JCTC Journal of Chemical Theory and Computation

Benchmarking of DFT Functionals for the Hydrolysis of Phosphodiester Bonds

António J. M. Ribeiro, Maria J. Ramos, and Pedro A. Fernandes*

Departamento de Química, Faculdade de Ciências do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

Received December 3, 2009

Abstract: Phosphodiester bonds are an important chemical component of biological systems, and their hydrolysis and formation reactions are involved in major steps throughout metabolic pathways of all organisms. In this work, we applied dimethylphosphate as a model for this kind of bonds and calculated the potential energy surface for its hydrolysis at the approximated CCSD(T)/CBS//B3LYP/6-311++G(2d,2p) level. By varying the nucleophile (water or hydroxide) and the medium (vacuum or aqueous implicit solvent) we obtained and described four reaction paths. These structures were then used in a DFT functional benchmarking in which we tested a total of 52 functionals. Furthermore, the performances of HF, MP2, MP3, MP4, and CCSD were also evaluated. This benchmarking showed that MPWB1K, MPW1B95, and PBE1PBE are the more accurate functionals to calculate the energies of dimethylphosphate hydrolysis as far as activation and reaction energies are concerned. If considering only the activation energies, MPWB1K, MPW1B95, and B1B95 give the lowest errors when comparing to CCSD(T). A basis set benchmarking on the same system shows that 6-311+G(2d,2p) is the best basis set concerning the relationship between computational time and accuracy. We believe that our results will be of great help to further studies on related phosphodiester systems. This includes not only pure chemical problems but also biochemical studies in which DNA, RNA, and phospholipids are required to be depicted at a quantum level.

1. Introduction

Phosphodiester bonds are central in biological systems not only because they are an essential part of nucleic acids and phospholipids but also due to their interaction with many enzymes. Their function exceeds simple bonding, since the negatively charged phosphate group adds unique properties to any molecule in which it is present. In nucleic acids, these bonds sustain the polymer covalent backbone by connecting (deoxy)riboses of adjacent nucleotides. Additionally, phosphodiester bonds make these molecules highly charged, a crucial factor in their folding and interaction with proteins, metal ions, and polyamines. Furthermore, the negative charge protects these bonds from nucleophilic attack and thus lowers the phosphodiester rate of hydrolysis: a desirable property when it comes to genetic information.[1,2] In phospholipids, the phosphodiester bond is the bridge between the hydrophilic and hydrophobic portions of the molecule. It is also an important hydrophilic factor and therefore half-responsible for phospholipids chemical behavior.[1] It is then clear the need for understanding the phosphodiester bond and its key properties, namely its geometry and hydrolysis/formation energies.

In this study, we were particularly interested in analyzing how DFT functionals describe the potential energy surface for phosphodiester hydrolysis, using approximated CCSD(T)/CBS energy values as reference (see Theoretical Methods for details). Our objective was to answer the following question: *If we want to study the kinetics and thermodynamics of the breakage/formation of a phosphodiester bond, which functional should we use*. This question is relevant to all sorts of mechanistic studies with biological systems, in which the use of accurate post-Hartree−Fock methods is not yet feasible. That includes a great number of enzyme families such as cyclic nucleotide phosphodiesterases, DNA/RNA

* Corresponding author e-mail: pafernan@fc.up.pt.

10.1021/ct900649e
Published on Web 07/08/2010

Table 1. DFT Functionals and Basis Sets Tested in This Work[a]

functionals			basis sets
LDA	**H-GGA**	**HM-GGA**	**Pople type**
SVWN[42,43]	B1LYP[44,45]	B1B95[45,46]	6-31G(d)
SVWN5[42,43]	B3LYP[44,47]	BB1K[44,46,48]	6-31+G(d)
GGA	B3P86[47,49]	M06[40,41]	6-311G(d)
BLYP[44,46]	B3PW91[47,50]	M06 - 2X[41,40]	6-31+G(d,p)
BP86[46,49]	B97−1[51]	M06 - HF[41,52]	6-311G(d,p)
BPBE[46,53]	B97−2[54]	MPW1B95[45,48,50,55]	6-311+G(d)
BPW91[46,50]	B98[56,57]	MPW1KCIS[50,55,58,59]	6-311++G(d)
G96LYP[44,60]	BhandH[44,46]	MPWB1K[45,48,50,55]	6-311+G(d,p)
HCTH93[51]	BhandHLYP[44,46]	MPWKCIS1K[50,55,58,59]	6-311++G(d,p)
HCTH147[51]	MPW1K[50,55,61]	PBE1KCIS[53,58,62]	6-311G(2d,2p)
HCTH407[51]	MPW1N[50,55,63]	TPSS1KCIS[58,62,64]	6-311+G(2d,2p)
MPWLYP[44,50,55]	MPW1PW91[50,55]	TPSSh[64]	6311++G(2d,2p)
MPWLYP1W[44,50,55,65]	MPW1S[50,55,66]	**M-GGA**	6311++G(3df,3pd)
MPWPW91[50,55]	MPW3LYP[44,50,55]	BB95[45,46]	**correlation-consistent**
OLYP[44,67]	O3LYP[44,67,68]	M06 - L[41,69]	aug-cc-pvdz
PBE1W[53,65]	PBE1PBE[53]	MPWB95[45,50,55]	aug-cc-pvtz
PBELYP1W[44,53,65]		TPSSLYP1W[44,64,65]	aug-cc-pvqz
PBEPBE[53]		TPSSTPSS[64]	
XLYP[44,46,70,71]		VSXC[72]	

[a] Functionals are organized by type: LDA - local density approximation; GGA - generalized gradient approximation; M-GGA - meta generalized gradient approximation; H-GGA - hybrid generalized gradient approximation; HM-GGA - hybrid meta generalized gradient approximation.

polymerases, nucleases, integrases, and phospholipases C/D, among others. We chose dimethylphosphate, the smallest molecule with this kind of bond, to perform our benchmarking calculations. We have tried to keep the system to a minimum to make it as general as possible and to serve as a model for all enzyme classes. Such a small system captures the essential of the chemical reaction, without the influence of a particular enzymatic scaffold. Both water and hydroxide were used as nucleophiles, and vacuum as well as aqueous solvent were employed, leading to four different systems. The effect of different basis sets was also studied, in a basis set benchmarking.

Numerous theoretical works concerning dimethylphosphate hydrolysis and some similar systems have been published in the past.[3−23] These studies describe the stationary states for the hydrolysis reaction path in a very detailed way, and at various levels of theory, namely at the density functional theory (always using B3LYP), second order Moller−Plesset (MP2) and Hartree−Fock (HF), combined with a whole range of split-valence basis sets. As B3LYP is widely used and produces generally good results[24] it is usually preassumed as a good choice for all types of calculations, and hence no effort has been made to justify its use as an alternative to others functionals. We note, however, that neither B3LYP has been properly validated for this so important chemical reaction nor have any others functionals been systematically tested. We therefore present this benchmarking to complement the lack of such information on phosphodiester systems.

2. Theoretical Methods

In order to obtain the structures for the benchmarking studies we started by modeling the molecules of the reactants and products involved in the two main reactions: hydroxide, water, dimethylphosphate, methoxide, methanol, and methyl phosphate. Structures of transition states and intermediates were then created by scanning the proper coordinates of the reactants systems. All these structures were fully optimized in vacuum and solvent at the B3LYP/6-311++G(2d,2p) level. Aqueous solvent systems were treated with the IEFPCM continuum model.[25−28] Frequency calculations were carried out for all structures to confirm their nature (minima/transition state) along the potential energy surface. Zero point energies and entropic corrections were also computed.

Single point energy values at the approximated CCSD(T)/CBS//B3LYP/6-311++G(2d,2p) level were employed as reference to evaluate the accuracy of the functionals. The approximated CCSD(T)/CBS energy was obtained in two main steps. First, the MP2/CBS energy was extrapolated by using the Helgaker scheme: the Hartree−Fock energy was extrapolated by fitting the energy of aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis sets to an exponential function and the energy of MP2 by fitting the values of aug-cc-pVTZ and aug-cc-pVQZ basis sets to a polynomial function.[29,30] Second, we summed the correlation energy difference between MP2 and CCSD(T), calculated with the 6-31+G(d) basis set.[31,32] Using the same methodology, we also calculated single point energies at the approximated CCSD/CBS level. We used these values to better understand the importance of triple excitations in CCSD(T), although it is not described in the literature if this kind of approximation can be applied precisely to CCSD.

Basis set benchmarking was used to assess the quality of the basis set indicated in Table 1. Single point energies with all those basis sets were calculated for all our systems using the same functional, B3LYP. MUE (mean unsigned error) and MSE (mean signed error) for these series were based on the comparison with the most complete basis set we have used, aug-cc-pvqz.

We tested the functionals shown in Table 1 by performing single point energy calculations for each one of the geometries mentioned above with the proper functionals and the 6-311++G(2d,2p) basis set. The same scheme

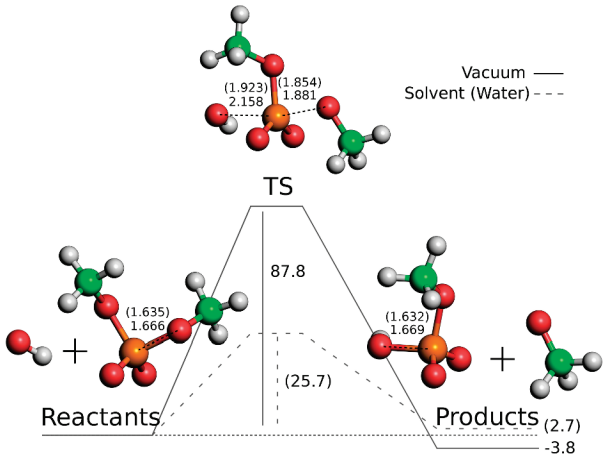

Figure 1. Reaction path for dimethylphosphate hydrolysis with hydroxide attack. Distances for formed and broken bonds are indicated in Å. Energies are in kcal/mol. Numbers inside brackets refer to solvent calculations. Structures shown are for vacuum systems. Solvent structures do not differ significantly from these.

was employed for Hartree−Fock and post-Hartree−Fock methods: MP2, MP3, MP4(SDQ), and CCSD. We have chosen the 6-311++G(2d,2p) basis set because it was the most complete basis set that could be used within most modern hardware capabilities. Calculations with more complete basis sets would become intolerably time-consuming. In addition, according to the basis set analysis we have made, this basis set gives rather reliable results for our system, with a basis set truncation error below 1 kcal/mol for both activation and reaction energies. Single point calculations with M06 functionals were carried out with JAGUAR.[33] As IEFPCM is not included in JAGUAR; all our M06 functionals calculations were performed in vacuum. All other calculations, including scans and optimizations, were done using GAUSSIAN03.[34]

3. Results and Discussion

3.1. Obtained Geometries. Structures obtained at the B3LYP/6-311++G(2d,2p) level are represented in Figures 1 and 2. Tables 2, 3, and 4 show relevant interatomic distances, angles, dihedrals, and Mulliken atomic charges associated with the four studied chemical pathways: dimethylphosphate hydrolysis with hydroxide or water as nucleophiles, both in the gas phase and implicit solvent. Phosphodiester hydrolysis via hydroxide attack is a one step mechanism. Barrier heights for this reaction, calculated at the approximated CCSD(T)/CBS//B3LYP/6-311++G(2d,2p) level, are 87.8 kcal/mol in vacuum and 25.7 kcal/mol in aqueous solvent. The products of the reaction are methylphosphate and methoxide.

The reaction with a water molecule as nucleophile, represented in Figure 2, is more complex. A total of three transition states and two intermediates are necessary. In the first transition state, water attacks the dimethylphosphate molecule and simultaneously loses a proton to it. In the second transition state, this hydrogen atom rotates ca. 180° to form an H-bond with the leaving group. In the last step, a methoxide ion dissociates from the pentacoordinated phosphorus center and simultaneously deprotonates one of the phosphate oxygen atoms, leaving as a neutral methanol molecule. Note that, unlike with the hydroxide mechanism, products and reagents here have favorable interaction energy. Transition states energies, calculated with the same level of theory as above are, respectively, 22.5, 29.0, and 25.4 kcal/mol in vacuum and 36.3, 36.5, and 37.62 kcal/mol in solvent.

When comparing the kinetics of the two nucleophilic reactions in solution, we need to be aware of the difference in the concentration of the two nucleophiles. The entropic cost of confining a molecule to a "solvent cage" in the prereactive state can be given by eq 1 below.[35] This term only affects the hydroxide whose concentration is very low (10^{-7} M) but not the water nucleophile, which is always present at the reactive center.

$$\Delta G^{\circ}_{cage} = -kb \cdot T \cdot \ln\left(\frac{Vf}{Vi}\right) \quad (1)$$

When taking this correction into account, and considering a final restriction volume of six water molecules, the activation energy associated with the hydroxide reaction raises 10.9 kcal/mol. Therefore, the activation energies for hydroxide attack and water attack involved, in solution, are very similar in the two cases. It does not make sense to consider this effect in the system with the water nucleophile since the dimethylphosphate is surrounded by water at all times and no energy cost is needed for a particular water molecule to be in position for attack.

Barrier heights reported in other studies are similar to ours.[8,19,21,22] Not surprisingly, geometries are also equivalent to those of previous studies with smaller basis sets.[7−9,12,17,19,22] It is known that geometric parameters converge very quickly with increasing basis sets and/or methods. Our protocol takes advantage of this, since for the remainder calculations we used these structures without further optimization. Performing optimizations for all functionals and basis sets would be a great computational effort and would bring negligible gains in terms of the calculated hydrolysis energies. A remark needs to be done, however, concerning the hydroxide attack in vacuum: for very small basis sets, STO-3G, 3-21G(d), and 3-21+G(d), this reaction path presents an intermediate and an additional transition state.[4,8,36] However, the depth of the intermediate well diminishes and finally disappears with basis sets of increasing size. Such an intermediate was not a stationary state at the B3LYP/6-311++G(2d,2p) level.

Tables 5, 6, and 7 show the energy values for the four reactions paths calculated with various methods as well as the Gibbs and zero point energy corrections. We will discuss relevant experimental data concerning the dimethyl phosphate hydrolysis, but we will not attempt at evaluating the quality of the various methods by direct comparison with it, as will become clear in the section below.

The B3LYP energy values are not in good agreement with the reference values, and so it is not an adequate functional to study the energetics of phosphodiester systems. Other functionals results will be discussed in the proper section. Pure Hartree−Fock is very unreliable, as the errors in barrier energies are, in most cases, higher than 10 kcal/mol. Regarding post-Hartree−Fock methods, it is very interesting

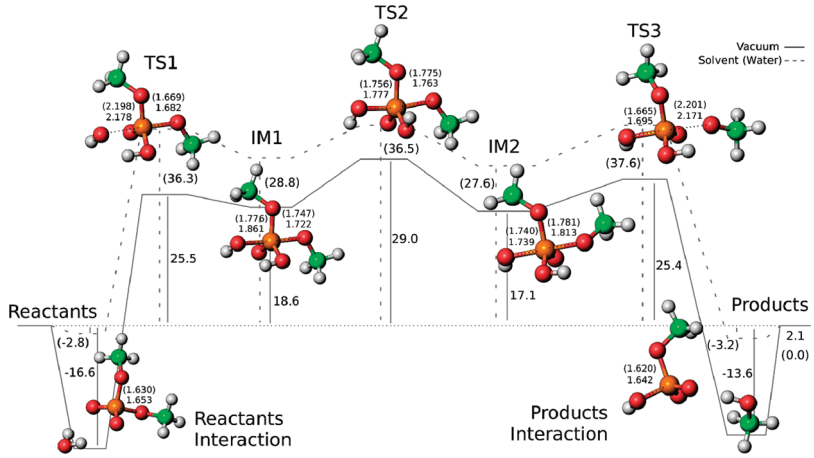

Figure 2. Reaction path for dimethylphosphate hydrolysis with a water molecule attack. Distances for formed and broken bonds are indicated in Å. Energies are in kcal/mol. Numbers inside brackets refer to solvent calculations. Structures shown are for vacuum systems. Solvent structures do not differ significantly from these.

Table 2. Relevant Distances, Angles, Dihedrals, and Atomic Charges for Dimethylphosphate Hydrolysis by Hydroxide Attack in Vacuum and Implicit Solvent[a]

		hydroxide attack					
		vacuum			implicit solvent		
		reactants	TS	products	reactants	TS	products
distances	On−P	---	2.158	1.669	----	1.923	1.923
	P−O1	1.666	1.881	----	1.635	1.854	----
angles	On−P−O1	----	167.6	----	----	170.1	----
dihedrals	O2−P−O1−C1	−73.4	−178.0	----	−66.5	−176.7	----
	C2−O2−P−O4	36.7	−32.5	−38.9	45.2	−20.3	−47.4
Mulliken charges	q1	−1.00	−0.70	−0.45	−1.00	−0.52	−0.41
	q2	−0.51	−0.88	−0.55	−0.55	−0.90	−0.59
	q3	−0.49	−0.41	−1.00	−0.45	−0.48	−1.00

[a] See Figure 3 for atom and charges denomination. Distances are indicated in Å; angles and dihedrals are in degrees. Geometrical parameters were optimized at B3LYP/6-311++G(2d,2p). Mulliken charges were calculated at the CCSD/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p) level.

Table 3. Important Distances, Angles, Dihedrals, and Charges in Dimethylphosphate Hydrolysis by Hydroxide Attack in Vacuum[a]

		water attack - vacuum								
		R.	R. I.	TS1	IM1	TS2	IM2	TS3	P. I.	P.
distances	On−P	----	3.359	2.198	1.776	1.765	1.740	1.665	1.620	1.632
	P−O1	1.635	1.630	1.669	1.747	1.775	1.781	2.201	3.931	----
	On−Hn2	0.971	0.971	1.709	2.040	2.761	3.209	3.208	3.987	----
	O4−Hn2	----	2.063	0.992	0.973	0.977	0.974	0.999	1.751	----
	O1−Hn2	----	4.046	3.248	3.252	2.751	1.975	1.648	0.981	0.973
angles	O1−P−On	----	128.5	165.7	168.1	166.4	166.6	163.9	118.1	----
dihedrals	O2−P−O1−C1	−66.5	−66.0	−179.7	176.4	179.1	68.2	134.3	−155.2	----
	C2−O2−P−O4	177.6	177.6	130.0	106.8	100.6	117.3	117.8	−72.3	−47.3
	On−P−O4−Hn2	----	0.0	−6.1	−10.7	84.3	174.6	173.2	−124.5	----
Mulliken charges	q1	0.00	−0.01	−0.73	−0.42	−0.32	−0.34	−0.29	−0.36	−0.41
	q2	−0.55	−0.59	−0.05	−0.25	−0.29	−0.27	0.02	0.64	−0.59
	q3	−0.45	−0.41	−0.22	−0.33	−0.39	−0.39	−0.72	−0.01	0.00

[a] See Figure 3 for atom and charges denomination. Distances are indicated in Å; angles and dihedrals are in degrees. Geometrical parameters were optimized at B3LYP/6-311++G(2d,2p). Mulliken charges were calculated at the CCSD/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p) level. R. - reactants, R. I. - reactants interaction, TS - transition state, IM -intermediate, P. I. - products interaction, P. - products.

to observe the difference between approximated CCSD/CBS and approximated CCSD(T)/CBS results, as an indication that the inclusion of triple excitations is crucial for a good description of our system. It is also important to note that the CCSD/6-311++G(2d,2p) energy is very different from the approximated CCSD/CBS. This means that CCSD energies converge slowly with the improvement of the basis sets, and even a triple-ζ basis set is still far from the complete basis set limit. The same could be said for all post-Hartree−Fock methods.

The thermal and entropic contributions for the Gibbs energy present large values (around 11 kcal/mol) when considering complexes as opposed to structures where the molecular species are set apart. This is due to a major loss

Table 4. Relevant Distances, Angles, Dihedrals, and Atomic Charges for Dimethylphosphate Hydrolysis by Hydroxide Attack in Implicit Solvent[a]

		water attack - implicit solvent								
		R.	R. I.	TS1	IM1	TS2	IM2	TS3	P. I.	P.
distances	On−P	----	3.360	2.178	1.861	1.777	1.739	1.695	1.642	1.669
	P−O1	1.666	1.653	1.682	1.722	1.763	1.813	2.171	3.772	----
	On−Hn2	0.961	0.971	1.668	1.921	2.720	3.213	3.247	4.111	----
	O4−Hn2	----	2.099	0.995	0.966	0.960	0.965	0.999	1.740	----
	O1−Hn2	----	4.049	3.238	3.221	2.751	1.883	1.632	0.985	0.960
angles	O1−P−On	----	129.8	165.5	167.8	166.4	166.9	167.4	121.1	----
dihedrals	O2−P−O1−C1	−73.4	−73.4	176.3	173.6	178.2	74.1	128.5	134.4	----
	C2−O2−P−O4	173.6	172.6	123.1	121.6	113.2	117.7	157.2	−64.2	−38.9
	On−P−O4−Hn2	----	0.3	−5.1	−7.4	82.1	175.2	176.7	−138.4	----
Mulliken charges	q1	0	−0.01	−0.67	−0.50	−0.36	−0.35	−0.30	−0.38	−0.45
	q2	−0.51	−0.64	−0.09	−0.24	−0.31	−0.24	−0.07	−0.61	−0.55
	q3	−0.49	−0.45	−0.24	−0.26	−0.33	−0.41	0.63	−0.02	0.00

[a] See Figure 3 for atom and charges denomination. Distances are indicated in Å; angles and dihedrals are in degrees. Geometrical parameters were optimized at B3LYP/6-311++G(2d,2p). Mulliken charges were calculated at the CCSD/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p) level. R. - reactants, R. I. - reactants interaction, TS - transition state, IM - intermediate, P. I. - products interaction, P. - products.

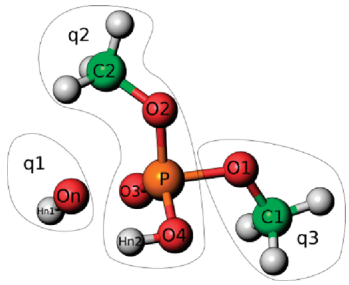

Figure 3. Division of the system in three groups, q1, q2, and q3, to analyze the changes in groups atomic charges along the reaction pathway. Hydroxide attacking systems do not have 'Hn2'. For water attacking systems, in reactants and products 'Hn2' belongs to q1 and q3, respectively. In the others structures belong to q2.

Table 5. Activation and Reaction Energies for Dimethylphosphate Hydrolysis by Hydroxide Attack in Vacuum and Implicit Solvent As Calculated by Various Methods[a]

	vacuum		implicit solvent	
	TS	products	TS	products
DFT - B3LYP	88.1	−8.6	30.1	0.4
Hartree−Fock	101.1	−8.2	41.3	0.2
MP2	86.7	−3.8	27.4	3.5
MP2/CBS	88.7	−2.9	25.0	4.2
MP3	89.1	−6.0	28.8	1.9
MP4(SDQ)	88.8	−5.0	29.2	2.6
CCSD	89.0	−5.3	29.2	2.4
CCSD/CBS	90.6	−4.0	28.1	2.3
CCSD(T)/CBS	87.8	−3.8	25.7	2.7
correction to the Gibbs free energy	9.4	−2.6	−2.0	0.8
zero point energy	0.4	−1.0	1.0	−0.5
solvent cage energy	----	----	10.9	----
total free energy	97.2	−6.4	34.6	3.5

[a] The zero point energy and the correction to the Gibbs free energy are also included. All values are in kcal/mol. In the implicit solvent case, the correction to the Gibbs free energy does not contain the translational and rotational entropic energies. Total free energy refers to the sum of CSSD(T)/CBS energy with the correction to the Gibbs free energy and the solvent cage energy, if applied. The methods not noted with CBS refer to calculations with the 6-311++G(2d,2p) basis set. TS - transition state.

of entropy during the formation of the complex. A detailed thermodynamic analysis of all molecules involved shows that the major contributions come from the translational and rotational entropy associated with the small nucleophiles, which are nearly completely lost in that step. We should not include this correction on the energy of solvated systems because the complex water−dimethylphosphate should always exist in solution and the nucleophile that attacks dimethylphosphate will substitute that water molecule. In this process, no significant energy loss or gain occurs: it remains the same for a water nucleophile and changes only 0.04 kcal/mol for a hydroxide nucleophile.

3.2. Comparison with Experimental Results. Dimethylphosphate hydrolysis in solution occurs mainly by C−O cleavage. The reaction, as we modeled it, accounts for less than 0.5% of total hydrolysis.[37] Thus, the activation energies reported for these reactions of 31.7 kcal/mol for hydroxide attack[38] and 34.9 kcal/mol for water attack[37] could only be seen as a bottom line reference. According to transition state theory, eq 2, we could roughly estimate that a reaction two hundred times slower than the above would have energies 3.1 kcal/mol higher, setting activation energies for P−O dimethylphosphate hydrolysis on 34.9 and 38.0 kcal/mol, respectively.

$$k_v = \frac{k_b \cdot T}{h} e^{-\Delta G^{\neq}/RT} \qquad (2)$$

A recent paper brought new light to this discussion.[39] By using a similar molecule to dimethylphosphate, dineopentyl phosphate, the reaction via P−O cleavage was isolated, as the attack on carbon atoms is avoided by steric hindrance. Experimentally obtained activation energies in this case are 37.7 kcal/mol for hydroxide and 38.1 kcal/mol for water attack.

Even though these values are equivalent to ours of 34.6 and 38.4 kcal/mol, respectively, a direct comparison should not be attempted because the experimental results do not treat exactly the reaction studied in our work. Moreover, we used many approximations, such as the treatment of the translational and rotational entropy and the continuum description of the solvent.

3.3. Functionals Benchmarking. The objective of this study is not to make a full evaluation of DFT functional but instead to evaluate the performance of DFT functionals with

Table 6. Activation and Reaction Energies for Dimethylphosphate Hydrolysis by Water Attack in Vacuum As Calculated by Various Methods[a]

	vacuum							
	R. I.	TS12	IM2	TS23	IM3	TS34	P. I.	P.
DFT - B3LYP	−15.3	25.1	23.5	33.8	22.4	27.3	−14.0	−0.2
Hartree−Fock	−14.4	36.4	32.8	43.1	31.2	38.5	−13.2	0.0
MP2	−16.9	23.0	20.6	31.5	19.4	26.4	−13.4	2.9
MP2/CBS	−16.6	22.2	18.8	29.4	17.5	25.5	−13.4	2.2
MP3	−16.9	25.4	21.7	32.3	20.3	28.3	−14.1	2.1
MP4(SDQ)	−16.5	25.3	22.4	33.0	21.1	28.4	−13.4	2.3
CCSD	−16.6	25.6	22.4	33.0	21.1	28.6	−13.6	2.2
CCSD/CBS	−16.2	24.5	20.2	30.6	20.6	27.4	−13.5	1.6
CCSD(T)/CBS	−16.6	22.5	18.6	29.0	17.1	25.4	−13.6	2.1
correction to the Gibbs free energy	10.2	12.1	13.2	12.0	13.1	10.5	9.0	−0.4
zero point energy	2.3	2.1	3.0	1.9	3.0	1.3	2.3	1.2
total free energy	−6.4	34.6	31.8	41.0	30.2	35.9	−4.6	1.7

[a] The zero point energy and the correction to the Gibbs free energy are also included. All values are in kcal/mol. Total free energy refers to the sum of CSSD(T)/CBS energy with the correction to the Gibbs free energy. The methods not noted with CBS refer to calculations with the 6-311++G(2d,2p) basis set. R. I. - reactants interaction, TS - transition state, IM - intermediate, P. I. - products interaction, P. - products.

Table 7. Activation and Reaction Energies for Dimethylphosphate Hydrolysis by Water Attack in Implicit Solvent As Calculated by Various Methods[a]

	implicit solvent							
	R. I.	TS12	IM2	TS23	IM3	TS34	P. I.	P.
DFT - B3LYP	−0.9	38.8	34.2	41.6	33.3	39.6	−3.8	−2.4
Hartree−Fock	1.0	51.2	44.3	52.3	43.5	52.3	−2.3	−2.8
MP2	−2.6	36.9	31.1	38.9	30.0	38.3	−2.9	0.2
MP2/CBS	−2.4	36.3	29.1	36.8	28.0	37.8	−3.0	0.1
MP3	−2.5	39.6	32.3	40.3	31.3	40.7	−3.5	−0.6
MP4(SDQ)	−2.2	39.2	32.8	40.7	31.8	40.4	−3.1	−0.4
CCSD	−2.2	39.5	32.9	40.8	31.9	40.7	−3.1	−0.5
CCSD/CBS	−2.1	38.5	30.5	38.2	29.5	39.9	−3.1	−0.5
CCSD(T)/CBS	−2.8	36.3	28.8	36.5	27.6	37.6	−3.2	0.0
correction to the Gibbs free energy	−2.2	1.7	−0.7	1.9	−0.7	−2.3	−3.9	1.2
zero point energy	2.2	1.6	2.5	1.3	2.5	1.2	2.0	0.4
total free energy	−5.0	38.0	28.1	38.4	26.9	35.3	−7.1	1.2

[a] The zero point energy and the correction to the Gibbs free energy are also included. All values are in kcal/mol. The correction to the Gibbs free energy does not contain the translational and rotational entropic energies. Total free energy refers to the sum of CSSD(T)/CBS energy with the correction to the Gibbs free energy. The methods not noted with CBS refer to calculations with the 6-311++G(2d,2p) basis set. R. I. - reactants interaction, TS - transition state, IM - intermediate, P. I. - products interaction, P. - products.

regards to phosphodiester bonds hydrolysis reactions. Hence, our results can only be examined in that context, and further conclusions should not be deduced. Due to the large heterogeneity of DFT functionals it is likely that our results can be significantly different from others concerning unrelated reactions.

The ranking criteria for the functional evaluation are the mean unsigned errors (MUE) and mean signed errors (MSE) between the electronic energy of each functional and their approximated CCSD(T)/CBS counterpart. The calculations were carried out on the same geometries. The unsigned error is used to assess the absolute deviation from the reference value, while the signed error measures a systematic deviation. When analyzing these two values we should be aware that a) a positive MSE means an average overestimation of energy values and a negative MSE an average underestimation; b) MUE must be larger or equal to MSE; c) if MSE is equal to MUE the functional always overestimated the energy in considered structures; and d) if MSE is symmetric to MUE the functional always underestimated the energy.

Tables 8, 9, and 10 show the MUE and MSE values for all the functionals tested. Table 8 contemplates all structures, Table 9 refers only to results in vacuum, and Table 10 refers to those in the solvent. Some general trends can be inferred from the analysis of these tables:

• Top ranked functionals are mostly H-GGA (hybrid-generalized gradient approximation) and HM-GGA (hybrid-meta-GGA). As a group, LDA (local density approximation) functionals give the poorest results. It should be noted, however, that the use of a H-GGA or a HM-GGA functional does not guarantee by itself improved estimated energies, since the middle and bottom of the discussed tables are also populated by those.

• When comparing MUEs for energy barriers and reaction energies, an unusual scenario arises. For the top results, reaction energies MUEs are larger than energy barriers MUEs. The probable source for this particularity may lie in the products structures, as a systematic deviation between DFT and CCSD(T) energies could be observed in the reactants energies but not in the activation energies. For the hydroxide attack system in vacuum, for example, the reaction energy calculated with approximated CCSD(T)/CBS is −3.8 kcal/mol, whereas for most DFT functionals is circa −9 kcal/

Table 8. MUEs and MSEs for the Best Studied Functionals for Vacuum and Solvent Systems[a]

		transition state		intermediate		minimum		all	
functional	type	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE
MPWB1K	HM-GGA	0.34	−0.10	0.16	−0.04	1.46	1.46	0.75	0.54
MPW1B95	HM-GGA	0.83	−0.79	0.10	−0.02	1.56	1.56	0.97	0.30
PBE1PBE	H-GGA	1.02	−0.67	0.70	0.7	2.01	2.01	1.35	0.68
BB1K	HM-GGA	1.05	1.02	1.25	1.25	1.85	1.85	1.41	1.39
TPSSh	HM-GGA	1.52	−1.44	0.26	0.26	1.93	1.93	1.43	0.25
B3P86	H-GGA	1.06	−0.22	1.37	1.37	2.02	2.02	1.50	0.99
B1B95	HM-GGA	0.88	0.53	1.54	1.54	2.12	−0.83	1.51	0.19
TPSS1KCIS	HM-GGA	1.14	−0.26	1.67	1.67	2.00	2.00	1.59	1.03
MPW1PW91	H-GGA	1.13	0.67	2.15	2.15	2.14	2.14	1.74	1.55
TPSSTPSS	M-GGA	2.40	−2.40	0.18	−0.15	1.97	1.97	1.79	−0.20
MPW1N	H-GGA	1.60	1.60	2.21	2.21	2.06	2.06	1.91	1.91
MPW1S	H-GGA	1.62	−0.79	1.87	1.87	2.26	2.26	1.93	0.96
MPW1K	H-GGA	1.88	1.88	2.21	2.21	2.02	2.02	2.00	2.00
BP86	GGA	1.92	−1.19	1.83	1.83	2.28	2.28	2.05	0.80
TPSSLYP1W	M-GGA	1.61	0.57	3.59	3.59	1.71	1.71	2.05	1.63
B97-1	H-GGA	1.57	1.14	3.23	3.23	1.98	1.98	2.06	1.89
MPWPW91	GGA	2.02	−1.34	1.72	1.72	2.31	2.31	2.07	0.73
MPWB95	M-GGA	3.12	−3.12	0.39	−0.39	1.93	1.93	2.10	−0.56
BB95	M-GGA	1.74	−1.03	1.76	1.76	2.66	2.66	2.11	1.00
PBEPBE	GGA	3.07	−3.07	0.19	−0.12	2.11	2.11	2.11	−0.41
PBE1W	GGA	1.88	−0.71	2.64	2.64	2.19	2.19	2.16	1.12
MPW3LYP	H-GGA	1.80	1.45	3.73	3.73	1.85	1.85	2.20	2.07
MPWLYP	GGA	1.96	0.30	4.05	4.05	2.04	2.04	2.41	1.74
B3PW91	H-GGA	1.92	1.80	3.66	3.66	2.33	2.33	2.43	2.38
PBE1KCIS	HM-GGA	2.11	1.86	3.97	3.97	2.02	2.02	2.44	2.34
B98	H-GGA	2.15	1.97	4.05	4.05	2.06	2.06	2.50	2.42
BPBE	GGA	1.79	0.58	3.66	3.66	2.90	2.90	2.61	2.13
BPW91	GGA	1.83	0.78	3.91	3.91	2.85	2.85	2.65	2.23
MPWLYP1W	GGA	2.13	1.14	4.98	4.98	2.11	2.11	2.69	2.30
MPW1KCIS	HM-GGA	2.56	2.37	5.12	5.12	2.27	2.27	2.96	2.88
B3LYP	H-GGA	2.94	2.94	5.28	5.28	2.02	2.02	3.04	3.04
PBELYP1W	GGA	2.80	2.02	6.08	6.08	2.09	2.09	3.17	2.86
BLYP	GGA	2.88	2.36	6.17	6.17	2.36	2.36	3.33	3.12
B97-2	H-GGA	3.51	3.51	5.58	5.58	2.43	2.43	3.49	3.49
B1LYP	H-GGA	3.78	3.78	5.95	5.95	1.98	1.98	3.49	3.49
MPWKCIS1K	HM-GGA	4.17	4.17	5.26	5.26	2.07	2.07	3.55	3.55
BhandHLYP	H-GGA	4.57	4.57	5.33	5.33	1.74	1.74	3.59	3.59
VSXC	M-GGA	5.15	−5.15	4.14	−4.14	3.40	3.40	4.25	−1.53
HCTH147	GGA	4.13	4.03	7.60	7.6	2.92	2.92	4.34	4.30
HCTH407	GGA	5.12	5.07	8.82	8.82	3.05	3.05	5.03	5.01

[a] All values are in kcal/mol. Functionals are ordered by the "all structures" MUE. M06 functionals are not included in this table, as they were not used for solvent systems.

mol. A bad description of some bond by the DFT functionals is likely to be responsible for this discrepancy.

• If we consider some recent benchmarking studies,[40,41] our ranking of functionals is not entirely expected. Notably, MPWB1K, the first rated functional, is more accurate than any of the M06 family functionals. As these functionals belong to the same lineage, with M06 functionals being the more recent, our results are out of the ordinary. We also fail to observe the assumption that a greater fraction of Hartree−Fock exchange leads to better results for activation energies. Furthermore, PBE1PBE also performs very well, whereas in those other studies that is not the case. This simply shows how DFT methods in general are biased toward certain reactions, due to their intrinsic parametrization and (many times) error cancelation. In these studies, carbon species are prevailing, and it is acceptable that our phosphodiester system does not agree exactly with those. This disagreement by itself is sufficient to prove the significance of this study.

• Excluding the above conclusions, we find it very difficult to make any other assumptions regarding the functionals benchmarking. Due to the great heterogeneity of functionals and their reliance on empirical parametrization, we believe that a systematic analysis would not reach appreciable conclusions. Because of this, we adopted a more descriptive approach, and we will not try to interpret why certain functionals give better or worse results.

According to the MUE for all the systems, MPWB1K is the more accurate functional for studying phosphodiester bond hydrolysis, followed by MPW1B95 and PBE1PBE. The first two of these functionals have chemical accuracy, meaning a MUE smaller than 1 kcal/mol. MPWB1K and MPW1B95 also give the lowest MUE when considering barrier heights, followed by B1B95. In this case, these three functionals have chemical accuracy. For reaction energies, MPWB1K, MPW1B95, and TPSSLYP1W give the best results but with errors larger than 1 kcal/mol. We should also point out that B3LYP did not perform particularly well and, therefore, is not suited to study the energetic of this type of reactions. Performance of the DFT functionals in solution and implicit solvent is very similar as shown in Tables 9 and 10. There is no major reorganization in their

Table 9. MUEs and MSEs for the Best Studied Functionals for Vacuum Systems[a]

		transition state		intermediate		minimum		all	
functional	type	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE
MPWB1K	HM-GGA	0.40	−0.40	0.20	−0.20	1.69	−1.61	0.88	−0.84
MPW1B95	HM-GGA	1.10	−1.10	0.10	−0.09	1.83	−1.50	1.19	−1.06
BB1K	HM-GGA	0.78	0.72	1.09	1.09	2.09	−1.33	1.36	−0.02
PBE1PBE	H-GGA	1.37	−1.09	0.56	0.56	2.29	−2.10	1.58	−1.16
B1B95	HM-GGA	0.90	0.23	1.48	1.48	2.44	−1.15	1.63	−0.07
B3P86	H-GGA	1.23	−0.64	1.20	1.20	2.27	−1.89	1.64	−0.77
TPSSh	HM-GGA	1.88	−1.81	0.23	0.23	2.19	−1.54	1.67	−1.30
MPW1N	H-GGA	1.15	1.15	1.93	1.93	2.30	−1.99	1.77	0.05
TPSS1KCIS	HM-GGA	1.39	−0.68	1.60	1.60	2.27	−1.57	1.78	−0.58
MPW1PW91	H-GGA	1.11	0.21	1.97	1.97	2.41	−1.87	1.80	−0.27
MPW1K	H-GGA	1.27	1.27	1.92	1.92	2.29	−2.01	1.80	0.09
M06	HM-GGA	1.82	−1.82	1.53	−1.53	1.96	−1.96	1.82	−1.82
TPSSTPSS	M-GGA	2.79	−2.79	0.19	−0.11	2.28	−1.49	2.06	−1.73
B97-1	H-GGA	1.52	0.65	3.07	3.07	2.26	−1.94	2.12	0.10
MPW3LYP	H-GGA	1.55	0.87	3.49	3.49	2.10	−1.89	2.16	0.29
TPSSLYP1W	M-GGA	1.72	0.01	3.50	3.50	1.97	−1.27	2.18	0.20
MPW1S	H-GGA	2.06	−1.27	1.81	1.81	2.58	−1.79	2.22	−0.86
B3PW91	H-GGA	1.59	1.34	3.48	3.48	2.61	−1.58	2.38	0.60
BP86	GGA	2.41	−1.63	1.82	1.82	2.62	−1.58	2.38	−0.92
PBE1KCIS	HM-GGA	1.84	1.35	3.78	3.78	2.29	−1.85	2.41	0.56
MPWPW91	GGA	2.55	−1.83	1.70	1.70	2.65	−1.79	2.42	−1.11
B98	H-GGA	1.80	1.44	3.84	3.84	2.34	−1.88	2.42	0.59
MPWB95	M-GGA	3.50	−3.50	0.27	−0.27	2.45	−1.39	2.43	−2.01
PBEPBE	GGA	3.53	−3.53	0.23	−0.09	2.47	−2.08	2.45	−2.27
PBE1W	GGA	2.37	−1.24	2.62	2.62	2.54	−1.85	2.49	−0.71
BB95	M-GGA	2.18	−1.38	1.87	1.87	3.19	−0.92	2.52	−0.55
MPWLYP	GGA	2.20	−0.37	3.95	3.95	2.37	−1.77	2.62	−0.07
B3LYP	H-GGA	2.37	2.37	5.03	5.03	2.25	−1.56	2.86	1.33
MPWLYP1W	GGA	2.31	0.46	4.87	4.87	2.44	−1.66	2.87	0.49
BPBE	GGA	2.00	0.12	3.64	3.64	3.40	−1.34	2.89	0.24
BPW91	GGA	2.00	0.30	3.88	3.88	3.34	−1.32	2.91	0.37
MPW1KCIS	HM-GGA	2.36	1.98	4.95	4.95	2.53	−1.60	2.94	1.14
M06-HF	HM-GGA	4.64	−0.53	1.51	1.51	2.08	−1.53	2.99	−0.52
B1LYP	H-GGA	3.20	3.20	5.67	5.67	2.20	−1.49	3.29	1.82
PBELYP1W	GGA	2.89	1.34	5.97	5.97	2.42	−1.65	3.32	1.07
BhandHLYP	H-GGA	3.99	3.99	4.89	4.89	1.94	−1.80	3.35	1.85
MPWKCIS1K	HM-GGA	3.64	3.64	4.91	4.91	2.31	−1.77	3.36	1.73
B97-2	H-GGA	3.00	3.00	5.36	5.36	2.75	−1.63	3.37	1.62
BLYP	GGA	2.77	1.73	6.06	6.06	2.76	−1.30	3.42	1.38
VSXC	M-GGA	5.13	−5.13	3.76	−3.76	2.74	1.37	3.90	−2.26

[a] All values are in kcal/mol. Functionals are ordered by the "all structures" MUE.

positions on these two tables, but, as a general trend, vacuum structures have a slightly larger MUE.

3.4. Basis Set Benchmarking. Basis set benchmarking analysis is a straightforward one: larger basis sets produce better results but are also more computationally demanding. This relationship is not linear though. Usually, when approaching the complete basis set limit, measured properties start converging to a point where computational effort does not justify the gain in terms of accuracy. Note that relative energies converge much faster than absolute energies. Dunning correlation consistent basis sets are especially useful to evaluate this effect, since they produce results in which the energy improves consistently as we go from double-ζ to triple-ζ and beyond, in particular when applied to post-Hartree−Fock methods.

Table 11 shows the results for the basis set benchmarking regarding phosphodiester hydrolysis. At the bottom of the table are the correlation consistent basis sets results. Aug-cc-pvqt was the most complete basis set we could use, conditioned by computational resources and our model size. It is noticeable, however, that in the transition from triple-ζ (aug-cc-pvtz) to quadruple-ζ (aug-cc-pvqz), MUE for energy values change around 1 kcal/mol. This is an acceptable convergence criterion and basically means that calculations with a quintuple-zeta basis set (aug-cc-pz5z) would not be significantly different.

Regarding all others basis sets, we could easily isolate the factors that most influence the results. Diffuse functions on heavy atoms seem to be the most prominent factors. Their inclusion diminishes the total MUE by more than 6 kcal/mol, even in an already large basis set such as 6-311G(2d,2p). This comes with no surprise though, since it is known that diffuse functions are very important in the electronic description of negatively charged molecules with more loosely bound electrons, as in these phosphodiester systems. Accordingly, hydroxide attacking systems originate larger MUEs as opposed to water attacking systems, since they have one extra negative charge. In the transition state of the hydroxide reactions the error due to the noninclusion of diffuse functions is superior to 20 kcal/mol.

The inclusion of polarization functions is very important, both in hydrogen and heavy atoms. The difference between 6 and 311++G(d,p) and 6-311++G(2d,2p) MUEs, for example, is superior to 1 kcal/mol, and it increases to about

Table 10. MUEs and MSEs for the Best Studied Functionals for Implicit Solvent Systems[a]

		transition state		intermediate		minimum		all	
functional	type	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE
MPWB1K	HM-GGA	0.27	0.20	0.12	0.12	1.23	−0.86	0.63	−0.24
MPW1B95	HM-GGA	0.56	−0.49	0.09	0.06	1.29	−0.81	0.75	−0.51
PBE1PBE	H-GGA	0.67	−0.26	0.84	0.84	1.74	−1.36	1.13	−0.48
TPSSh	HM-GGA	1.16	−1.07	0.30	0.3	1.67	−0.94	1.19	−0.74
B3P86	H-GGA	0.88	0.19	1.54	1.54	1.77	−1.2	1.37	−0.10
B1B95	HM-GGA	0.85	0.83	1.61	1.61	1.80	−0.5	1.38	0.45
TPSS1KCIS	HM-GGA	0.90	0.17	1.75	1.75	1.73	−0.94	1.40	0.04
BB1K	HM-GGA	1.31	1.31	1.40	1.40	1.60	−0.62	1.45	0.56
TPSSTPSS	M-GGA	2.02	−2.02	0.18	−0.18	1.66	−0.92	1.51	−1.21
MPW1S	H-GGA	1.18	−0.31	1.92	1.92	1.95	−1.1	1.63	−0.18
MPW1PW91	H-GGA	1.15	1.12	2.33	2.33	1.88	−1.13	1.67	0.46
BB95	M-GGA	1.30	−0.69	1.65	1.65	2.14	−0.34	1.70	−0.08
BP86	GGA	1.43	−0.75	1.84	1.84	1.94	−0.97	1.71	−0.32
MPWPW91	GGA	1.49	−0.84	1.74	1.74	1.97	−1.11	1.73	−0.43
MPWB95	M-GGA	2.75	−2.75	0.51	−0.51	1.40	−0.76	1.76	−1.51
PBEPBE	GGA	2.61	−2.61	0.15	−0.15	1.76	−1.4	1.78	−1.64
PBE1W	GGA	1.40	−0.18	2.67	2.67	1.85	−1.17	1.83	−0.01
TPSSLYP1W	M-GGA	1.51	1.14	3.68	3.68	1.44	−0.71	1.92	0.91
B97-1	H-GGA	1.62	1.62	3.39	3.39	1.69	−1.2	2.01	0.85
MPW1N	H-GGA	2.05	2.05	2.50	2.50	1.81	−1.18	2.05	0.85
MPW1K	H-GGA	2.50	2.50	2.51	2.51	1.76	−1.14	2.20	1.04
MPWLYP	GGA	1.73	0.97	4.15	4.15	1.71	−1.09	2.20	0.78
MPW3LYP	H-GGA	2.04	2.04	3.96	3.96	1.59	−1.18	2.25	1.14
BPBE	GGA	1.58	1.04	3.68	3.68	2.41	−0.7	2.33	0.87
BPW91	GGA	1.67	1.25	3.94	3.94	2.36	−0.69	2.40	1.01
PBE1KCIS	HM-GGA	2.37	2.37	4.16	4.16	1.75	−1.12	2.48	1.33
B3PW91	H-GGA	2.26	2.26	3.83	3.83	2.05	−0.87	2.49	1.32
MPWLYP1W	GGA	1.96	1.82	5.10	5.10	1.78	−0.99	2.52	1.35
B98	H-GGA	2.50	2.50	4.26	4.26	1.79	−1.12	2.57	1.40
MPW1KCIS	HM-GGA	2.76	2.76	5.30	5.30	2.01	−0.93	2.97	1.79
PBELYP1W	GGA	2.70	2.70	6.19	6.19	1.77	−0.99	3.03	1.93
B3LYP	H-GGA	3.52	3.52	5.52	5.52	1.78	−0.88	3.22	2.16
BLYP	GGA	3.00	3.00	6.28	6.28	1.96	−0.67	3.24	2.19
B97-2	H-GGA	4.02	4.02	5.79	5.79	2.11	−0.86	3.61	2.42
B1LYP	H-GGA	4.36	4.36	6.23	6.23	1.76	−0.79	3.69	2.67
MPWKCIS1K	HM-GGA	4.71	4.71	5.61	5.61	1.84	−0.96	3.74	2.62
BhandHLYP	H-GGA	5.15	5.15	5.77	5.77	1.55	−0.99	3.83	2.82
HCTH147	GGA	4.66	4.66	7.74	7.74	2.38	−0.83	4.37	3.08
VSXC	M-GGA	5.18	−5.18	4.52	−4.52	4.06	2.47	4.60	−1.99
HCTH407	GGA	5.82	5.82	8.99	8.99	2.53	−0.92	5.14	3.76

[a] All values are in kcal/mol. Functionals are ordered by the "all structures" MUE. M06 functionals are not included in this table, as they were not used for solvent systems.

Table 11. MUEs for the Analyzed Basis Sets[a]

			hydroxide		water	
	basis set	all structures	barrier energies	reaction energies	barrier energies	reaction energies
140	6-31G(d)	8.69	25.22	16.86	6.82	6.08
172	6-31+G(d)	1.98	4.16	0.42	2.39	2.02
176	6-311G(d)	8.34	27.10	16.27	5.94	5.55
196	6-31+G(d,p)	1.98	3.85	0.28	2.61	1.97
200	6-311G(d,p)	8.26	27.06	16.11	6.09	5.29
208	6-311+G(d)	2.23	4.92	1.43	2.43	2.15
216	6-311++G(d)	2.27	5.14	1.63	2.44	2.15
232	6-311+G(d,p)	1.44	3.82	0.29	1.75	1.26
240	6-311++G(d,p)	1.50	4.07	0.51	1.78	1.27
264	6-311G(2d,2p)	6.88	23.61	16.69	4.27	3.91
296	6-311+G(2d,2p)	0.39	1.09	0.41	0.36	0.33
304	6311++G(2d,2p)	0.42	1.37	0.68	0.38	0.26
464	6311++G(3df,3pd)	0.27	0.89	0.83	0.13	0.14
260	aug-cc-pvdz	4.53	7.78	0.02	6.93	4.18
556	aug-cc-pvtz	0.70	1.25	0.01	1.08	0.62
1012	aug-cc-pvqz	0.00	0.00	0.00	0.00	0.00

[a] Values are in kcal/mol. MSEs were not included because they have always the same absolute value as MUEs but with negative sign.

2 kcal/mol when we consider transition state structures. Further improvement of the basis set with superior angular momentum polarizable functions does not justify the computational effort. With the 6-311++G(3df,3pd) basis set the

number of basis functions raises more than 30% in relation to 6-311++G(2d,2p) and the MUE barely improves.

The splitting of the sp shell has lesser significance than we would have expected, as there is little or no improvement in the energy values when shifting from double-ζ to triple-ζ basis sets.

4. Conclusions

During this work we were able to describe the reaction path of the phosphodiester bond hydrolysis of dimethylphosphate at the approximated CCSD(T)/CBS//B3LYP/6-311++G(2d,2p) level. We modeled the reactions with hydroxide and water as nucleophiles, both in the gas phase and in implicit solvent. The geometries and energies obtained are in agreement with previous studies by other groups.

We then used these structures to perform a benchmarking study on 52 density functionals. HF, MP2, MP3, MP4, and CCSD single point energies were also computed. Approximated CCSD(T)/CBS energy values were used as reference. We concluded that, overall, MPWB1K, MPW1B95, and PBE1PBE give the lowest MUEs of all functionals: 0.75, 0.97, and 1.35 kcal/mol, respectively. Taking into account just the activation energies, MPWB1K, MPW1B95, and B1B95 are the best positioned functionals, with MUEs of 0.34, 0.83, and 0.88 kcal/mol, respectively.

A basis set benchmarking, carried out on the same systems, showed us that diffuse functions on heavy atoms are a very important variable when choosing a basis set for the systems. Regarding polarization functions, having two additional functions of higher angular momentum for each atom is the best option. As a result of our calculations, we are able to point out the 6-311+G(2d,2p) basis set as having the best equilibrium between accuracy and computational cost.

We believe that our benchmarking results will be of great help to further studies on related phosphodiester systems. This includes not only pure chemical problems but also biochemical studies in which DNA, RNA, and phospholipids are required to be depicted at a quantum level.

Acknowledgment. The authors acknowledge the FCT (Fundação para a Ciência e Tecnlogia) for financial support through project PTDC/QUI/68302/2006.

References

(1) Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*, 4th ed.; Palgrave Macmillan: 2004; pp 276−277.

(2) Westheimer, F. H. Why Nature Chose Phosphates. *Science* **1987**, *235*, 1173–1178.

(3) Lim, C.; Karplus, M. Nonexistence of Dianionic Pentacovalent Intermediates in an Abinitio Study of the Base-Catalyzed-Hydrolysis of Ethylene Phosphate. *J. Am. Chem. Soc.* **1990**, *112*, 5872–5873.

(4) Dejaegere, A.; Lim, C.; Karplus, M. Dianionic Pentacoordinate Species in the Base-Catalyzed-Hydrolysis of Ethylene and Dimethyl-Phosphate. *J. Am. Chem. Soc.* **1991**, *113*, 4353–4355.

(5) Uchimaru, T.; Tanabe, K.; Nishikawa, S.; Taira, K. Abinitio Studies of a Marginally Stable Intermediate in the Base-Catalyzed Methanolysis of Dimethyl-Phosphate and Nonexistence of the Stereoelectronically Unfavorable Transition-State. *J. Am. Chem. Soc.* **1991**, *113*, 4351–4353.

(6) Lim, C.; Tole, P. Concerted Hydroxyl Ion Attack and Pseudorotation in the Base-Catalyzed-Hydrolysis of Methyl Ethylene Phosphate. *J. Phys. Chem.* **1992**, *96*, 5217–5219.

(7) Liang, C. X.; Ewig, C. S.; Stouch, T. R.; Hagler, A. T. Abinitio Studies of Lipid Model Species 0.1. Dimethyl-Phosphate and Methyl Propyl Phosphate Anions. *J. Am. Chem. Soc.* **1993**, *115*, 1537–1545.

(8) Dejaegere, A.; Liang, X. L.; Karplus, M. Phosphate Ester Hydrolysis - Calculation of Gas-Phase Reaction Paths and Solvation Effects. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 1763–1770.

(9) Landin, J.; Pascher, I.; Cremer, D. Ab-Initio and Semiempirical Conformation Potentials for Phospholipid Head Groups. *J. Phys. Chem.* **1995**, *99*, 4471–4485.

(10) Hu, C. H.; Brinck, T. Theoretical studies of the hydrolysis of the methyl phosphate anion. *J. Phys. Chem. A* **1999**, *103*, 5379–5386.

(11) Mercero, J. M.; Barrett, P.; Lam, C. W.; Fowler, J. E.; Ugalde, J. M.; Pedersen, L. G. Quantum mechanical calculations on phosphate hydrolysis reactions. *J. Comput. Chem.* **2000**, *21*, 43–51.

(12) Murashov, V. V.; Leszczynski, J. A comparison of the B3LYP and MP2 methods in the calculation of phosphate complexes. *J. Mol. Struct. - Theochem* **2000**, *529*, 1–14.

(13) Lopez, X.; Dejaegere, A.; Karplus, M. Solvent effects on the reaction coordinate of the hydrolysis of phosphates and sulfates: Application of Hammond and anti-Hammond postulates to understand hydrolysis in solution. *J. Am. Chem. Soc.* **2001**, *123*, 11755–11763.

(14) Bianciotto, M.; Barthelat, J. C.; Vigroux, A. Reactivity of phosphate monoester monoanions in aqueous solution. 1. Quantum mechanical calculations support the existence of "anionic zwitterion" MeO+(H)PO3 (2-) as a key intermediate in the dissociative hydrolysis of the methyl phosphate anion. *J. Am. Chem. Soc.* **2002**, *124*, 7573–7587.

(15) Bianciotto, M.; Barthelat, J. C.; Vigroux, A. Reactivity of phosphate monoester monoanions in aqueous solution. 2. A theoretical study of the elusive zwitterion intermediates RO+(H)PO32-. *J. Phys. Chem. A* **2002**, *106*, 6521–6526.

(16) Lopez, X.; York, D. M.; Dejaegere, A.; Karplus, M. Theoretical studies on the hydrolysis of phosphate diesters in the gas phase, solution, and RNase A. *Int. J. Quantum Chem.* **2002**, *86*, 10–26.

(17) Kolandaivel, P.; Kanakaraju, R. Structure, stability and interaction studies on nucleotide analogue systems. *Int. J. Mol. Sci.* **2003**, *4*, 486–502.

(18) Wang, Y. N.; Topol, I. A.; Collins, J. R.; Burt, S. K. Theoretical studies on the hydrolysis of mono-phosphate and tri-phosphate in gas phase and aqueous solution. *J. Am. Chem. Soc.* **2003**, *125*, 13265–13273.

(19) Chen, X.; Zhan, C. G. Theoretical determination of activation free energies for alkaline hydrolysis of cyclic and acyclic phosphodiesters in aqueous solution. *J. Phys. Chem. A* **2004**, *108*, 6407–6413.

(20) Range, K.; McGrath, M. J.; Lopez, X.; York, D. M. The structure and stability of biological metaphosphate, phosphate, and phosphorane compounds in the gas phase and in solution. *J. Am. Chem. Soc.* **2004**, *126*, 1654–1665.

(21) Iche-Tarrat, N.; Barthelat, J. C.; Rinaldi, D.; Vigroux, A. Theoretical studies of the hydroxide-catalyzed P-O cleavage reactions of neutral phosphate triesters and diesters in aqueous solution: Examination of the changes induced by H/Me substitution. *J. Phys. Chem. B* **2005**, *109*, 22570–22580.

(22) Imhof, P.; Fischer, S.; Kramer, R.; Smith, J. C. Density functional theory analysis of dimethylphosphate hydrolysis: effect of solvation and nucleophile variation. *J. Mol. Struct. - Theochem* **2005**, *713*, 1–5.

(23) Lopez, X.; Dejaegere, A.; Leclerc, F.; York, D. M.; Karplus, M. Nucleophilic attack on phosphate diesters: A density functional study of in-line reactivity in dianionic, monoanionic, and neutral systems. *J. Phys. Chem. B* **2006**, *110*, 11525–11539.

(24) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. General performance of density functionals. *J. Phys. Chem. A* **2007**, *111*, 10439–10452.

(25) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. *J. Chem. Phys.* **2002**, *117*, 43–54.

(26) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. Ab initio study of ionic solutions by a polarizable continuum dielectric model. *Chem. Phys. Lett.* **1998**, *286*, 253–260.

(27) Mennucci, B.; Tomasi, J. Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries. *J. Chem. Phys.* **1997**, *106*, 5151–5158.

(28) Cances, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

(29) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.

(30) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N-2, and H2O. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

(31) Jurecka, P.; Hobza, P. On the convergence of the (Delta E-CCSD(T)-Delta E-MP2) term for complexes with multiple H-bonds. *Chem. Phys. Lett.* **2002**, *365*, 89–94.

(32) Hobza, P.; Sponer, J. Toward true DNA base-stacking energies: MP2, CCSD(T), and complete basis set calculations. *J. Am. Chem. Soc.* **2002**, *124*, 11802–11808.

(33) Schrodinger. *Jaguar, version 7.5*; New York, 2008.

(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al.Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision C.02*; Gaussian, Inc.: Wallingford, CT, 2004.

(35) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley-Interscience: 1997; pp *136*−140.

(36) Gorenstein, D. G.; Luxon, B. A.; Findlay, J. B. Stereoelectronic Effects in the Reactions of Phosphate Diesters - Abinitio Molecular-Orbital Calculations of Reaction Surfaces. 2. *J. Am. Chem. Soc.* **1979**, *101*, 5869–5875.

(37) Wolfenden, R.; Ridgway, C.; Young, G. Spontaneous hydrolysis of ionized phosphate monoesters and diesters and the proficiencies of phosphatases and phosphodiesterases as catalysts. *J. Am. Chem. Soc.* **1998**, *120*, 833–834.

(38) Kumamoto, J.; Cox, J. R.; Westheimer, F. H. Barium Ethylene Phosphate. *J. Am. Chem. Soc.* **1956**, *78*, 4858–4860.

(39) Schroeder, G. K.; Lad, C.; Wyman, P.; Williams, N. H.; Wolfenden, R. The time required for water attack at the phosphorus atom of simple phosphodiesters and of DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 4052–4055.

(40) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(41) Zhao, Y.; Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **2008**, *41*, 157–167.

(42) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin-Density Calculations - a Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.

(43) Slater, J. C. *Quantum Theory of Molecular and Solids. Vol. 4: The Self-Consistent Field for Molecular and Solids*; McGraw-Hill: New York, 1974.

(44) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.

(45) Becke, A. D. Density-functional thermochemistry 0.4. A new dynamical correlation functional and implications for exact-exchange mixing. *J. Chem. Phys.* **1996**, *104*, 1040–1046.

(46) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098–3100.

(47) Becke, A. D. Density-Functional Thermochemistry 0.3. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(48) Zhao, Y.; Truhlar, D. G. Hybrid meta density functional theory methods for thermochemistry, thermochemical kinetics, and noncovalent interactions: The MPW1B95 and MPWB1K models and comparative assessments for hydrogen bonding and van der Waals interactions. *J. Phys. Chem. A* **2004**, *108*, 6908–6918.

(49) Perdew, J. P. Density-Functional Approximation for the Correlation-Energy of the Inhomogeneous Electron-Gas. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *33*, 8822–8824.

(50) Perdew, J. P. Unified Theory of Exchange and Correlation Beyond the Local Density Aproximation. In *Electronic*

Structure of Solids; Zieche, P., Eschrig, H., Eds.; Berlin, Germany, 1991; pp 11−20.

(51) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. Development and assessment of new exchange-correlation functionals. *J. Chem. Phys.* **1998**, *109*, 6264–6271.

(52) Zhao, Y.; Truhlar, D. G. Density functional for spectroscopy: No long-range self-interaction error, good performance for Rydberg and charge-transfer states, and better performance on average than B3LYP for ground states. *J. Phys. Chem. A* **2006**, *110*, 13126–13130.

(53) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(54) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. Hybrid exchange-correlation functional determined from thermochemical data and ab initio potentials. *J. Chem. Phys.* **2001**, *115*, 9233–9242.

(55) Adamo, C.; Barone, V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models. *J. Chem. Phys.* **1998**, *108*, 664–675.

(56) Schmider, H. L.; Becke, A. D. Optimized density functionals from the extended G2 test set. *J. Chem. Phys.* **1998**, *108*, 9624–9631.

(57) Becke, A. D. Density-functional thermochemistry 0.5. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560.

(58) Krieger, J. B.; Chen, J. Q.; Iafrate, G. J.; Savin, A. Construction of an accurate self-interaction-corrected correlation energy functional based on an electron gas with a gap. *Electron Correl. Mater. Prop., [Proc. Int. Workshop], 1st* **1999**, 463–477.

(59) Zhao, Y.; Gonzalez-Garcia, N.; Truhlar, D. G. Benchmark database of barrier heights for heavy atom transfer, nucleophilic substitution, association, and unimolecular reactions and its use to test theoretical methods. *J. Phys. Chem. A* **2005**, *109*, 2012–2018.

(60) Gill, P. M. W. A new gradient-corrected exchange functional. *Mol. Phys.* **1996**, *89*, 433–445.

(61) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. Adiabatic connection for kinetics. *J. Phys. Chem. A* **2000**, *104*, 4811–4815.

(62) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43–52.

(63) Kormos, B. L.; Cramer, C. J. Adiabatic connection method for X-+RX nucleophilic substitution reactions (X = F, Cl). *J. Phys. Org. Chem.* **2002**, *15*, 712–720.

(64) Tao, J. M.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(65) Dahlke, E. E.; Truhlar, D. G. Improved density functionals for water. *J. Phys. Chem. B* **2005**, *109*, 15677–15683.

(66) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. Effectiveness of diffuse basis functions for calculating relative energies by density functional theory. *J. Phys. Chem. A* **2003**, *107*, 1384–1388.

(67) Handy, N. C.; Cohen, A. J. Left-right correlation energy. *Mol. Phys.* **2001**, *99*, 403–412.

(68) Hoe, W. M.; Cohen, A. J.; Handy, N. C. Assessment of a new local exchange functional OPTX. *Chem. Phys. Lett.* **2001**, *341*, 319–328.

(69) Zhao, Y.; Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006**, *125*, 194101.

(70) Xu, X.; Goddard, W. A. The X3LYP extended density functional for accurate descriptions of nonbond interactions, spin states, and thermochemical properties. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 2673–2677.

(71) Perdew, J. P.; Wang, Y. Accurate and Simple Analytic Representation of the Electron-Gas Correlation-Energy. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1992**, *45*, 13244–13249.

(72) Van Voorhis, T.; Scuseria, G. E. A never form for the exchange-correlation energy functional. *J. Chem. Phys.* **1998**, *109*, 400–410.

CT900649E