

10
3-24-97 JS(1)

SANDIA REPORT

SAND97-8217 • UC-401
Unlimited Release
Printed January 1997

M97052277

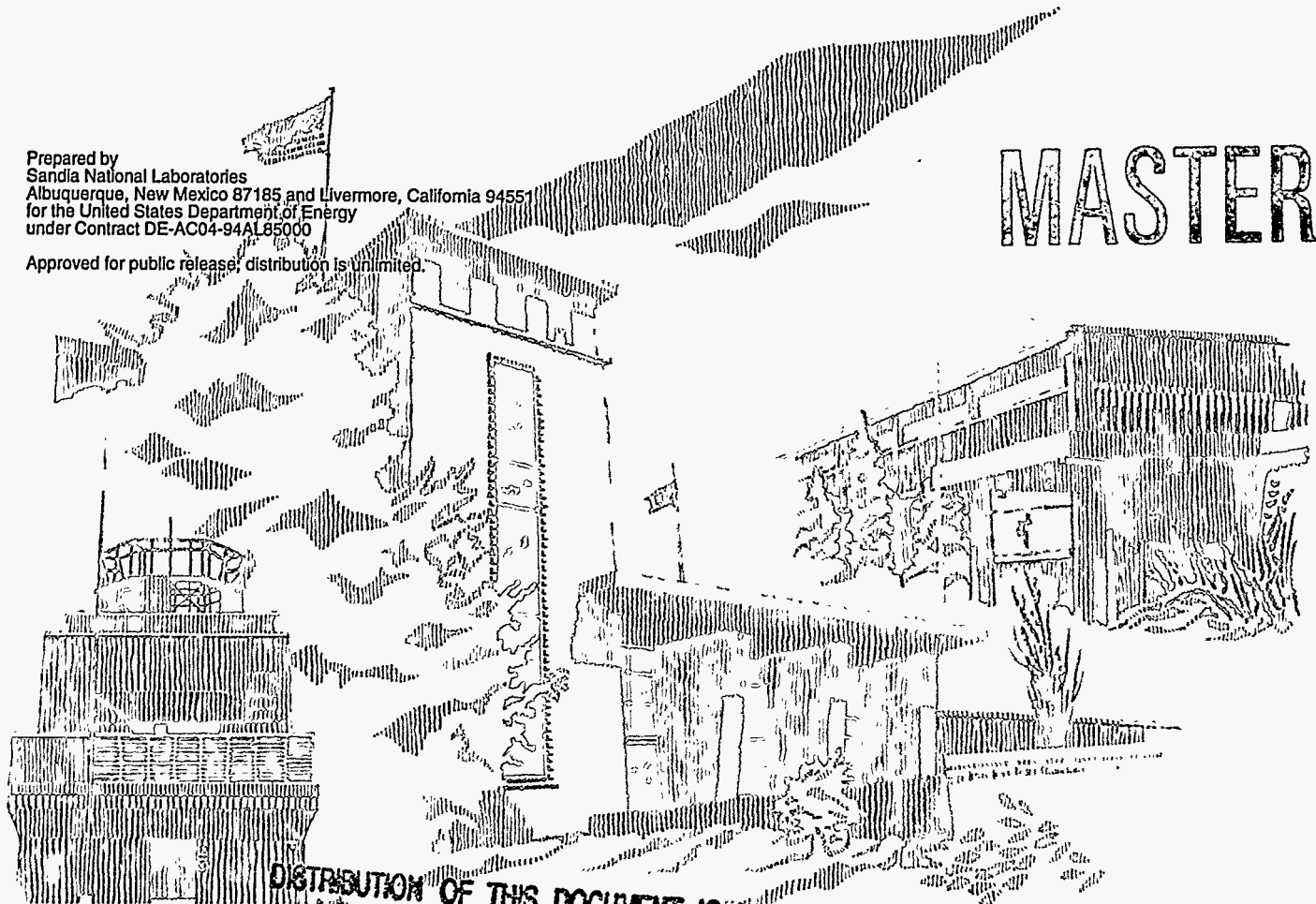
Grid-Based Methods for Biochemical *Ab Initio* Quantum Chemical Applications

Michael E. Colvin, Jeffrey S. Nelson, Eugenia Mori, Curtis L. Janssen,
Carl F. Melius

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94551
for the United States Department of Energy
under Contract DE-AC04-94AL85000

Approved for public release; distribution is unlimited.

MASTER



DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

UH

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of the contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors or subcontractors.

This report has been reproduced from the best available copy.

Available to DOE and DOE contractors from:

Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge TN 37831

Prices available from (615) 576-8401, FTS 626-8401.

Available to the public from:

National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd.
Springfield, VA 22161

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Grid-based Methods for Biochemical *Ab Initio* Quantum Chemical Applications

Michael E. Colvin, Jeffrey S. Nelson, Eugenia Mori, Curtis L. Janssen
and Carl F. Melius

Contribution from Scientific Computing Department, MS 9214, Sandia National Laboratories, Livermore, CA 94551 and (JSN) Semiconductor Materials and Device Sciences Department, MS 0601, Sandia National Laboratories, Albuquerque, NM 87123

Abstract

Ab initio quantum chemical methods are seeing increased application in a large variety of real-world problems including biomedical applications ranging from drug design to the understanding of environmental mutagens. The vast majority of these quantum chemical methods are “spectral”, that is they describe the charge distribution around the nuclear framework in terms of a fixed analytic basis set. Despite the additional complexity they bring, methods involving grid representations of the electron or solvent charge can provide more efficient schemes for evaluating spectral operators, inexpensive methods for calculating electron correlation, and methods for treating the electrostatic energy of solvation in polar solvents. The advantage of mixed or “pseudospectral” methods is that they allow individual non-linear operators in the partial differential equations, such as coulomb operators, to be calculated in the most appropriate regime. Moreover, these molecular grids can be used to integrate empirical functionals of the electron density. These so-called density functional methods (DFT) are an extremely promising alternative to conventional post-Hartree Fock quantum chemical methods. The introduction of a grid at the molecular solvent-accessible surface allows a very sophisticated treatment of a polarizable continuum solvent model (PCM). Where most PCM approaches use a truncated expansion of the solute’s electric multipole expansion, e.g. net charge (Born model) or dipole moment (Onsager model), such a grid-based boundary-element method (BEM) yields a nearly exact treatment of the solute’s electric field. This report describes the use of both DFT and BEM methods in several biomedical chemical applications.

Section 1: Biomedical Applications of Density Functional Methods

Cisplatin:

We have applied plane-wave based density functional methods to the potent anticancer drug cisplatin (*cis*-diamminedichloroplatinum(II)) and its inactive isomer transplatin. These calculations included the first *ab initio* optimizations available for this important compound and will be used to help understand its mode of activity. Since the accidental discovery of its antitumor activity, *cis*-diamminedichloroplatinum(II) (cisplatin) has become an important drug in curative treatments of several forms of cancer.¹ Cisplatin is known to work by binding to the DNA of cancer cells which inhibits cell division and eventually leads to cell death. The reason why cisplatin preferentially kills cancer cells is still unknown, but an interesting clue to the mechanism of cisplatin is that its isomer, *trans*-diamminedichloroplatinum(II) (*trans*-DDP) exhibits no anticancer activity. Studies of the differential activity of the two isomers indicate that they bind to DNA at the same rate and have similar rates of cellular uptake.² These results suggest that the difference in anti-cancer activity might arise from different rates of cellular repair of their DNA adducts. This in turn strongly suggests that the *cis*- and *trans*-isomers form structurally different adducts (and further that the DNA binding is bifunctional).

Numerous experimental studies have been performed on the binding of Pt compounds to DNA.²⁻⁴ These have been augmented by theoretical studies of the Pt-DNA adducts, including molecular-mechanics(MM),^{5, 6} and MM augmented by extended Hückel⁷ and *ab initio* calculations.⁸ The Pt molecular mechanics potentials have been derived from experimental crystal and spectroscopic data as well as from relativistic-pseudopotential self-consistent-field calculations.⁹ Single determinant determinants are generally known to be poor descriptions of transition metal compounds and the *ab initio* calculation of spectroscopically accurate properties of even PtH is challenging.¹⁰

Density Functional Theory (DFT) is a promising alternative to conventional *ab initio* methods for including electron correlation.¹¹ In particular, local-spin density approximation (LDA) DFT has been shown to yield reasonable molecular properties including equilibrium structures and dissociation energies for a wide range of transitional metal compounds.¹²⁻¹⁴ In this study we applied a plane-wave basis LDA to optimize all molecular structural parameters of both *cis*- and *trans*- diamminedichloroplatinum(II). (cisplatin and trans-DDP).

The electronic and structural properties of cisplatin and trans-DDP were calculated within the LDA of DFT^{15, 16}, using the generalized norm-conserving pseudopotentials of Hamann¹⁷⁻¹⁹, the exchange-correlation potential of Ceperley-Alder²⁰, and a highly converged plane-wave basis. The eigenstates are obtained using a state-by-state preconditioned conjugate gradient algorithm proposed by Teter, Payne and Alan.²¹ The calculations were performed on a nCUBE2-1024 multiple-instruction-multiple-data(MIMD) parallel supercomputer (see Nelson et. al.²² for details of the parallel implementation). For our purposes here, the nCUBE2 has 1024 processor, each with a peak performance of about 2 Mflops and 4 Mbytes of memory.

In a plane-wave basis, the quality of the calculation is directly related to only a small number of input parameters: the convergence of the plane-wave expansion; the quality of the ionic pseudopotentials; and the size of the periodic supercell(in the case of a molecular calculation). The plane-wave expansion for N is by far the most demanding due to the lack of core-orthogonality for the 2p-state; although Pt contains 5d-electrons the orthogonality with the core d-states weakens the pseudopotential. In Table 1 we establish the convergence and quality of our basis by performing equilibrium calculations

of NH_3 , Cl_2 and crystalline Pt. With a plane-wave expansion of 120 Ry for NH_3 and 40 Ry for Pt and Cl_2 , the bond lengths are converged to less than 0.02\AA and within 1% of the experimental values. This level of agreement with experiment is typical of highly converged LDA calculations. For the cisplatin and trans-DDP we employed a plane-wave expansion of 120 Ry (72000 pw) and a supercell of dimension (18.0 a.u., 18.0 a.u., 10.0 a.u). We found this to be sufficient to effectively isolate the molecule ("no interaction" with periodic replicas) and also allowed a convenient Fast-Fourier-Transform parameter set of (128,128,64).

The predicted structural properties of both isomers are given in Table 2 along with the X-ray crystal structures. The Pt-N bond lengths agree to within the experimental error bars for both isomers. The Pt-Cl bond lengths are $\sim 0.06\text{-}0.09\text{\AA}$ shorter than those in the crystal structure. The bond angles exhibit considerable differences between the two isomers and between experiment and theory. Both experimental structures are found to be very nearly square structures. These large bond angle differences may result from the formation of different hydrogen bonds between the chlorines and amine hydrogens in the crystal and the *in vacuo* theoretical structure. Milburn and Truter²³ cite structural evidence for the presence of intermolecular hydrogen bonds connecting both amines on a single cisplatin molecule to the same chlorine atom on the adjacent molecule in the crystal structure. Such intermolecular bonds would act to decrease the $\angle\text{N-Pt-N}$ bond angle. In our predicted structure, both amines are oriented to allow a such a bond between a hydrogen and the chlorine (see Figure 1). Although there are no rigorous criteria defining hydrogen bonds ($\text{A-H}\cdots\text{B}$), they are usually indicated for appropriate donor-acceptor pairs when the $\text{H}\cdots\text{B}$ distances is less than the sum of Van der Waals radii for H and B.²⁴ In this case the $\text{Cl}\cdots\text{H}$ distance is 2.337\AA as compared to a sum of Van der Waals radii of

3.0Å.²⁵ The presence of these intramolecular hydrogen bonds will be to decrease the \angle Cl-Pt-N bond angles.

These results clearly demonstrate the utility of the LDA DFT method to molecules containing transition metals bound to second and third row atoms. Moreover, these results indicate that intermolecular interactions within the cisplatin and trans-DDP crystal structures are likely to be leading to a nearly 0.01Å lengthening of the Pt-Cl bond.

Table 1: Convergence and quality of plane-wave basis.

N-H Bond Length (Å)		Crystalline Pt Lattice Parameter(Å)		Cl ₂ bond length (Å)	
50 Ry	1.05	40 Ry	3.88	40 Ry	1.96
90 Ry	1.03	80 Ry	3.88	exp	1.988
120 Ry	1.02	exp	3.92		
exp	1.01				

Table 2: DFT-calculated structural properties of *cis*- and *trans*-diamminedichloroplatinum

Structural Parameters	This work		Experiment ²³	
	<i>cisplatin</i>	<i>trans</i> -DDP	<i>cisplatin</i>	<i>trans</i> -DDP
Pt-Cl	2.24 Å	2.26 Å	2.328(0.009)Å 2.333(0.009)Å	2.32(0.01)Å
Pt-N	2.02 Å	1.99 Å	1.95(0.03)Å 2.05(0.04)Å	2.05(0.04)Å
\angle Cl-Pt-Cl	96.7°	-	91.9(0.3)°	-
\angle N-Pt-N	97.3°	-	87(1.5)°	-
\angle Cl-Pt-N	83.0°	88.4°	88.5(0.9)° 92(1)°	88.5(1.0)° 91.5(1.0)°

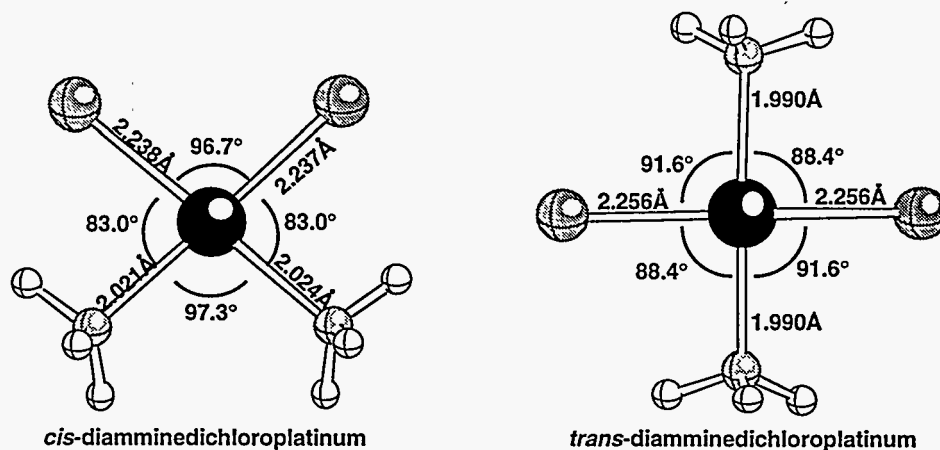


Figure 1: Local density functional optimized structures of *cis*- and *trans*-diamminedichloroplatinum(II)

Glycine Zwitterion and Neutral:

Glycine zwitterion is a particularly good test case since it involves intramolecular hydrogen bonding--known to be a difficulty for fully local density functionals. For this reason we have run a series of test calculations on this species using both spectral Hartree-Fock and fully numerical DFT. These results show clearly the need for non-local corrections to the exchange-correlation term in DFT. For example, at the fully optimized 6-31G* Hartree-Fock geometry (where all residual forces are zero), the local density approximation DFT method finds significant non-zero terms in the gradient matrix. We extended this DFT study of glycine and glycine zwitterion by performing optimizations of these structures using the non-local Becke-Lee-Yang-Parr exchange-correlation functional.^{26, 27} These results were significant in showing that the non-local DFT agreed well with the MP2 optimized structures in finding no zwitterionic minimum on the gas-phase potential energy surface. Table 3 compares the glycine neutral structures for HF, MP2 and DFT (using the BLYP functional), using the following atomic numbering shown below. As can be seen in this table, the DFT optimization yields structural properties that are closer to the MP2 results than to the HF. Except for the C-C bond length, the DFT parameters are close to the MP2 result. Interestingly, the MP2 C-C bond length is nearly identical to the HF result, contradicting the usual result that electron correlation acts to increase bond lengths. This result should be tested using (as yet unfeasible) more highly correlated *ab initio* methods, which may well concur better with the inexpensive DFT bond length.

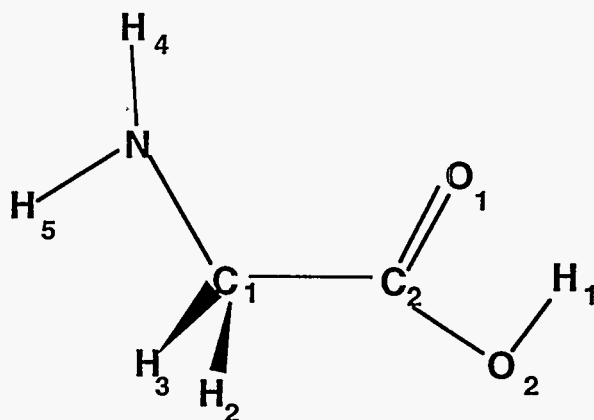


Table 3: Structural properties of glycine zwitterion calculated using HF, MP2 and DFT(BLYP) using a 6-31G** basis set.

Structural Parameter	HF/6-31G**	MP2/6-31G**	BLYP/6-31G**
C ₁ -C ₂	1.508 Å	1.509 Å	1.527 Å
C ₂ -O ₁	1.187 Å	1.218 Å	1.224 Å
C ₁ -O ₂	1.327 Å	1.356 Å	1.371 Å
O ₂ -H ₁	0.948 Å	0.971 Å	0.984 Å
C ₁ -N	1.441 Å	1.454 Å	1.470 Å
∠C ₁ -C ₂ -O ₁	125.11°	124.98°	125.06°
∠C ₁ -C ₂ -O ₂	111.95°	111.55°	111.66°
∠N-C ₁ -C ₂	110.31°	109.20°	109.83°
∠N-C ₁ -C ₂ -O ₁	20.75°	24.59°	22.88°

Guanine Exocyclic Amines:

Additionally, we have used a mixed spectral/grid DFT to study the structure of the exocyclic amines in heterocyclic biological compounds including the nucleic acid base guanine. Even for these moderately large-sized molecules, the non-local DFT optimizations will require a small fraction of the CPU time of the MP2 optimizations. Since the DFT has a much lower computational scaling, much larger savings will be achieved for larger molecules.

Figure 2 shows the structures for both fully optimized and (constrained) planar guanine using three methods: Hartree-Fock, second-order Møller-Plesset perturbation theory (MP2) and a density functional theory using the Becke-Lee-Yang-Parr exchange-correlation functional. As can be seen in this figure, there is considerable variation in the geometrical properties of the exocyclic amine between the different methods. A useful measure of the amine planarity is the sum of three R-N-R bond angles denoted by " Σ ". (For a planar amide $\Sigma = 360^\circ$ for a typical amine $\Sigma \approx 330^\circ$) For the HF guanine structure, $\Sigma=348.6^\circ$, compared to 338.9° and 341.6° for the MP2 and DFT structures. Another measure of the treatment of the exocyclic amine is the C-N bond distance connecting the amine to the guanine C₂. For HF this bond distance is 1.359Å, with the MP2 and DFT giving more similar lengths of 1.383Å and 1.391Å, respectively. These results illustrate that the DFT methods, which have nearly the same computational cost as HF, yield structural results in much closer agreement with MP2 optimizations that are more than ten times more costly for this size molecule.

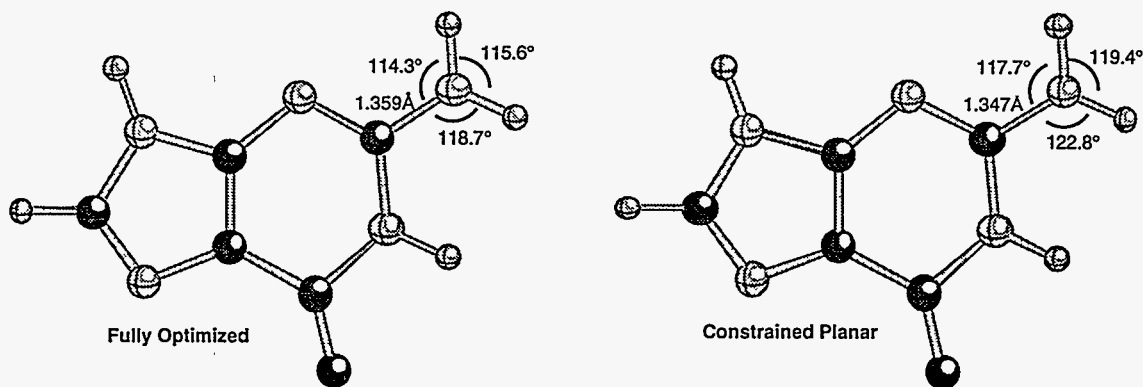


Figure 2a: HF/6-31G** guanine structures

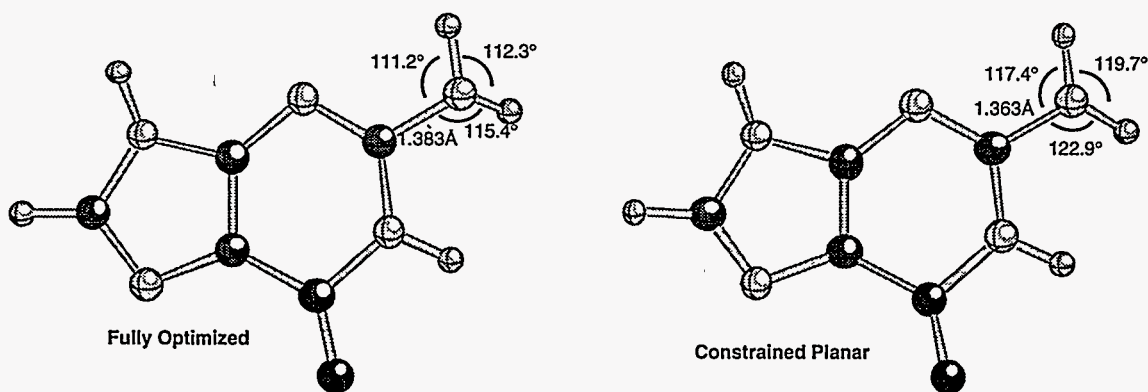


Figure 2b: MP2/6-31G** guanine structures

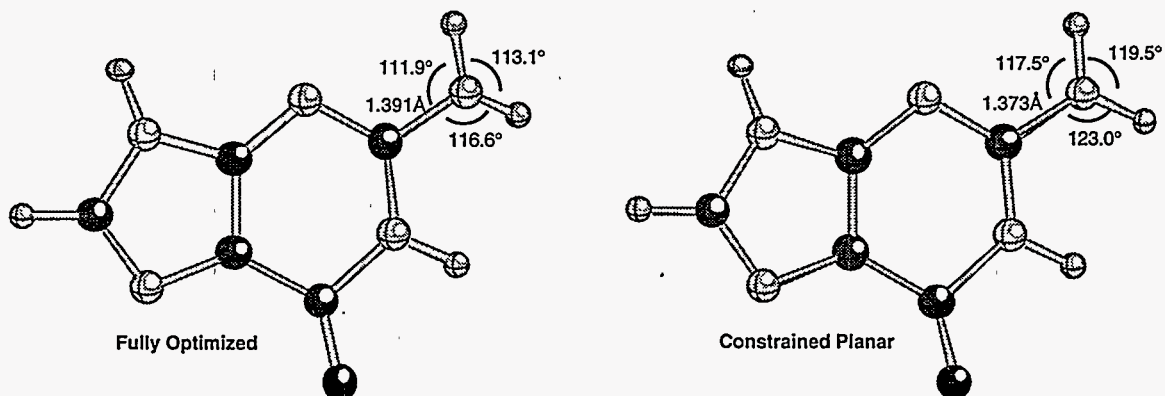


Figure 2c: DFT BLYP/6-31G** guanine structures

Figure 2: Guanine structures optimized with Hartree-Fock, Second order Møller-Plesset Perturbation Theory (MP2), and density functional theory (BLYP functional) using a 6-31G** basis set

The structural differences in the exocyclic amines between the three computational methods is also evident in the relative energies of the planar versus pyramidal amines. Table 4 lists the total and relative energies of these structures. The HF method yields a much lower potential energy barrier to planarity (0.38 kcal/mole) than either the MP2 or DFT method. However, there is significant variance between the MP2 and DFT barriers, 1.48 vs. 0.88 kcal/mole, suggesting that an appropriate protocol would be to use the DFT with the BLYP functional for the expensive geometry optimization, followed by a single-point MP2 energy calculation to determine accurate energy differences. Indeed, using this approach for the guanine barrier to amine planarity yields 1.46 kcal/mole (see last row of Table 4), in excellent agreement with the MP2-optimized results.

Table 4: Total energies and energy barrier between the constrained planar and fully optimized guanine structures using three different *ab initio* methods. The final row gives the results for MP2 energies calculated at the BLYP optimized structures. All calculations were performed with a 6-31G** basis set. Total energies in Hartrees, barrier energies in kcal/mole.

Method	E(planar)	E(fully opt)	Barrier
HF	-539.412009	-539.412563	0.35
MP2	-541.038787	-541.041144	1.48
BLYP	-542.416867	-542.418276	0.88
MP2//BLYP	-541.035992	-541.038313	1.46

Section 2: Grid-based Boundary Element Methods

Application to Acetic Acid Carboxylate Conformation:

Solvating a molecule can have profound effects on its energetics and can affect its preferred geometry relative to its optimal geometry in the gas phase. The effects of solvation of molecules in water in particular have been most widely studied and are critical in many problems of biological interest. To investigate how well aqueous solvation can be treated with grid-based boundary element solvation models (BEM), we have studied the conformational preference of acetic acid in the gas phase and in solution phase using *ab initio* calculations for the acetic acid molecule coupled with various continuum models for water. Acetic acid is a relatively simple molecule which contains the important carboxylate functional group. The purpose of this study is to compare the gas phase geometry and energetics with those in the solution phase and to comment on the changes acetic acid assumes in solution as well as to compare the various continuum solvent models for consistency or differences among methods. We also examine the energetics of an acetic acid dimer complex with the same methods in the gas and solution phases.

The carboxylate moiety has two local energy minima: the syn form, Fig 3a, which has the acidic hydrogen located between the two carboxylate oxygens, and the anti form, Fig 3b, (also referred to as the E- and Z-forms, respectively) in which the acidic hydrogen is located to the outside of the oxygen to which it is attached. A transition state exists between the two states obtained by rotating the acidic hydrogen about the O-C bond out of the plane of symmetry of the molecule. In the gas phase, the syn form is calculated to be about 5 kcal/mol more stable than the anti form,^{28, 29} a fact rationalized by the beneficial interactions of the slightly positively charged hydrogen with two partially negatively charged oxygens. In the solution phase, the energy difference between the two forms is less pronounced because the anti form has a larger dipole moment which solvates

favorably in a polar medium. Previous Monte Carlo simulations³⁰⁻³³ have calculated the syn conformer still to be the preferred geometry in the solution phase by 1.1 +/- 0.3 kcal/mol³⁰ and by 1.7 +/- 0.3 kcal/mol.³¹ We performed gas phase and continuum solvent calculations and confirmed these results. Our calculations find the solvated syn confirmation to be 1.7 to 3.2 kcal/mol more stable than the anti. We further find that in solution a dimer complex (Figure 4) is favored over two syn monomers by a few kcals/mol.

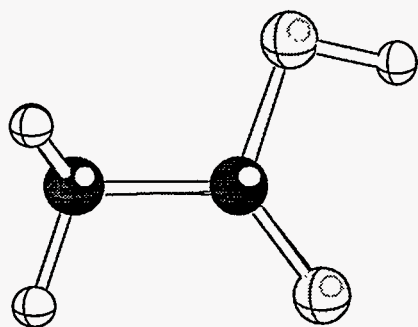


Figure 3a: Acetic acid (syn form)

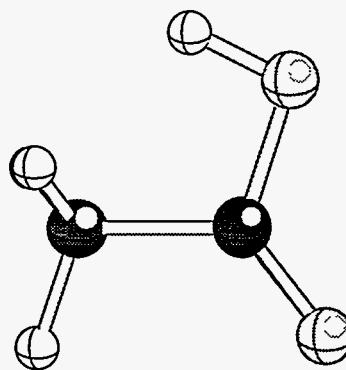


Figure 3b: Acetic acid (anti form)

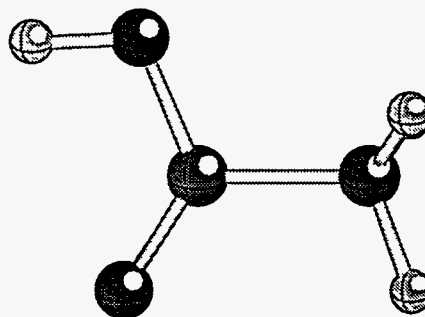
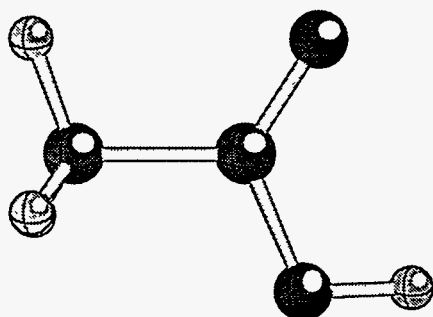


Figure 4: Acetic acid dimer, HF/6-31G optimized geometry.**

Hartree-Fock gas phase geometry optimizations were performed with Gaussian-92³⁴ for the carboxylic acid monomer and with Gaussian-94³⁵ for the dimer. Geometry optimizations using the Onsager solvent model (SCRF) were performed using Gaussian-92 for the monomer and Gaussian-94 for the dimer. Our implementation of a grid-based

boundary element method (called MSURF) was used with both the gas phase and the self-consistent reaction field geometries to calculate a solvation energy by generating a solvent accessible surface of the molecule and calculating a set of self-consistent surface charges that interact with a dielectric medium. The program MSURF,³⁶ couples with Gaussian-92 to calculate the final wave-function and energy. Finally, the implementation of Tomasi's polarizable continuum solvation model in Gaussian-94 (SCIPCM) was used to optimize geometries for both the monomer and dimer in the presence of a dielectric medium. Frequency calculations were done on all final states to ensure that each state was a true energy minimum.

The calculated energies of the syn and anti forms of carboxylic acid, the transition state separating them, and the dimer are presented in Table 5. The relative free energies, ΔG , of the same complexes, relative to the corresponding gas phase energy, are shown in Table 6 for at 298K. Table 6 also lists an experimentally measured solvation free energy of -6.7 kcal/mol.³⁷ The contributions of the entropy and zero-point vibrational energy (ZPE) to the free energies ($G = E - TS + ZPE$) are similar for all the monomers and do not change the relative energy orderings of any of the calculations. Table 7 shows the energy difference between the syn and anti forms for the various calculations ($\Delta G(\text{anti-syn}) = G_{\text{anti}} - G_{\text{syn}}$) as well as the difference in energy between the transition state and the syn form ($\Delta G(\text{TS-syn}) = G_{\text{TS}} - G_{\text{syn}}$).

Table 6 shows the expected stabilizing effect that solvation has on this polar molecule. The SCRF calculations have the least solvation stabilization relative to the gas phase, ~1-5 kcal/mol; the SCIPCM calculations have a moderate amount of stabilization with ~5-10 kcal/mol; and the MSURF calculations have the most solvent stabilization at ~13-19 kcal/mol.

As seen in Table 7, in all calculations, the syn monomer is more stable than the anti in both the gas phase and solution phase. The syn is favored over the anti by 6.9 kcal/mol in the HF gas phase calculation. In the solution phase, the syn is still favored, although due to the better solvation of the more polar anti form, the syn is favored by lesser amounts ranging from 1.7 kcal/mol in the MSURF calculation performed with the SCRF structure to 3.2 kcal/mol in the SCIPCM calculation. Solvation also lowers the energy barrier height separating the syn and anti forms as solvation stabilizes the polar transition state. The gas phase barrier between the transition state and the syn form is 12.5 kcal/mol, but in solution the barrier is reduced to 10.6 kcal/mol in the SCRF calculation and to 6.7 kcal/mol in the SCIPCM calculation. The SCRF results are a direct outcome of the higher dipole moments for the anti and TS structures (see last row of Table 8). The increased solvation energy predicted by the MSURF BEM and SCIPCM methods indicate the importance of the higher electric moments in describing the solvent-solute electronic interactions.

We also calculate a potential energy curve for the conversion of the syn monomer to the anti structure obtained by rotating the acidic hydrogen from the syn position to the anti. Figure 4 shows the electronic energies calculated by the different methods as a function of the carbon-hydroxyl dihedral angle. A dihedral angle of 0 degrees corresponds to the syn structure and a dihedral angle of 180 degrees corresponds to the anti structure. In these calculations, the dihedral angle is held fixed and all other coordinates are optimized. The energies displayed are all relative to the syn conformer for that type of calculation. Again, all solvated models offer more stabilization for the conversion with the SCIPCM offering the most and the SCRF offering the least. The MSURF calculations performed on the 2 different geometries, the gas phase and the SCRF optimized geometry, are nearly identical, supporting that the structural differences

between the gas phase and solution phase geometry (*vide infra*) are not significant compared to the energetic changes induced by the solvation.

The structural differences between the syn and anti form and differences in the atomic charges as calculated by natural population analysis (NPA) are not large. The largest difference in bond length between the SCIPCM syn and anti structures, for example, is 0.0047 Å in the C-C distance (1.5047 Å in the anti form, 1.5000 Å in the syn). The largest bond angle difference is 4.04 degrees for the C-C-hydroxyl oxygen angle (116.28 degrees in the anti, 112.24 degrees in the syn). And the largest difference in the atomic charges is 0.03 electron charges on the carbonyl oxygen (-0.70 a.u. in the anti, -0.67 in the syn).

The structural relaxation of acetic acid in the gas phase versus in solution is also small. Comparing the syn Hartree-Fock gas phase structure with the SCIPCM solvated structure, the largest change in a bond length is 0.0083 Å for the C-hydroxyl oxygen bond length (1.3307 Å gas phase, 1.3224 Å solvated). The largest change in a bond angle is 0.43° (125.71° gas phase, 125.28° solvated). And the largest difference in NPA charges is 0.004 on the carbonyl oxygen (-0.698 au gas phase, -0.702 au solvated). The energies of the solvated gas phase structure and of the solvated SCRF structure differed by less than 0.5 kcal/mol. Structural changes as well as changes in charges induced by solvation changes are not large for this molecule.

In the syn form, we investigate the nature of the interaction of the acidic hydrogen with the carbonyl oxygen to determine whether it could be classified as a hydrogen bond. Using the criteria described by Vinogradov,²⁴ hydrogen bonds are characterized by an attractive interaction energy of 3-10 kcal/mol, a hydrogen-carbonyl oxygen distance less than the sum of their van der Waals radii, an O-H--O angle of 20° within linear, and finally a decreased OH stretching frequency of typically 200 cm⁻¹. The syn monomer is

more stable than the anti, but by margins between 1.7 and 2.9 kcal/mol for the different calculations. In the SCIPCM optimized syn geometry, the carbonyl oxygen - hydrogen distance is 2.8889 Å; the OH-O bond angle is 73.00 degrees; and the OH stretching frequency is 4129.02 cm^{-1} . In the SCIPCM optimized anti structure, the O-H bond length is 0.9435 Å which is even larger than the syn form, and the O-H stretching frequency is 4188.00 cm^{-1} which is not appreciably tighter than the syn form. Based on these findings, we would not classify the interaction of the acidic hydrogen with the carbonyl oxygen as a hydrogen bond.

The energies and free energies of the dimer calculated by the same methods are shown on the fourth lines of Table 5 and Table 6. Since the dimer complex has no overall dipole, the SCRF method, which approximates a solvent interaction energy based only on the overall dipole moment, gives the same optimized geometry and energy as the Hartree Fock method. In all cases, the dimer free energy is lower than that of two separate syn monomers which suggests that acetic acid may pair into dimers in aqueous solution.

By the criteria described above, the interaction in the dimer complex between the acidic hydrogen on one monomer and the carboxylic oxygen is a typical hydrogen bond. For the SCIPCM calculation, the H--O bond distance is 1.8318 Å, O-H--O bond angle is 178.08°, and the OH stretching frequencies of 3813.15 cm^{-1} and 3860.72 cm^{-1} are approximately 200 cm^{-1} lower in energy than that for the syn monomer. Furthermore, the C=O stretches in the dimer appear at 1937.75 cm^{-1} and 1977.18 cm^{-1} , whereas it appears at 2038.19 cm^{-1} in the monomer.

This study of acetic acid and its dimer shows that a grid-based boundary element continuum model of aqueous solvation can give yield results in good agreement with explicit-water Monte Carlo methods. However, the simpler Onsager (SCRF) model does not give nearly as good agreement with these MC results and underestimates the total solvation energies of the acetic acid conformations. Finally, we found that optimization

in the presence of the dielectric continuum has very small effect on the structural or energetic properties of acetic acid.

Table 5: Electronic energies, in Hartrees, of acetic acid in the syn form, the anti form, the transition state, and a dimer complex as computed by different methods.

E(a.u.)	HF	SCRF	MSURF (HF geom)	MSURF (SCRF geom)	SCIPCM
syn	-227.82217	-227.82327	-227.84396	-227.84413	-227.83231
anti	-227.81082	-227.81822	-227.84085	-227.84140	-227.82769
TS	-227.80070	-227.80465	-227.82758	-227.82769	*
dimer	-455.66904	-455.66904	-455.69275	-455.69275	-455.67735

*could not be calculated

Table 6: Free energies of acetic acid in the syn form, anti form, the transition state between the conversion of the two forms, and a dimer complex in kcal/mol as computed by the different methods, relative to the gas-phase HF energy.

G _{rel}	HF.	SCRF	MSURF (HF geom)	MSURF (SCRF geom)	SCIPCM
syn	0	-0.72	-13.68	-13.81	-6.73
anti	0	-4.55	-18.85	-19.10	-10.46
TS	0	-2.63	-16.87	-17.08	*
dimer	0	0	-14.88	-14.88	-5.513

*could not be calculated

Table 7: Free energy differences, in kcal/mol, between the anti and syn conformers of acetic acid and the activation free energy for their conversion as computed by different methods. Monte Carlo (MC) simulations of the same quantities in others' works are listed for comparison.

	HF	SCRF	MSURF (HF geom)	MSURF (SCRF geom.)	SCIPCM	MC (Gao)	MC (Pranata)
$\Delta G(\text{anti-syn})$	6.95	3.12	1.78	1.65	3.21	1.1	1.7
$\Delta G(\text{TS-syn})$	12.54	10.64	9.35	9.27	*	8.3	

*could not be calculated

Table 8: The NPA charges, in units of electron charges, of the carboxylate oxygen, the carboxylate carbon, the hydroxyl oxygen, and the hydroxyl hydrogen for the acetic acid in the syn form, the anti form, and the transition state (TS) between them in the HF gas phase calculation. And the net dipole moments in units of Debye

	syn	TS	anti
O= charge	-0.698	-0.631	-0.662
C charge	0.986	0.973	0.977
O charge	-0.795	-0.832	-0.781
H charge	0.518	0.518	0.506
Dipole moment	1.83	3.56	4.74

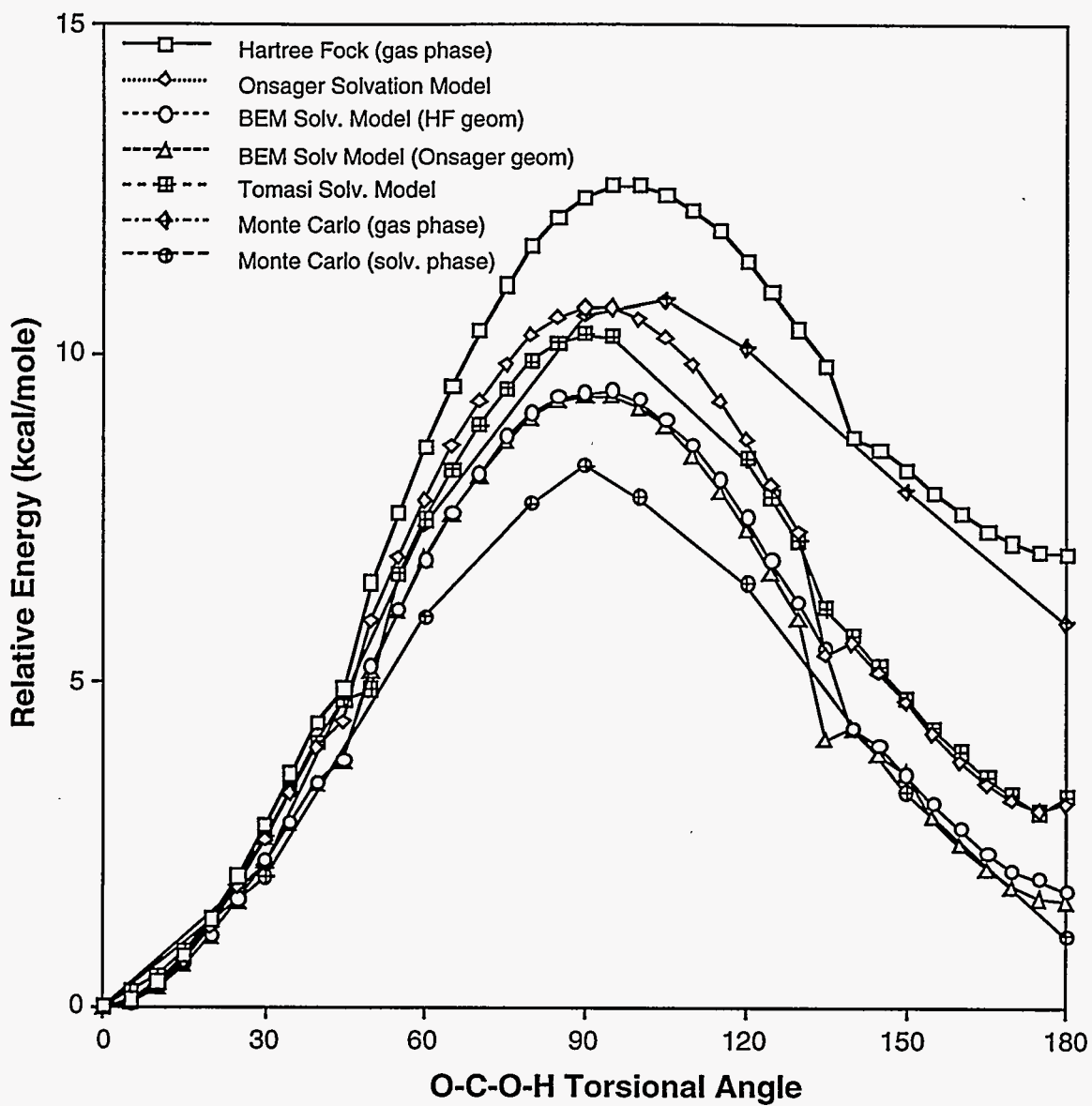


Figure 5: Torsional energy surfaces for Acetic Acid calculated using a variety of different methods.

Application to Pyrophosphate Hydrolysis:

The hydrolysis of pyrophosphate to form two orthophosphates is coupled to virtually all biosynthetic reactions. In many cases, such as the synthesis of polynucleotides, this hydrolysis determines the direction of the reaction. Numerous measurements have indicated that the reaction free energy of inorganic pyrophosphate hydrolysis is approximately -5 kcal/mole. However, despite these empirical results, a detailed understanding of the energetic factors contributing to this reaction energy is still lacking. In particular, it is unknown to what extent this reaction energy is determined by intermolecular versus intramolecular effects. The high energy phosphate bond was originally thought due to intramolecular effects ranging from simple electrostatic repulsion to more complex "opposing resonance" effects. This hypothesis was challenged by George *et al.*³⁸ who found that the reaction enthalpy for pyrophosphate hydrolysis was *inversely* related to the total charge of the phosphates with $\Delta H^0 = -7.6$ for the hydrolysis of $H_4P_2O_7$ and $\Delta H^0 = -3.7$ for the hydrolysis of $P_2O_7^{4-}$, exactly the opposite of that expected if electrostatic repulsion was dominating the reaction energetics. Experiments on the hydrolysis reaction in non-polar solvents suggest that the negative reaction energy may be entirely a result of differing solvation energies of the reactant and products. Romero³⁹ determined that for pyrophosphate hydrolysis, $\Delta G^\circ = -5.1$ kcal/mole in H_2O , $\Delta G^\circ = -1.6$ kcal/mole in 50% Me_2SO , and $\Delta G^\circ = +0.1$ kcal/mole in 50% poly(ethylene glycol).

Pyrophosphate hydrolysis is a particularly difficult problem for *ab initio* studies because the reactants and products are multiply deprotonated at neutral pH. To date,

most predictions of the reaction enthalpy have been for the fully protonated 1gas phase×
reaction:



with the goal of establishing whether the unsolvated reaction is endothermic or exothermic. These studies have not unambiguously resolved this issue, predicting a wide range of gas-phase hydrolysis enthalpies ranging from -0.4 to -13.4 kcal/mole.⁴⁰⁻⁴³

Methods:

In this report we describe large basis set *ab initio* calculations of the aqueous-phase reaction enthalpy and free energy of pyrophosphate hydrolysis using several methods based on a dielectric continuum model of the aqueous solvent to estimate to solvation energies of the reactants and products. The molecular structures of all protonation states of pyrophosphate and orthophosphate were optimized within the Hartree-Fock approximation using both a 6-31G* and a larger 6-311++G** basis set. Single point second order Moller Plesset Perturbation theory single point energy calculations (with the core electrons frozen) where performed using both basis sets at the HF/6-31G* and HF/6-311++G** optimized geometries. Additionally, the entropy of formation of each species was estimated using standard thermodynamic approximations.⁴⁴ All *ab initio* calculations were performed using Gaussian 92.³⁴

In order to model solvation effects and to calculate accurately atomic charge distributions, we have used several theoretical methods based on the polarizable continuum model for solvents. The first of these is a first-principles method developed

independently by Miertus and Tomasi⁴⁵ and Rashin.⁴⁶ This method is described in detail elsewhere, so only a brief description is given here.³⁶ The surrounding medium is treated as a continuum dielectric in a procedure involving several steps. First, the molecular geometry and electronic wave function are obtained from the Hartree-Fock calculation. Next, the interface between the molecule and the dielectric continuum of the surrounding medium is defined by the contact surface of the solvent molecule rolling around the scaled Van der Waal's surface of the molecule. The boundary-element-method is used to represent the triangulated surface, as developed by Zauhar and Morgan.⁴⁷ The electric potential is used to determine effective atomic charges (using a least-squares fit with constraints for net charge and dipole moments). Applying Gauss' law, the resulting electric field at the molecule surface (obtained from the atomic charges) is used to determine the effective charge of the dielectric continuum at the surface. These induced charges at the surface are then included in the one-electron portion of the HF Hamiltonian and the wavefunction is reoptimized. New atomic charges are calculated from the reoptimized wavefunction and the whole procedure is iterated to convergence (typically 8 iterations). A single point MP2 energy calculation including the converged image charges is used to determine the total solvated energy. The surface area and volume of the molecule is used in an approximate expression⁴⁸ to determine the free energy of cavitation (the non-electrostatic energy to create a cavity in the dielectric continuum containing the molecule). The theoretical free energy of solvation is the sum of the free energy of hydration (calculated as the difference between the *in vacuo* and solvated MP2 energies) and the free energy of cavitation. Note that an additional term of 1.84 kcal/mole

is added for the work required to compress the solute from the standard gas-phase volume of 22.4 liters to the standard aqueous-phase volume of 1 liter.

A study of the performance of this model³⁶ indicates that it predicts absolute ΔG_{soln} s accurate to 5 kcal/mole for several neutral, ionic, and zwitterionic compounds. Larger errors of up to 11 kcal/mole occur for small non-polar molecules such as neutral trimethylamine where ΔG_{soln} is not dominated by electrostatic solute-solvent interactions. For these compounds, dispersion and solvent reorganization energies, not included in these predictions, act to cancel out the cavitation energy.

Gas-Phase Results:

In the literature for there is wide spread disagreement about the optimal configuration of gas-phase $\text{H}_4\text{P}_2\text{O}_7$. Our no-symmetry HF/6-31G* and HF/6-311++G** optimizations of pyrophosphate yielded a C_2 symmetry structure with the two intramolecular P O ... H O hydrogen bonds linking opposite ends of the pyrophosphate. The optimized $\text{H}_4\text{P}_2\text{O}_7$ determined in this work is shown in Figure 6. $\text{H}_4\text{P}_2\text{O}_7$ structure differs from that of Saint-Martin.⁴³ (and all earlier structures) primarily by the presence of the hydrogen bonds. Interestingly, Ewig and Van Waser⁴² found a very similar optimal conformation for diphosphonic acid ($\text{H}_4\text{P}_2\text{O}_5$) but did not find any intramolecular hydrogen bonds for pyrophosphate.

In aqueous solution at physiological pH (~7.4), orthophosphate and pyrophosphate exist largely as the deprotonated ions HPO_4^{2-} and $\text{HP}_2\text{O}_7^{3-}$. As has been widely noted in the literature, gas phase reaction enthalpies of anionic phosphates provide little insight into the aqueous phase reaction for two reasons. First, the repulsion between

the unscreened negatively charged phosphates leads to extremely distorted pyrophosphate conformations. Second, since multiply charged anions will spontaneously ionize in the gas phase, *ab initio* calculations of the structure and energetics are extremely suspect. These problems should be particularly acute for the di- and tri-anionic orthophosphates, HPO_4^{2-} and PO_4^{3-} , and the tri- and tetra-anionic pyrophosphates, $\text{P}_2\text{O}_7^{4-}$ and $\text{HP}_2\text{O}_7^{3-}$, that have occupied HF molecular orbitals with positive orbital energies indicating that inclusion of plane-wave basis functions would lead to spontaneous ionization of these anions. Hence, the gas-phase wavefunctions and energies of these polyanions have little meaning. Such multiply charged anions can exist in the aqueous phase because water or counter-ions act to neutralize the net charge. In this study we assume that the gas-phase anionic structures are accurate, but the electronic wavefunction will be re-optimized in the presence of the continuum solvent. The optimized PCM/*ab initio* wavefunctions have well-defined Fermi levels and positive ionization energies for all occupied orbitals.

In accordance with earlier studies emphasizing the need for diffuse basis functions to accurately model anionic systems we used a large basis set with diffuse functions (6-311++G**) to optimize all protonation states of pyrophosphate and orthophosphate. The larger basis set with diffuse functions has little effect on the structure of the orthophosphate moieties, changing the P–O bond lengths by only ~ 0.005 Å. For the P–O–P bridge; however, the effects are more pronounced. For example in $\text{P}_2\text{O}_7^{4-}$, HF/6-31G* optimization yields a P–O bond length of 1.667 Å and a P–O–P angle of 159° compared to 1.639 Å and 180° from the HF/6-311++G** optimization. Like the neutral

$\text{H}_4\text{P}_2\text{O}_7$, all of the proton-containing anionic pyrophosphates were found to have intramolecular hydrogen bonds linking the phosphate moieties.

The energetics of the gas-phase hydrolysis reactions for the mono- and dianionic species are given in Table 9. As anticipated by earlier theorists, the poly-anionic pyrophosphates have large, negative hydrolysis enthalpies due to the repulsion between the negatively charged phosphate moieties. An exception to this trend is $\text{H}_3\text{P}_2\text{O}_7^{-1}$ which is predicted to have a higher reaction enthalpy than the neutral species. This is a result of a combination of factors: the repulsion between the neutral and anionic ends of $\text{H}_3\text{P}_2\text{O}_7^{-1}$ is not significantly greater than in the neutral species, and the $\text{H}_2\text{PO}_4^{-1}$ product is less able to stabilize a negative charge than the $\text{H}_3\text{P}_2\text{O}_7^{-1}$ reactant.

Aqueous Phase Results

George et al.³⁸ predicted that the extremely large solvation energies for the anionic phosphates would dominate the hydrolysis reaction energies. The components of the solvation energy of each phosphate are given in Table 10. The total aqueous solvation energies calculated using all three solvation models are compared with George's empirical estimates in Table 11. As can be seen in this table there is surprising variation in the predicted solvation energies, even for the neutral and mono-anionic species. The AM1-SM2 method, in particular, predicts up to 90 kcal/mole larger solvation energies than the other methods. Although experimental results must be the ultimate arbiter of these numbers, the AM1-SM2 results of -110. kcal/mole for H_2PO_4^- and -146. kcal/mole for $\text{H}_3\text{P}_2\text{O}_7^-$ are unreasonably large for mono-anionic molecules. These anomalous values

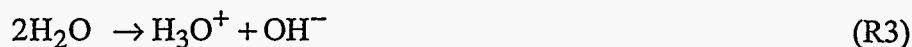
arise from unreasonable large atomic charges predicted for the phosphorus and oxygen atoms. These erroneous atomic charges (and solvation energies) have been seen in a series of second row oxides.⁴⁹

Although there is little experimental data available for the solvation energies for the pyro- and orthophosphates, precise results are available for the aqueous phase acid constants (pK_a) of these compounds. *Ab initio* prediction of pK_a values is a notoriously difficult problem for several reasons. Firstly, practical basis set limitations make it extremely difficult to calculate with equal accuracy the gas-phase energies of the acid and its conjugate base. Secondly, accurate solvation energies are difficult to calculate. Finally, the entire range of typical pK_a values spans a range of reaction energies of about 10 kcal/mole, hence requiring extreme accuracy in all terms. Nevertheless, studies have found that within families of related compounds, systematic errors can be identified and relative pK_a values can be determined within a few pK_a units.⁵⁰

Table 12 lists the *ab initio* and semi-empirical gas-phase proton affinities (PA) for the pyro- and orthophosphates, where the experimental ΔH_f for the proton of 367.2 kcal/mole was used for the AM1 and PM3 PA's.⁵¹ The estimated PA's of George³⁸ are listed in the last column of this table. The predicted pK_a 's are given in Table 13. The PCM/*ab initio* values are calculated from the proton transfer reaction:



where the experimental solvation energy of -259.5 kcal/mole was used for the proton.⁵² The imbalance in molecular charges between the reactants and products (neutral (or anion) + neutral → anion + cation) will lead to less accurate results. For example, the error in the free energy of the gas-phase proton transfer in water:



is 29.1 kcal/mole at this level of theory. (Note that the bulk of this discrepancy arises from unusually large errors in OH⁻.)

The PCM/*ab initio* method predicts the correct ordering of the pK_a values for all the ortho- and pyrophosphates. Quantitatively, however, both the relative and absolute PCM/*ab initio* pK_a's have errors of 0-9 pK_a units, with the largest errors seen for the most highly charged species.

To be consistent with the pK_a predictions of Cramer and Truhlar⁵³ the pK_a's were also calculated using the relation of Pearson⁵²:

$$-\Delta G^\circ_{\text{X}^-} = \text{PA} - 267 + -\Delta G^\circ_{\text{HX}} - 1.36 \text{ pK}_a$$

This expression yields PCM/*ab initio* pK_a's that are generally much better than those calculated directly, particularly for the orthophosphates. The errors in the relative pK_a values for the pyrophosphates from the AM1-SM2 method are actually only about ~5 pK_a units, but these results may be fortuitous since the pK_a ordering is incorrect and the AM1-SM2 orthophosphate results have much larger relative errors. (Note that our

calculated pK_a 's for H_3PO_4 and $H_2PO_4^-$ differ marginally from those reported by Cramer and Truhlar.⁵³⁾

Table 14 lists the predicted aqueous-phase hydrolysis energies using the *ab initio* - polarizable continuum model and the AM1-SM2 and PM3-SM3 methods. As expected, the aqueous solvent greatly mitigates the repulsive interactions in the anionic systems. In fact, the PCM/*ab initio* and AM1-SM2 methods predict positive hydrolysis enthalpies for almost all anionic pyrophosphates, indicating that the solvation model has overestimated the aqueous phase stability of the pyrophosphate. The PM3-SM3 hydrolysis energies are 20-30 kcal/mole lower than calculated with the other two methods, because PM3-SM3 underestimates the solvation energies of the pyrophosphates relative to the orthophosphates.

The experimental hydrolysis free energy for pyrophosphate is -5.1 kcal/mole at pH 7.0.³⁹ At this pH the pyrophosphates will exist primarily in its trianionic and dianionic forms. The predicted hydrolysis free energies using the PCM/*ab initio* method is 6.6 kcal/mole for the prominent trianionic $H_1P_2O_7^{-3}$ and 2.7 kcal/mole for the dianionic $H_2P_2O_7^{-2}$. Hence, our predictions are high by as much as 12 kcal/mole. There are several sources of error in the calculation that could lead to this final error. First, as mentioned above, the gas phase HF optimizations may lead to unreliable structures and energies due to the inherent instability of polyanionic compounds in the gas phase. Secondly, even if the gas-phase anionic structures are reliable, they are likely to change upon aqueous solvation, an effect not yet incorporated into our solvation model. Thirdly, the predicted electrostatic solvation energy is somewhat sensitive to the details of the solvent-

accessible surface. For example, using a smaller probe radius (the radius of the sphere that is “rolled” across the Van der Waal’s surface) of 0.8 Å, rather than 1.4 Å, to generate the solvent accessible surface causes the predicted aqueous phase hydrolysis energies of $\text{H}_4\text{P}_2\text{O}_7$ and $\text{H}_1\text{P}_2\text{O}_7^{-3}$ to increase to -8.7 and 7.8 kcal/mole respectively, while the $\text{H}_2\text{P}_2\text{O}_7^{-2}$ hydrolysis energy decreases very slightly to 2.6 kcal/mole. These results will exhibit a similar sensitivity to the choice of Van der Waal’s radii. Reasonable arguments can be made for a range of Van der Waal’s⁵⁴ and probe radii^{46, 55}, but it is clear that the choice of these radii will have a large effect on the results for highly charged compounds. Finally, the PCM/*ab initio* solvation model does not include dispersion terms, so that it is likely to underestimate the solvation energies.

Of the three methods employed, PM3-SM3 predicts negative free energies for all of the hydrolysis reactions shown in Table 14. Although PM3-SM3 predicts the correct sign of the reaction energy, this is likely to be fortuitous in light of the large errors in the predicted pK_a ’s of the reactants and products. This result must be kept in mind when this method is applied to large biologically important polyphosphates such as adenosine triphosphate (ATP) which are too large to be practical for *ab initio* investigation.

Returning to the original question of whether intra- or intermolecular effects lead to the negative free energy of pyrophosphate hydrolysis, this study demonstrates that a simple answer is elusive. In the gas phase, the hydrolysis energy for the neutral $\text{H}_4\text{P}_2\text{O}_7$ is positive due to intramolecular hydrogen bonds linking the phosphate groups. For the polyanionic species in the gas phase, intramolecular repulsion overwhelms the hydrogen bonding to yield strongly negative reaction energies. In the aqueous phase, this

intramolecular repulsion is mitigated by electrostatic solvent interactions. These results suggest that the most accurate overall description of the hydrolysis is that there is a delicate balance between intramolecular repulsion and the intermolecular solvent interactions.

The balance between the intra- and intermolecular effects is evident in the relation between the charge on the pyrophosphate and the aqueous phase hydrolysis energies. The PCM/*ab initio* results in Table 14 qualitatively agree with the enthalpy results of George *et al.*³⁸ who found that in the aqueous phase, neutral $\text{H}_4\text{P}_2\text{O}_7$ had a hydrolysis enthalpy 4 kcal/mole lower (more negative) than $\text{P}_2\text{O}_7^{-4}$ indicating that the solvent interaction has more than compensated for the very large intramolecular repulsion for the tetra-anionic species. Note that George does not find the same trend for the free energies of hydrolysis due to an unusually high ΔS of +23 e.u. for the $\text{P}_2\text{O}_7^{-4}$ hydrolysis compared with -7 to +5 e.u. for the other protonation states. We predict a much more narrow range of hydrolysis entropies of (+4.2 to +11.1 e.u.) with the $\text{P}_2\text{O}_7^{-4}$ hydrolysis having the lowest ΔS in this range.

The goal of this study was to determine if the negative hydrolysis free energy of pyrophosphate is caused by intra- or intermolecular effects. These calculations find that in the gas-phase the hydrolysis of the fully-protonated pyrophosphate is disfavored by 5 kcal/mole. The origin of this unfavorable free energy is a pair of intramolecular hydrogen bonds that link the two phosphate moieties. These H-bonds have the structural effects of reducing the O-P-O bridging angle by 10-20° relative earlier predicted structures.

As expected, for the anionic forms of pyrophosphate that exist near neutral pH, the gas-phase hydrolysis energies are strongly negative due to electrostatic repulsion. All of the aqueous solvation models used act to cancel this repulsion, leading to hydrolysis energies in the range -25 to +10 kcal/mole. The *ab initio* aqueous phase result which we anticipate to be most reliable, yields hydrolysis energies of 3 to 7 kcal/mole for the protonation states predominant near physiological pH.

Considering the large solvation energies predicted for the anionic phosphates, it is reasonable to take the view that the near-zero free energies found in most of the hydrolysis reactions is purely fortuitous. Indeed the errors in the predicted pK_a 's suggests that there is considerable room for continued refinement in the prediction of aqueous phase reaction energies.

Continuum-based solvation methods yield reasonable solvation phase reaction energies. However, more accurate methods or carefully calibrated empirical corrections are required to achieve predictive accuracy for acid constant calculations. The complexity and subtle energetics of biochemical systems raise many challenges to their theoretical modeling. A careful analysis of the accuracy all of the gas-phase and aqueous-phase energy terms is a necessary first step to predictive simulations of these processes.

Table 9: Pyrophosphate hydrolysis enthalpies (*ab initio* values include ZPE) in the gas phase (kcal/mole).

	$\Delta H(\text{Gas})$ AM1	$\Delta H(\text{Gas})$ PM3	$\Delta H(\text{Gas})$ SCF/ 6-31G**	$\Delta H(\text{Gas})$ MP2/ 6-31G**	$\Delta H(\text{Gas})$ SCF/ 6-311++G**	$\Delta H(\text{Gas})$ MP2/ 6-311++G**
$\text{H}_4\text{P}_2\text{O}_7 + \text{H}_2\text{O} \rightarrow 2 \text{H}_3\text{PO}_4$	3.5	-2.0	0.0	4.4	1.7	5.4
$\text{H}_3\text{P}_2\text{O}_7^- + \text{H}_2\text{O} \rightarrow \text{H}_3\text{PO}_4 + \text{H}_2\text{PO}_4^-$	14.3	6.1	20.1	26.1	19.8	23.9
$\text{H}_2\text{P}_2\text{O}_7^{2-} + \text{H}_2\text{O} \rightarrow 2 \text{H}_2\text{PO}_4^-$	-67.9	-80.2	-53.8	-48.3	-52.9	-49.0

Table 10: Components of the polarizable continuum method/*ab initio* solvation energies. Note that the conversion from standard volume to 1 liter (1.84 kcal/mole = RT ln(22.4/1.0)) has been added to the total solvation energies (columns 4 and 5).

Molecule	Total solution phase free energy (a.u.)	Electrostatic (kcal/mole)	Cavitation (kcal/mole)	Solvation E (kcal/mole)
H ₄ P ₂ O ₇	-1209.8849	-31.2	15.0	-14.3
H ₃ P ₂ O ₇	-1209.4700	-78.0	14.9	-61.3
H ₂ P ₂ O ₇ ²	-1209.0436	-211.8	14.2	-195.8
H ₁ P ₂ O ₇ ³	-1208.6085	-458.2	14.5	-441.9
P ₂ O ₇ ⁴	-1208.1557	-791.9	15.0	-775.1
H ₃ PO ₄	-643.0965	-28.2	10.5	-15.9
H ₂ PO ₄	-642.6664	-85.0	10.1	-73.0
H ₁ PO ₄ ²	-642.2242	-263.0	9.9	-251.2
PO ₄ ³	-641.7708	-556.9	9.7	-545.4
H ₂ O	-76.2918	-12.1	5.0	-5.3

Table 11: Comparison of total solvation energies calculated by three different methods. (Energies in kcal/mole)

Molecule	PCM/ <i>ab initio</i>	AM1-SM2	PM3-SM3	Empirical Estimates ³⁸
H ₄ P ₂ O ₇	-14.3	-104.1	-59.7	NA
H ₃ P ₂ O ₇	-61.3	-146.3	-96.6	-84.
H ₂ P ₂ O ₇ ²	-195.8	-282.2	-230.2	-134.
H ₁ P ₂ O ₇ ³	-441.9	-540.7	-471.2	-358.
P ₂ O ₇ ⁴	-775.1	-891.0	-806.3	-584.
H ₃ PO ₄	-15.9	-52.5	-34.5	NA
H ₂ PO ₄	-73.0	-110.3	-86.3	-76.
H ₁ PO ₄ ²	-251.2	-299.9	-266.2	-299.
PO ₄ ³	-545.4	-629.2	-570.9	-637.
H ₂ O	-5.3	-6.3	-6.3	-6.3 (Expt.)

Table 12: Gas phase proton affinities (kcal/mole).

Reaction	MP2/ 6-311++G**	AM1	PM3	Estimated Gas-phase PA ³⁸
$\text{H}_4\text{P}_2\text{O}_7 \rightarrow \text{H}_3\text{P}_2\text{O}_7 + \text{H}^+$	307.3	328.0	338.2	311.
$\text{H}_3\text{P}_2\text{O}_7 \rightarrow \text{H}_2\text{P}_2\text{O}_7^2 + \text{H}^+$	398.7	421.0	432.5	488.
$\text{H}_3\text{PO}_4 \rightarrow \text{H}_2\text{PO}_4 + \text{H}^+$	325.8	338.8	346.3	311.
$\text{H}_2\text{PO}_4 \rightarrow \text{H}_1\text{PO}_4^2 + \text{H}^+$	455.6	477.7	485.8	488.

Table 13: Aqueous phase acid constants (pK_a units).

Reaction	PCM <i>lab</i> <i>initio</i>	PCM <i>lab</i> <i>initio</i> (Pearson)	AM1- SM2	PM3- SM3	Expt. ⁵⁶
$\text{H}_4\text{P}_2\text{O}_7 \rightarrow \text{H}_3\text{P}_2\text{O}_7 + \text{H}^+$	0.7	-4.9	13.8	25.3	0.8
$\text{H}_3\text{P}_2\text{O}_7 \rightarrow \text{H}_2\text{P}_2\text{O}_7^2 + \text{H}^+$	5.9	-2.0	13.3	23.5	2.2
$\text{H}_2\text{P}_2\text{O}_7^2 \rightarrow \text{H}_1\text{P}_2\text{O}_7^3 + \text{H}^+$	10.0	5.5	10.6	32.4	6.7
$\text{H}_1\text{P}_2\text{O}_7^3 \rightarrow \text{P}_2\text{O}_7^4 + \text{H}^+$	18.1	13.9	13.5	31.7	9.4
$\text{H}_3\text{PO}_4 \rightarrow \text{H}_2\text{PO}_4 + \text{H}^+$	7.7	1.3	10.3	20.2	2.1
$\text{H}_2\text{PO}_4 \rightarrow \text{H}_1\text{PO}_4^2 + \text{H}^+$	13.2	7.7	15.6	28.6	7.2
$\text{H}_1\text{PO}_4^2 \rightarrow \text{PO}_4^3 + \text{H}^+$	18.4	11.8	14.7	34.8	12.4

Table 14: Pyrophosphate hydrolysis energies in the aqueous phase (kcal/mole).

Reaction	$\Delta\text{G}(\text{Aq})$ MP2/ 6-311++G**	$\Delta\text{G}(\text{Aq})$ AM1-SM2	$\Delta\text{G}(\text{Aq})$ PM3-SM3
$\text{H}_4\text{P}_2\text{O}_7 + \text{H}_2\text{O} \rightarrow 2 \text{H}_3\text{PO}_4$	-9.3	9.6	-3.4
$\text{H}_3\text{P}_2\text{O}_7 + \text{H}_2\text{O} \rightarrow \text{H}_3\text{PO}_4$ $+ \text{H}_2\text{PO}_4$	0.1	4.1	-11.0
$\text{H}_2\text{P}_2\text{O}_7^2 + \text{H}_2\text{O} \rightarrow 2$ H_2PO_4	2.7	0.7	-12.9
$\text{H}_1\text{P}_2\text{O}_7^3 + \text{H}_2\text{O} \rightarrow$ $\text{H}_2\text{PO}_4 + \text{H}_1\text{PO}_4^2$	6.6	4.0	-21.3
$\text{P}_2\text{O}_7^4 + \text{H}_2\text{O} \rightarrow 2 \text{H}_1\text{PO}_4$ 2	-0.8	8.2	-24.6

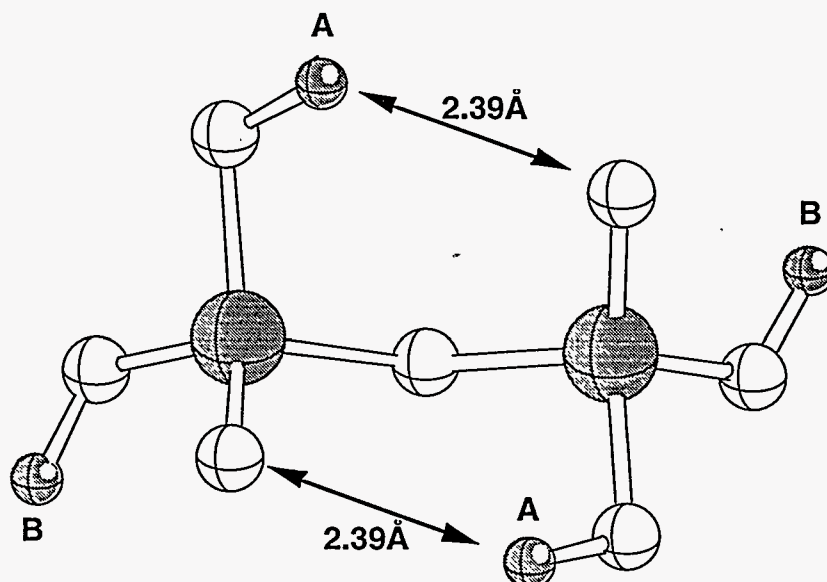


Figure 6) H₄P₂O₇ Hartree-Fock/6-311++G** global minimum. The C₂ symmetry structure has two symmetry distinct hydrogens labeled "A" and "B". Those labeled "A" are hydrogen bonded to the opposite phosphates.

Conclusions:

This report has presented the results of applying quantum chemical methods employing the methods of density functional theory and a grid-based boundary-element solvent model to a series of biomedically relevant chemical studies. We demonstrated that local density approximation based, plane-wave basis DFT can accurately predict the structural properties of crystalline, platinum compounds, such as the anticancer drug cisplatin. Additionally, we showed that a grid-based DFT method using a gradient-corrected exchange-correlation functional (BLYP) can predict structural properties of biochemicals such as glycine and guanine with an accuracy close to that of second-order Møller-Plesset Perturbation Theory, for far less computational cost. Using the grid-based boundary-element solvation, coupled to an *ab initio* quantum chemical description of the solute molecules, we found very good agreement for the solution-phase energies of the *syn*- and *anti*-configurations of the carboxylate moiety of acetic acid with Monte Carlo simulations including explicit waters. We also showed that this solvation model yielded

reasonably accurate acid constants and hydrolysis energies for the important biochemical group pyrophosphate.

Acknowledgments: This work was carried out in part at the Sandia National Laboratory under contract from the U.S. Department of Energy and supported by its Division of Basic Energy Sciences and by the Sandia Laboratory Directed Research and Development Funding.

References:

1. Reed, E. & Kohm, K.W. in *Cancer Chemotherapy* (eds. Chabner, B.A. & Collins, J.M.) 465-490 (J. B. Lippincott Co., Philadelphia, 1990).
2. Sherman, S.E. & Lippard, S.J. *Chemical Reviews* **87**, 1153-1181 (1987).
3. Sherman, S.E., Gibson, D., Wang, A.H.-J. & Lippard, S.J. *Journal of the American Chemical Society* **110**, 7368-7381 (1988).
4. Yohannes, P.G., Zon, G., Doetsch, P.W. & Marzilli, L.G. *Journal of the American Chemical Society* **115**, 5105-5110 (1993).
5. McCarthy, S.L., *et al.* *Biopolymers* **29**, 823-826 (1990).
6. Kozelka, J., Petsko, G.A. & Lippard, S.J. *Journal of the American Chemical Society* **107**, 4079-4081 (1985).
7. Krogh-Jespersen, M.-B. & Altonen, A. *Inorganic Chemistry* **26**, 2084-2090 (1987).
8. Krauss, M., Basch, H. & Miller, K.J. *Journal of the American Chemical Society* **110**, 4517-4521 (1988).
9. Kozelka, J., Savinelli, R., Berthier, G., Flament, J.-P. & Lavery, R. *Journal of Computational Chemistry* **14**, 45-53 (1993).
10. Gropen, O., Almlöf, J. & Wahlgren, U. *Journal of Chemical Physics* **96**, 8363-8366 (1992).
11. Parr, R.G. & Yang, W. *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

12. Ziegler, T., Tschinke, V. & Becke, A. *Journal of the American Chemical Society* **109**, 1351-1358 (1987).
13. Delley, B., Freeman, A.J. & Ellis, D.E. *Physical Review Letters* **50**, 488-491 (1983).
14. Bernholc, J. & Holzwarth, N.A.W. *Physical Review Letters* **50**, 1451-1454 (1983).
15. Hohenberg, P. & Kohn, W. *Physical Review* **136**, B864 (1964).
16. Kohn, W. & Sham, L.J. *Physical Review* (1965).
17. Bachelet, G.B., Hamann, D.R. & Schluter, M. *Physical Review* **B24**, 4190 (1982).
18. Hamann, D.R., Schluter, M. & Chiang, C. *Physical Review Letters* **43** (1979).
19. Hamann, D.R. *Physical Review* **B40**, 2980 (1989).
20. Ceperley, D.M. & Alder, B.J. *Physical Review Letters* **45**, 566-569 (1980).
21. Teter, M.P., Payne, M.C. & Allan, D.C. *Physical Review* **B40**, 12255 (1989).
22. Nelson, J.S., Plimpton, S.J. & Sears, M.P. *Physical Review* **B47**, 1765 (1993).
23. Milburn, G.H.W. & Truter, M.R. *Journal of the Chemical Society* , 1609-1616 (1966).
24. Vinogradov, S.N. & Linnell, R.H. *Hydrogen Bonding* (Van Nostrand Reinhold, New York, 1971).
25. Pauling, L. *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, 1960).
26. Becke, A.D. *Physical Review A* **38**, 3098-3100 (1988).
27. Lee, C., Yang, W. & Parr, R.G. *Physical Review B* **37**, 785-789 (1988).
28. Allen, F.H. & Kirby, A.J. *Journal of the American Chemical Society* **113**, 8829-8831 (1991).
29. Wiberg, K.B. & Laidig, K.E. *Journal of the American Chemical Society* **109**, 5935-5943 (1987).
30. Gao, J. *Journal of Physical Chemistry* **96**, 6432-6439 (1992).
31. Pranata, J. *Journal of Computational Chemistry* **14**, 685-690 (1993).

32. Gao, J. & Pavelities, J.J. *Journal of the American Chemical Society* **114**, 1912-1914 (1992).
33. Briggs, J.M., Nguyen, T.B. & Jorgensen, W.L. *Journal of Physical Chemistry* **95**, 3315-3322 (1991).
34. Frisch, M.J., *et al.* (Gaussian Inc., Pittsburg, 1992).
35. Frisch, M.J., *et al.* (Gaussian Inc., Pittsburg, PA, 1994).
36. Colvin, M.E. & Melius, C.F. (Sandia National Laboratories, 1993).
37. Cramer, C.J. & Truhlar, D.G. in *Reviews in Computational Chemistry* (eds. Lipkowitz, K.B. & Boyd, D.B.) 1-72 (VCH Publishers, New York, 1995).
38. George, P., *et al.* *Biochimica et Biophysica Acta*. **223**, 1-15 (1970).
39. Romero, P.J. & de Meis, L. *Journal of Biological Chemistry* **264**, 7869-7873 (1989).
40. Hayes, D.M., Kenyon, G.I. & Kollman, P.A. *Journal of the American Chemical Society* **100**, 4331-4340 (1978).
41. O'Keefe, M., Domenges, B. & Gibbs, G.V. *Journal of Physical Chemistry* **89**, 2304-2309 (1985).
42. Ewig, C.S. & Van Waser, J.R. *Journal of the American Chemical Society* **110**, 79-86 (1988).
43. Saint-Martin, H., Otega-Blake, I., Lez, A. & Adamowitz, L. *Biochimica et Biophysica Acta*. **1080**, 205-214 (1991).
44. Hout Jr., R.F., Levi, B.A. & Hehre, W.J. *Journal of Computational Chemistry* **3**, 234-250 (1982).
45. Miertus, S., Scrocco, E. & Tomasi, J. *Chemical Physics* **55**, 117-129 (1981).
46. Rashin, A.A. & Namboodiri, K. *Journal of Physical Chemistry* **91**, 6003-6012 (1987).
47. Zauhar, R.J. & Morgan, R.S. *Journal of Molecular Biology* **186**, 815-820 (1985).

48. Halicioglu, T. & Sinanolgu, O. *Annals New York Academy of Sciences* **158**, 308-317 (1969).
49. Evleth, E., Akacem, Y. & Colvin, M.E. *Chemical Physics Letters* **227**, 412-418 (1994).
50. Rashin, A.A., Rabinowitz, J.R. & Banfelder, J.R. *J. Am. Chem. Soc.* **112**, 4133-4137 (1990).
51. Stull, D.R. & Prophet, H. *JANAF Thermochemical Tables* (NSRDS-NBS 37, 1971).
52. Pearson, R.G. *Journal of the American Chemical Society* **108**, 6109-6114 (1986).
53. Cramer, C.J. & Truhlar, D.G. *Journal of Computer-Aided Molecular Design* **6**, 629-666 (1992).
54. Marcus, Y. *Chemical Reviews* **88**, 1475 (1988).
55. Ooi, T., Oobatake, M., Némethy, G. & Scheraga, H.A. *Proc. Nat. Acad. Sci.* **85**, 3086-3090 (1987).
56. Martell, A.E. & Smith, R.M. *Critical Stability Constants* (Plenum, New York, 1982).

UNLIMITED RELEASE

INITIAL DISTRIBUTION:

MS 9001 Tom Hunter, Attn:
2200 J. B. Wright
5200 J. F. Ney (Acting)
8100 M. E. John
8200 L. A. West
8300 W. J. McLean
8400 R. C. Wayne
8500 P. N. Smith
8700 T. M. Dyer
8800 P. E. Brewer
8900 D. L. Crawford

MS 0188 C. E. Meyers, 4523
MS 9214 L. M. Napolitano, Jr., 8117
MS 9214 M. E. Colvin, 8117 (10)

MS 9021 Technical Communications Department, 8815 for OSTI (10)
MS 9021 Technical Communications Department, 8815/Technical Library, MS 0899,
4414

MS 0899 Technical Library, 4414 (4)
MS 9018 Central Technical Files, 8940-2 (3)