in Table 1 and indicated by the superscripted Roman numeral V. Figure 3 shows the gas-phase optimized structures of selected acyclic and cyclic phosphoranes.

The equatorial P−O bond lengths are similar in the cyclic and acyclic phosphoranes (1.63−1.66 Å), whereas the axial P−O bond lengths are longer and show larger variation (1.67−1.73 Å). The variation in axial P−O bond lengths is due largely to the presence of intramolecular hydrogen bonding that involves equatorial hydroxyl groups. In structures that lack this hydrogen bonding, the axial P−O bond lengthens and shows smaller variation. In the cyclic phosphoranes, the axial P−O(C) bond involved in the ring is slightly elongated by around 0.03−0.06 Å relative to axial P−O bonds not part of the ring structure. Moreover, the dipole moments for the cyclic phosphoranes are typically larger than the values for the acyclic phosphoranes. In all the neutral phosphoranes, the axial O−P−O angle is nearly linear (173.8−179.6°).

Figure 4 plots the solvation energy as a function of the number of carbon atoms for the neutral acyclic and cyclic phosphoranes. As with the protonated phosphates, for a given solvation model, there is a nearly linear relationship between the solvation energy and number of carbons. The difference in solvation energy between different phosphorane isomers is small, especially for the SM5 model where typically the difference is less than 0.3 kcal/mol (the main exception being the cyclic isomers $P_C^{(V)}$−**4** and $P_C^{(V)}$−**5** where the difference is 0.7 kcal/mol). For the same solvation model, the slopes of the

(87) Lide, D. R., Ed. *CRC Handbook of Chemistry and Physics*, 83rd ed.; CRC Press LLC: Boca Raton, FL, 2003.
(88) Davies, J.; Doltsinis, N.; Kirby, A.; Roussev, C.; Sprik, M. *J. Am. Chem. Soc.* **2002**, *124*, 6594−6599.
(89) Foloppe, N.; MacKerell, A. D., Jr. *J. Phys. Chem. B.* **1999**, *103*, 10955−10964.

Table 2. Calculated Relative Gas Phase Enthalpies and Free Energies of Deprotonation and Estimated Microscopic Solution pK_a Values for Metaphosphates, Phosphates, and Phosphoranes[a]

			pK_a		
molecule	ΔH	ΔG	PCM	COSMO	SM5
Metaphosphates (Trivalent Phosphorus)					
$P_A^{(III)}$−**1**	−19.09	−19.84	−9.70	−12.43	−10.69
Acyclic Phosphates (Tetravalent Phosphorus)					
$P_A^{(IV)}$−**1**	−1.32	−1.97	2.65	0.68	0.03
$P_A^{(IV)}$−**2**	0.19	−0.51	0.26	0.84	0.87
$P_A^{(IV)}$−**3**	0.00	0.00	0.99	0.99	0.99
Cyclic Phosphates (Tetravalent Phosphorus)					
$P_C^{(IV)}$−**1**	−0.70	−1.83	−0.19	−0.09	−0.22
Acyclic Phosphoranes (Pentavalent Phosphorus)					
eq $P_A^{(V)}$−**1**	9.14	8.10	13.32	8.76	8.46
ax $P_A^{(V)}$−**1**	20.40	18.93	16.19	14.07	15.69
eq $P_A^{(V)}$−**2**	8.37	7.58	8.82	7.57	7.89
ax $P_A^{(V)}$−**2**	18.58	17.69	17.44	14.69	15.29
eq $P_A^{(V)}$−**3**	14.97	14.23	12.08	10.30	11.06
ax $P_A^{(V)}$−**3**	21.87	20.57	14.74	15.25	16.29
eq $P_A^{(V)}$−**4**	8.60	7.28	9.12	7.73	7.85
eq $P_A^{(V)}$−**5**	14.30	12.97	13.01	10.06	10.44
ax $P_A^{(V)}$−**5**	21.27	20.03	16.98	15.81	16.61
eq $P_A^{(V)}$−**6**	14.98	14.03	12.59	10.89	11.78
eq $P_A^{(V)}$−**7**	12.74	11.91	11.85	9.72	10.11
eq $P_A^{(V)}$−**8**	12.82	12.05	11.58	10.35	10.25
ax $P_A^{(V)}$−**8**	19.11	16.85	15.85	14.18	14.45
ax $P_A^{(V)}$−**9**	38.93	37.10	23.70	21.78	25.13
eq $P_A^{(V)}$−**10**	12.03	10.79	12.17	9.87	9.33
ax $P_A^{(V)}$−**11**	19.35	16.64	17.27	14.07	14.30
Cyclic Phosphoranes (Pentavalent Phosphorus)					
eq $P_C^{(V)}$−**1**	12.56	10.95	10.16	9.36	10.24
ax $P_C^{(V)}$−**1**	20.87	20.20	16.25	15.14	16.08
eq $P_C^{(V)}$−**2**	10.81	9.77	12.46	8.75	8.84
eq $P_C^{(V)}$−**3**	14.52	13.68	9.24	9.42	10.50
ax $P_C^{(V)}$−**3**	21.17	19.48	15.71	14.10	15.37
eq $P_C^{(V)}$−**4**	15.04	13.71	8.68	9.66	10.35

[a] ΔH and ΔG are in kcal/mol and are calculated for reactions of the type shown in eq 14 where AH is the neutral molecule listed in the table and BH is DMPH. Thus, the ΔH values above can be interpreted as gas-phase proton affinities relative to DMPH. Microscopic solution pK_a values are predicted from eq 16 using the microscopic pK_a value for DMPH (0.99)[20] as pK_a^{known}. Missing are entries that were found to be unstable with respect to geometry optimization (i.e., that were not energy minima). For the definition of the neutral compound abbreviations, see Table 1.

Table 3. Comparison of Calculated and Experimental Solvation Energies and Predicted pK_a Values[a]

molecule	PCM	COSMO	SM5	EXP
		ΔG_{sol}		
HOH	−6.92	−6.01	−5.96	−6.3
OH^-	−112.13	−107.04	−108.29	−110
CH_3OH	−4.94	−5.03	−5.39	−5.1
CH_3O^-	−86.76	−85.48	−83.90	−95
$H_2PO_4^-$	−63.18	−62.77	−74.56	−68
$OP(OCH_3)_3$	−3.56	−3.12	−7.48	−8.7
$OP(OC_2H_5)_3$	−4.77	−3.63	−6.69	−7.8
$OP(OC_3H_7)_3$	−2.00	−1.04	−5.34	−6.1
		pK_a		
DMPH/HOH	8.78	7.53	3.81	1.29[b]
H_3PO_4/HOH	9.96	6.75	2.33	2.12[c]
	23.23	20.46	10.68	7.21[c]
	29.38	26.61	8.58	12.67[c]
$CH_3OPO_3H_2$/HOH	7.57	6.90	3.21	1.541[b]
	23.79	21.06	11.37	6.312[b]
eq $P(OH)_5$/HOH	20.33	14.53	10.51	8.6[d]
ax $P(OH)_5$/HOH	23.38	20.02	17.91	13.5[d]
H_3PO_4/DMPH	2.47	0.50	−0.25	2.12[c]
	15.74	14.21	8.16	7.21[c]
	21.83	20.37	6.05	12.67[c]
$CH_3OPO_3H_2$/DMPH	0.38	0.96	0.99	1.54[b]
	16.60	15.12	9.15	6.31[b]
eq $P(OH)_5$/DMPH	12.84	8.28	7.98	8.6[c]
ax $P(OH)_5$/DMPH	15.89	13.77	15.39	13.5[c]

[a] Solvation energies are in kcal/mol. Solvation energies and relative pK_a values were calculated with the polarizable continuum model (PCM), conductor-like screening model (COSMO) and SM5 solvation model (SM5). See text for more details. Experimental values for solvation energies were taken from ref 57. Relative pK_a values are taken with respect to water (HOH) and protonated dimethyl phosphate (DMPH), which is $P_A^{(IV)}$−**3** in Table 1. The predicted pK_a values were calculated from eqs 16 and 17 using the *macroscopic* pK_a values for water and DMPH as pK_a^{known}. [b] Reference 85. [c] Reference 87. [d] Experimental estimate of pK_a values for equitorial and axial protons in tetracyclohexyloxyhydroxyphosphorane from bond length − pK_a correlations based on crystal structures of cyclohexanol derivatives.[88]

solvation energies with number of carbons are nearly the same for acyclic and cyclic phosphoranes (i.e., the fitted lines are nearly parallel), indicating that methyl substitution has a similar effect on the relative solvation energy. As for protonated phosphates, the cyclic phosphoranes are more highly solvated that the corresponding acyclic phosphoranes by around 3−6 kcal/mol. Although all of the solvation models predict a near-linear relationship between solvation energy and number of carbons, a notable difference is observed in the slopes of the plots in Figure 4 between different solvation models. In particular, the slopes for the SM5 model (2.1 and 2.2 kcal/mol/carbon for acyclic and cyclic phosphoranes, respectively) differ significantly from those of the PCM and COSMO models (3.3 and 3.6−4.0 kcal/mol/carbon for acyclic and cyclic phosphoranes, respectively).

The importance of intramolecular hydrogen bonding on the structure and stability of neutral and monoanionic phosphoranes is an important issue with regard to discussions of transition states and intermediates in phosphate hydrolysis reactions, especially under acidic conditions. Intramolecular hydrogen bonding can cause considerable elongation and weakening of the axial P−O bonds, in particular, the P−O bond involved in endocyclic cleavage (see Table 5). In the neutral cyclic phosphoranes the endocyclic axial P−O bond elongates to 1.78−1.79 Å when doubly hydrogen bonded, whereas the exocyclic P−O bond is 1.64 Å. This situation is even more dramatic for the monoanionic system where the endocyclic bond lengthens to 1.89−2.01 Å when doubly hydrogen bonded, whereas the exocyclic P−O bond is 1.70−1.71 Å. Moreover, the axial bond angle becomes considerably distorted in the monoanionic phosphoranes (164.1−168.4° and 160.4−160.9° for $P_C^{(V)}$−**1** and $P_C^{(V)}$−**2**, respectively) relative to the neutral phosphoranes (176.3−179.6° and 175.2−177.5° for $P_C^{(V)}$−**1** and $P_C^{(V)}$−**2**, respectively). The distortion is slightly greater, by roughly 4°, for the monoanionic forms of $P_C^{(V)}$−**2** that has a methoxy group in the axial exocyclic position.

4. Discussion

In this section, the data for phosphate and phosphoranes are used to calculate a variety of reaction energies and other properties that are relevant to understanding their differential stability and reactivity. Emphasis is placed on four main aspects: reaction energies for isomerization of neutral phosphoranes that involve axial/equatorial exchange of hydroxyl/methoxy groups, reaction energies for ligand substitution reactions, pK_a values in solution, and bond energies.

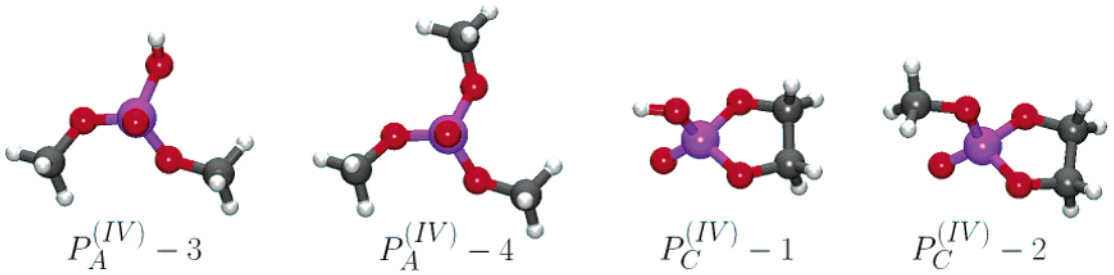


Figure 1. Gas-phase optimized structure of selected acyclic and cyclic protonated phosphates.

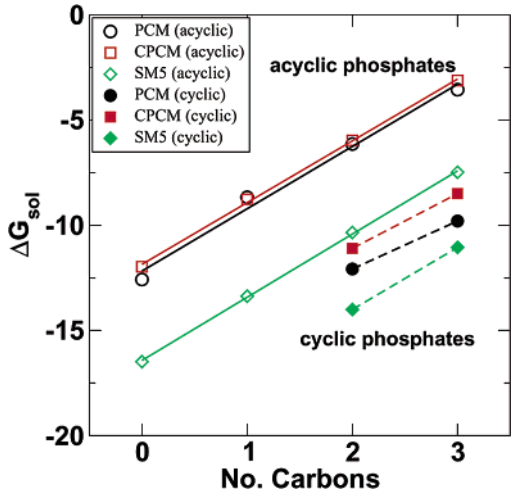


Figure 2. Solvation energies of protonated acyclic phosphates (molecules $P_A^{(IV)}$ **−1**- $P_A^{(IV)}$ **−6**) as a function of number of carbons.

Table 4. Relative Conformational Free Energies of Dimethyl Phosphate in the Gas Phase and in Solution[a]

				ΔG_{aq}		
conformation (α,ζ)	ΔH	$T\Delta S$	ΔG	PCM	COSMO	SM5
g−g (73.7,73.6)	-	-	-	-	-	-
TS1 (132.9,73.1)	1.234	−0.973	2.207	1.267	2.297	2.294
g−t (−165.3,71.4)	1.108	0.362	0.746	−0.474	0.516	0.294
TS2 (155.0,131.2)	2.156	−0.536	2.692	1.282	2.332	2.474
t−t (147.4,147.6)	2.765	1.093	1.672	2.302	1.222	1.358
TS3 (180.0,180.0)	2.254	−0.206	2.460	0.800	1.610	1.178

[a] Energetic quantitites are in kcal/mol. For comparison with other ab initio results see, for example, ref 12.

4.1. Isomerization Reactions of Phosphoranes. This section describes results of isomerization of neutral phosphoranes in the gas phase and in solution that involve the exchange of hydroxyl and methoxy groups from axial to equatorial positions. In both the gas phase and in solution, methoxy groups (OCH_3) prefer an axial position relative to hydroxyl (OH) groups. Rules for axial and equatorial ligand preferences in phosphoranes have been discussed extensively elsewhere.[90] A predominant feature is the preference for electronegative groups for axial positions. Here, the difference in polarity between P−OH and P−OCH_3 bonds is fairly small; however, there is a steric effect associated with the larger methoxy group.

The gas-phase thermodynamic quantities for phosphorane isomerization reactions are summarized in Table 6. There is considerable enthalpy/entropy compensation that occurs in the isomerization reactions. For single exchanges ($\Delta_{ax\rightarrow eq}^{OCH_3}$ values of 1), the enthalpy favors axial methoxy groups over hydroxyl groups by 0.7−3.8 and 1.1−4.3 kcal/mol for acyclic and cyclic phosphoranes, respectively, whereas the entropy contribution favors axial hydroxyl groups to a lesser degree from 0.5 to 1.0 and 0.4 to 0.6 kcal/mol for acyclic and cyclic phosphoranes, respectively. The free energy in the gas phase indicates a preference in all cases for axial methoxy groups ranging from 0.2 to 3.1 and 0.7 to 3.8 kcal/mol for single $\Delta_{ax\rightarrow eq}^{OCH_3}$ exchanges for the acyclic and cyclic phosphoranes, respectively.

In the aqueous phase, the qualitative picture remains the same (Table 6): methoxy groups prefer to occupy the axial positions relative to hydroxyl groups. This is in accord with the observation noted earlier that the solvation energy does not significantly change between isomers.

4.2. Ligand Substitution Reactions. This section describes results of ligand substitution reactions of the form

$$A(OH) + CH_3OH \rightarrow A(OCH_3) + HOH \qquad (20)$$

Formally, the above reaction is a ligand substitution combined with a relative proton affinity; however, for brevity this type of reaction is henceforth referred to simply as a ligand substitution reaction.

Table 7 shows the thermodynamic quantities for ligand substitution reactions for the neutral phosphates and phosphoranes. In the gas phase, the free energy of hydroxyl substitution by a methoxy group in the neutral phosphate (relative to water and methanol) is favorable (negative), ranging from −0.5 to −1.4 kcal/mol. However, for the phosphoranes, the corresponding free energy is positive, ranging from 1.1 to 5.4 kcal/mol. This arises from a balance between bonding interactions and steric effects. For the gas-phase phosphates, it is preferable for the methoxy groups to be bonded in the phosphate and for the hydroxyl group to form a water molecule. For the gas-phase phosphoranes, however, the situation is reversed: it is preferable for the hydroxyl groups to be bonded in the phosphorane and for the methoxy group to form a methanol molecule. The larger methoxy group is more sterically unfavorable in the pentavalent phosphoranes than in the tetravalent phosphates. This is evident by the trend that the gas-phase free energy for ligand substitution becomes increasingly more positive for the phosphoranes as the ligand environment becomes more sterically crowded (i.e., as more methoxy groups are present).

The preference for a methoxy group in the gas-phase ligand substitution reaction of the cyclic phosphate (−1.4 kcal/mol) is enhanced relative to the corresponding reactions of the acyclic phosphates (around −0.6 kcal/mol). For the acyclic and cyclic phosphoranes, the magnitudes of the free energy for gas-phase ligand substitution reactions span a similar range (1.1−4.7 and 2.0−5.4 kcal/mol, respectively) with the entropic term being an important and sometimes dominant contribution.

(90) Holmes, R. R. *J. Am. Chem. Soc.* **1978**, *100*, 433−446.

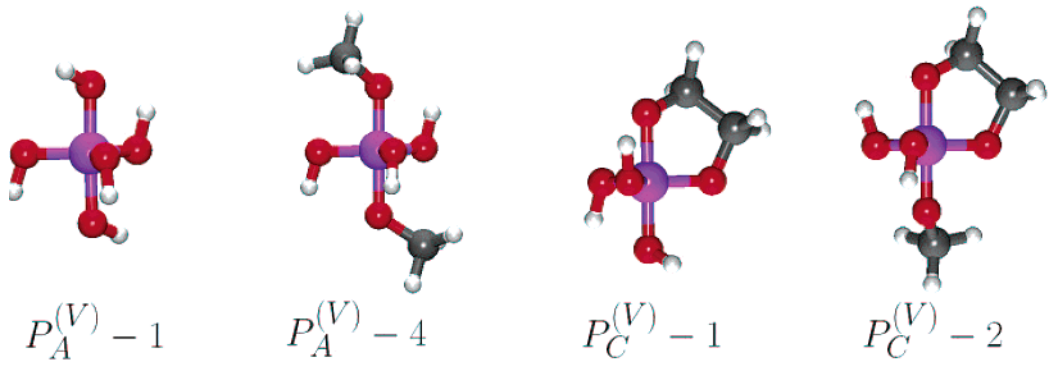


Figure 3. Gas-phase optimized structure of selected acyclic and cyclic phosphoranes.

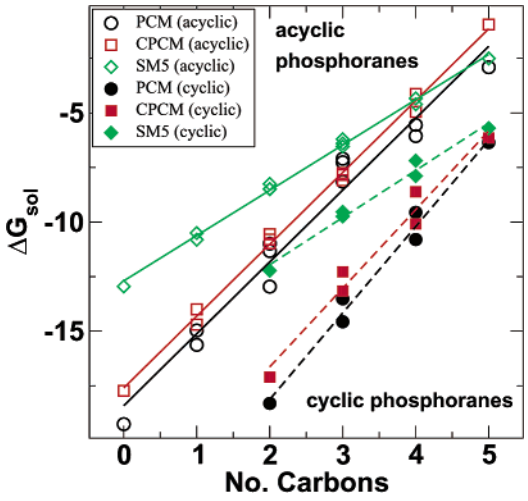


Figure 4. Solvation energies of protonated acyclic and cyclic phosphoranes (molecules $P_A^{(V)}$−**1**, $P_A^{(V)}$−**12**, and $P_C^{(V)}$ − $P_C^{(V)}$) as a function of number of carbons.

Table 5. Structures for Different Protonation States of Neutral and Monoanionic Phosphoranes[a]

O_{1P}	O_{2P}	P–O_{ax}^{ring}	P–O_{ax}	O_{ax}–P–O_{ax} ∠
		Neutral $P_C^{(V)}$−1		
↑	↑	1.7824	1.6447	176.3
↑	↓	1.7252	1.6788	176.6
↓	↑	1.7240	1.6767	179.6
↓	↓	1.6805	1.7263	176.7
		Monoanionic $P_C^{(V)}$-1 (R⁻/S⁻)		
↑	0	2.0052	1.7028	164.1
0	↑	1.9204	1.7040	164.4
↓	0	1.7788	1.7866	168.4
0	↓	1.7830	1.7835	164.6
		Neutral $P_C^{(V)}$−2		
↑	↑	1.7940	1.6381	177.0
↑	↓	1.7314	1.6756	177.5
↓	↑	1.7289	1.6732	175.2
↓	↓	1.6853	1.7208	175.7
		Monoanionic $P_C^{(V)}$-2 (R⁻/S⁻)		
↑	0	1.8933	1.7149	160.9
0	↑	1.8934	1.7149	160.9
↓	0	1.7798	1.7894	160.4
0	↓	1.7795	1.7897	160.4

[a] Bond lengths are in Å, angles in degrees. For the definition of the neutral compound abbreviations, see Table 1.

In the aqueous phase, the equilibrium for the phosphate ligand substitution equilibrium is reversed from the gas phase (i.e., the free energies are positive, with the exception of the PCM predicted value for the cyclic phosphate ligand substitution reaction). The effect of solvation on the free energy is to cause a positive shift from the gas-phase values, leaving the relative values largely unchanged. This is predominantly due to the preferable solvation of water versus methanol. As discussed earlier, for a given solvation model, the solvation free energy increases linearly (becomes less negative) with addition of successive methyl groups through OH → OCH_3 substitutions. This is illustrated in Figures 2 and 4.

Table 6. Calculated Thermodynamic Data for Isomerizations of Pentavalent Phosphorus Compounds[a]

molecule	$\Delta_{ax\rightarrow eq}^{OCH_3}$	ΔE	ΔG	ΔG_{aq} PCM	ΔG_{aq} COSMO	ΔG_{aq} SM5
Acyclic Phosphoranes (Pentavalent Phosphorus)						
$P_A^{(V)}$−**2** → $P_A^{(V)}$−**3**	1	0.813	0.188	0.848	0.888	0.495
$P_A^{(V)}$−**4** → $P_A^{(V)}$−**5**	1	0.946	0.305	2.275	0.725	0.536
$P_A^{(V)}$−**4** → $P_A^{(V)}$−**6**	2	3.996	2.942	4.572	3.122	2.931
$P_A^{(V)}$−**7** → $P_A^{(V)}$−**8**	1	3.408	2.559	3.609	2.809	2.225
$P_A^{(V)}$−**7** → $P_A^{(V)}$−**9**	2	4.742	3.372	4.262	3.712	3.183
$P_A^{(V)}$−**10** → $P_A^{(V)}$−**11**	1	4.030	3.061	3.571	2.241	2.793
Cyclic Phosphoranes (Pentavalent Phosphorus)						
$P_C^{(V)}$−**2** → $P_C^{(V)}$−**3**	1	1.318	0.731	1.781	1.611	0.936
$P_C^{(V)}$−**4** → $P_C^{(V)}$−**5**	1	4.568	3.751	2.511	2.281	3.049

[a] Energetic quantities are in kcal/mol. The designation $D_{ax\rightarrow eq}^{OCH_3}$ gives the number of methoxy groups that have exchanged axial → equatorial positions with hydroxyl groups in the isomerization reaction. For the definition of the neutral compound abbreviations, see Table 1.

4.3. Solvation Energies and Microscopic pK_a Values. The pK_a values for phosphates and phosphoranes have important implications into mechanisms of transesterification and phosphate hydrolysis catalyzed by ribozymes. Moreover, the quantification of gas-phase proton affinities for phosphates and phosphoranes with high-level quantum methods provides important benchmarks that can be used to design new semiempirical Hamiltonian models that can be applied in hybrid QM/MM calculations.

Comparisons of the calculated and experimental solvation energies and pK_a values are given in Table 3 for several molecules relevant to the present work. With the exception of the three neutral phosphate compounds of the form $OP(OR)_3$ where R is a hydrocarbon side chain, the solvation energies are similar in accuracy. For the neutral $OP(OR)_3$ phosphate compounds, all the solvation models under-predict the magnitude of the solvation energy. The SM5 model matches the experimental values most closely, having an error of around 1 kcal/mol, whereas the PCM and COSMO models have errors around 3−5 kcal/mol.

Table 7. Calculated Thermodynamic Data for Ligand Substitution Reactions of Phosphorus Compounds[a]

molecule	ΔE	ΔG	ΔG_{aq} PCM	ΔG_{aq} COSMO	ΔG_{aq} SM5
Acyclic Phosphates (Tetravalent Phosphorus)					
$P_A^{(IV)}$−**1** + $CH_3OH \rightarrow P_A^{(IV)}$−**2** + HOH	−1.551	−0.745	1.184	1.474	1.791
$P_A^{(IV)}$−**2** + $CH_3OH \rightarrow P_A^{(IV)}$−**3** + HOH	−1.266	−0.459	0.090	1.360	1.992
$P_A^{(IV)}$−**3** + $CH_3OH \rightarrow P_A^{(IV)}$−**4** + HOH	−1.001	−0.470	0.129	1.399	1.819
Cyclic Phosphates (Tetravalent Phosphorus)					
$P_C^{(IV)}$−**1** + $CH_3OH \rightarrow P_C^{(IV)}$−**2** + HOH	−0.496	−1.354	−1.054	0.265	1.023
Acyclic Phosphoranes (Pentavalent Phosphorus)					
$P_A^{(V)}$−**1** + $CH_3OH \rightarrow P_A^{(V)}$−**2** + HOH	2.132	2.778	4.428	4.828	4.346
$P_A^{(V)}$−**2** + $CH_3OH \rightarrow P_A^{(V)}$−**4** + HOH	−0.232	1.122	1.812	3.862	2.873
$P_A^{(V)}$−**4** + $CH_3OH \rightarrow P_A^{(V)}$−**7** + HOH	3.069	3.338	6.168	5.248	5.041
$P_A^{(V)}$−**7** + $CH_3OH \rightarrow P_A^{(V)}$−**10** + HOH	3.598	4.096	4.206	7.076	5.392
$P_A^{(V)}$−**10** + $CH_3OH \rightarrow P_A^{(V)}$−**12** + HOH	4.363	4.654	5.824	6.854	5.917
Cyclic Phosphoranes (Pentavalent Phosphorus)					
$P_C^{(V)}$−**1** + $CH_3OH \rightarrow P_C^{(V)}$−**2** + HOH	1.298	2.284	4.264	4.494	3.628
$P_C^{(V)}$−**2** + $CH_3OH \rightarrow P_C^{(V)}$−**4** + HOH	1.275	2.041	5.071	5.611	4.019
$P_C^{(V)}$−**4** + $CH_3OH \rightarrow P_C^{(V)}$−**6** + HOH	5.060	5.382	6.602	6.862	6.299

[a] Energetic quantities are in kcal/mol. For the definition of the neutral compound abbreviations, see Table 1.

The overall reliability of the models is even more pronounced in the comparison of pK_a values calculated from relative pK_a values with respect to water (HOH) and dimethyl phosphate (DMPH). In this case the PCM and COSMO models employed here lead to artificially elevated pK_a values, especially for the multiply charged anions. A problem, however, with the SM5 model occurred with the highest (third) pK_a for H_3PO_4 that was predicted to be lower than the second pK_a.

In the case of the phosphorane species, theoretical methods have predicted a values of 7.9 for the macroscopic pK_{a1} of ethylene phosphorane from localized basis set density functional methods,[20] and values of 9.8 and 14.2 have been predicted for the equatorial and axial positions of pentahydroxyphosphorane from Car−Parinello density functional calculations.[88] Proposed experimental values for pK_{a1} for oxyphosphoranes range from 6.5 to 11.0.[1] A recent estimate of pK_a values for equatorial and axial protons in tetracyclohexyloxyhydroxyphosphorane from bond length−pK_a correlations based on crystal structures of cyclohexanol derivatives gives values of 8.62 and 13.5, respectively.[88] The COSMO model gives phosphorane pK_a values for equatorial/axial protons (when measured relative to dimethyl phosphate) of 8.28/13.77 that falls within the proposed range; the SM5 model gives similar values of 7.98/15.39, having a slightly elevated pK_a for the axial proton. The PCM model gives pK_a values that are considerably higher and lie outside the proposed range for phosphoranes. It is clear that there remains important work to be done in the development of solvation methods that can accurately reproduce and reliably predict multiple pK_a values for biological phosphates and phosphoranes.

Table 2 lists the predicted microscopic pK_a values for metaphosphates, acyclic and cyclic phosphates, and phosphoranes. The estimates are based on the calculation of microscopic pK_a shifts for each molecule with respect to dimethyl phosphate according to eq 16. The extensive dataset covered in this work allow for the identification of interesting trends in pK_a's for this set of compounds. These trends are summarized below.

It is evident from Table 2 the trend in pK_a values for metaphosphates, phosphates, and phosphoranes:

metaphosphate pK_a ≪ **phosphate** pK_a ≪ **phosphorane** pK_a

For the phosphates, the cyclic structures have lower pK_a values than the corresponding acyclic structures by around 1 pK_a unit:

cyclic phosphate pK_a < **acyclic phosphate** pK_a

This is related to the preferential stabilization of the more highly solvated cyclic phosphates that allow them to remain as solvated anions at lower pH. Recall that the solvation energies of the neutral cyclic phosphates were more favorable (greater in magnitude) than for the corresponding acyclic molecules. This derived from the enhanced dipole moment of the former. Upon ionization, the preferential solvation for the cyclic phosphates becomes even more pronounced due to their more compact structure (i.e., smaller effective Born radius) and results in lower pK_a values relative to the corresponding acyclic phosphates.

For the phosphoranes, there is less of a difference between acyclic and cyclic structures; however, there is considerable difference between deprotonation of the axial and equatorial protons, deprotonation of the equatorial position being favored by several pK_a units:

equatorial phosphorane pK_a < **axial phosphorane** pK_a

The average intramolecular pK_a shift for axial/equatorial protons (i.e., $\langle pK_a^{axial} - pK_a^{equatorial} \rangle$) was 4.99, 5.35, and 5.85 for the PCM, COSMO, and SM5 models, respectively. These values are comparable to reported values predicted from Car−Parinello density functional calculations (4.4) and experimental estimates (4.9) of phosphoranes.[88] In two cases ($P_C^{(V)}$−**5** and $P_A^{(V)}$−**6**), no values are given in Table 2 for the pK_a of the axial proton. This is because a local minimum could not be trapped: the geometry optimization procedure would consistently cause the species to pseudorotate so that the deprotonated oxygen was in an equatorial position. This is further evidence of the instability of deprotonation at the axial positions.

4.4. Bond Energy Model. To draw general conclusions about the average bond strengths of biological phosphorus compounds, a bond energy model was employed.[86] The bond energy model

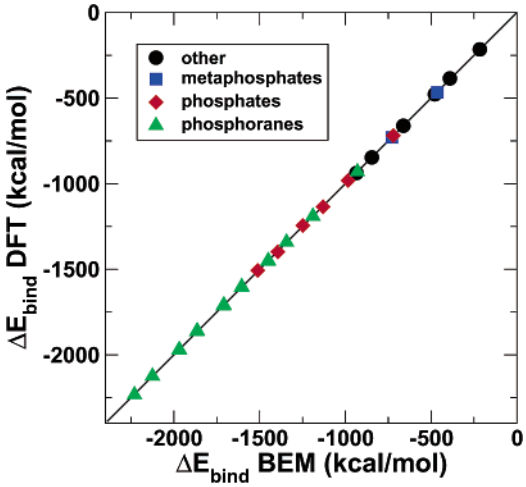


Figure 5. Regression of the molecular binding energies (ΔE_{bind}) for neutral molecules calculated with density functional theory (DFT) and with the bond energy model (BEM).

Table 8. Calculated Bond Energies and Bond Lengths for Biological Phosphorus compounds[a]

bond type	energy (kcal/mol)	⟨*bl*⟩ (Å)	(σ_{bl}) (Å)
C−C	81.6	1.530	3.0×10^{-3}
C−O(C)	83.0	1.433	9.5×10^{-3}
C−O(H)	83.0	1.426	9.4×10^{-4}
H−C	98.6	1.094	2.4×10^{-3}
H−O(H,C)	110.1	0.965	4.8×10^{-4}
H−O(H,H)	110.1	0.965	4.9×10^{-7}
H−O(H,P)	110.1	0.968	2.1×10^{-3}
O(C)−P$^{(III)}$	79.3	1.591	–
O(C)−P$^{(IV)}$	85.4	1.613	1.2×10^{-2}
O(C)−P$^{(V)}$ eq	79.3	1.641	1.4×10^{-2}
O(C)−P$^{(V)}$ ax	70.5	1.703	2.5×10^{-2}
O(H)−P$^{(III)}$	83.9	1.603	–
O(H)−P$^{(IV)}$	86.6	1.613	3.0×10^{-3}
O(H)−P$^{(V)}$ eq	83.9	1.636	6.4×10^{-3}
O(H)−P$^{(V)}$ ax	72.8	1.689	1.9×10^{-2}
O═P$^{(III)}$	140.2	1.474	3.6×10^{-3}
O═P$^{(IV)}$	140.2	1.482	3.7×10^{-3}

[a] Bond energies ("Energy") for each "Bond type" were calculated by fitting the bond energy model to the molecular enthalpy of binding at 298 K (see text) and given here in kcal/mol. Bond lengths ("bl") and their standard deviations ("σ_{bl}") for each bond type are given in a. All C−O, H−O, and P═O bonds energies were constrained to be equal, as were the O(H)−P and O(C)−P bonds of trivalent phosphorus and (equatorial) pentavalent phosphorus.

was fit to the calculated enthalpy of formation of each molecule resulting from the binding of the isolated atoms at 298 K (referred to as the molecular "binding energy", ΔE_{bind}) and to the relative energies associated with the isomerization and ligand substitution reactions. The details of the procedure are described elsewhere.[86] A regression of ΔE_{bind} values from density functional calculations and estimated from the bond energy model are shown in Figure 5.

The bond types, bond energies, average bond lengths, and root-mean-square deviations for covalent bonds are listed in Table 8. The two types of intramolecular hydrogen bonds found in this data set (H···O(C)−P and H···O(H)−P) have average bond lengths of 2.064 ± 0.061 and 2.091 ± 0.036 Å and bond angles of 93.6 ± 1.5° and 92.2 ± 2.0°, respectively. (See Table S10 in the Supporting Information.) The bond energy model presented here takes into account in an average way conformational and stereoelectronic effects of each type of bond.

The determination of bond strengths or related bond energies from experiment is often difficult, and various methods have been employed for their calculation; however, in the area of phosphorus chemistry, there have been relative few systematic studies.[91] In this work, the calculation of a complete series of metaphosphates, phosphates, and phosphoranes at a consistent level of quantum theory allows the estimation of bond energies specifically for relevant biological phosphorus compounds.

Comparison of the bond lengths and bond energies reveals several trends:

- Metaphosphates have slightly shorter P−O single and P═O double bonds than those of the phosphates and phosphoranes.
- P−O single bonds in phosphates are slightly shorter and have larger bond energies than corresponding bonds in phosphoranes.
- P−O(H) and P−O(C) bonds in phosphates are similar in bond length and bond energy.
- Axial P−O bonds in phosphoranes are longer and weaker than equatorial P−O bonds.
- The difference between axial/equatorial P−O(C) bonds (8.8 kcal/mol) is smaller than the difference between axial/equatorial P−O(H) bonds (11.1 kcal/mol), in support of the observation that methoxy groups prefer axial positions.
- Intramolecular hydrogen bond lengths are fairly long (2.06−2.09 Å), form nearly right angles (92.2−93.6°), and provide a moderate stabilization of the axial P−O bonds.

The bond energy values of the P−O single bonds range from 70.5 kcal/mol for axial P−O(C) phosphorane bonds to 86.6 kcal/mol for P−O(H) phosphate bonds. The phosphate bond energy values (85.4 and 86.6 kcal/mol for P−O(C) and P−O(H) bonds, respectively) are close to the average P−O bond energy (86 kcal/mol) reported elsewhere.[91] The bond energy for the P═O double bond in the metaphosphate and phosphate molecules (140.2 kcal/mol) is similar to the tentative P═O bond energy (130 kcal/mol) reported elsewhere[91] and close to the value of the experimental bond dissociation energy of the PO molecule (140 kcal/mol).[91]

5. Conclusions

The present work reports results of theoretical calculations of a series metaphosphates and acyclic and cyclic phosphates and phosphoranes relevant to the study of RNA catalysis and phosphate hydrolysis reactions in solution. Solvation effects were treated and compared using three different solvation models: PCM,[53−55] COSMO,[56,65] and SM5.[57] The structure and stability of these compounds have been characterized through thermodynamic quantities and solvation energies. Cyclic phosphates and phosphoranes were found to be preferentially stabilized by solvent relative to the analogous acyclic molecules. Solvation energy was observed to have a nearly linear relationship with number of methoxy groups (carbon atoms) for both the acyclic and cyclic phosphates and phosphoranes. Stationary points on the potential energy surface of dimethyl phosphate were identified, and intramolecular hydrogen bonding of neutral and monoanionic cyclic phosphoranes were analyzed. Intramolecular hydrogen bonds cause considerable elongation of the axial P−O bonds, especially in the monoanionic cyclic phos-

(91) Corbridge, D. E. C. *Phosphorus 2000: Chemistry, Biochemistry and Technology*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 2000.

phoranes, which also showed considerable distortion of the $O_{ax}-P-O_{ax}$ bond angle.

The calculated thermodynamic quantities for isomerization and ligand substitution reactions indicate that methoxy groups prefer axial positions, and in solution, prefer to be bound to phosphorus with a water molecule as opposed to forming a hydroxyl-bound ligand and methanol. Analysis of gas-phase proton affinities and solution pK_a values indicate several trends, including (1) the pK_a values of metaphosphates are considerably lower than for phosphates, which are considerably lower than phosphoranes, (2) cyclic phosphates and phosphoranes have lower pK_a values than corresponding acyclic molecules, and (3) protonation of the equatorial position of phosphoranes is around 4 pK_a units lower than the axial positions. Finally, the extensive series of data allowed the estimation of bond energies. It was found that (1) P−O single bonds in phosphates are stronger than in phosphoranes, (2) axial P−O bonds are considerably weaker than equatorial P−O bonds by around 10 kcal/mol, and (3) P−O(C) bonds are more apicophilic in that the difference in bond energy between axial/equatorial P−O(C) bonds is smaller than for P−O(H) bonds.

The results reported here provide quantitative insight into the structure and stability of biological phosphorus compounds relevant for RNA catalysis. In addition, this work serves as the first step of construction of a high-level quantum database from which improved quantum models, such as new semiempirical Hamiltonians, can be derived that are more useful for applications to complex reactions catalyzed by ribozymes. These applications may involve, for example, large-scale linear-scaling electronic structure calculations or activated dynamics simulations along multidimensional reaction coordinates using hybrid QM/MM potentials that would not be feasible in the near future at the level of quantum theory employed here. Consequently, the development of new-generation models that provide very accurate description of bonding, polarization, and quantum many-body effects, yet that are highly efficient and can be combined with enhanced configurational sampling techniques, may present a significant step toward the reliable theoretical treatment and detailed understanding of ribozyme catalysis.

Acknowledgment. D.Y. is grateful for financial support provided by the National Institutes of Health (Grant 1R01-GM62248-01A1) and the donors of the Petroleum Research Fund, administered by the American Chemical Society. X.L. thanks the MSI for a Research Scholar Award, the Basque Government (Eusko Jaurlaritza), and the University of the Basque Country (Euskal Herriko Unibertsitatea) for financial support. M.M. was partially supported by a fellowship from the Chemical Physics program at the University of Minnesota. Computational resources were provided by the Minnesota Supercomputing Institute.

Supporting Information Available: Table S1: Calculated gas-phase thermodynamic data for metaphosphates, phosphates and phosphoranes. Table S2: Calculated solvation energies for metaphosphates, phosphates and phosphoranes. Table S3: Gas-phase geometrical quantities and dipole moments for protonated phosphates. Table S4: Gas-phase geometrical quantities and dipole moments for phosphoranes. Table S5: Structures and energies for protonation states of neutral and monoanionic phosphoranes. Table S6: Calculated thermodynamic data for isomerizations of pentavalent phosphorus compounds. Table S7: Calculated thermodynamic data for ligand substitution reactions of phosphorus compounds. Table S8: Calculated intramolecular hydrogen bond lengths and angles for neutral phosphoranes (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0356277