

Are mixed explicit/implicit solvation models reliable for studying phosphate hydrolysis? A comparative study of continuum, explicit and mixed solvation models.

Shina C. L. Kamerlin^{*[a]}, Maciej Haranczyk^[b] and Arieh Warshel^{*[a]}

[a] Department of Chemistry
University of Southern California
3620 McClintock Ave., Los Angeles, CA-90089 (USA)
Fax: (+1) 213 740 2701
E-mail: L. Kamerlin@gmx.at, warshel@usc.edu

[b] Computational Research Division
Lawrence Berkeley National Laboratory
One Cyclotron Road, Mail Stop 50F-1650, Berkeley, CA-94720 (USA)

Abstract: Phosphate hydrolysis is ubiquitous in biology. However, despite intensive research on this class of reactions, the precise nature of the reaction mechanism remains controversial. In this work, we have examined the hydrolysis of three homologous phosphate diesters. The solvation free energy was simulated by means of either an implicit solvation model (COSMO), hybrid quantum mechanical / molecular mechanical free energy perturbation (QM/MM-FEP) or a mixed solvation model in which N water molecules were explicitly included in the *ab initio* description of the reacting system (where $N=1-3$), with the remainder of the solvent being implicitly modelled as a continuum. Here, both COSMO and QM/MM-FEP reproduce ΔG_{obs} within an error of about 2kcal/mol. However, we demonstrate that in order to obtain any form of reliable results from a mixed model, it is essential to carefully select the explicit water molecules from short QM/MM runs that act as a model for the true infinite system. Additionally, the mixed models tend to be increasingly inaccurate the more explicit water molecules are placed into the system. Thus, our analysis indicates that this approach provides an unreliable way for modelling phosphate hydrolysis in solution.

Keywords: phosphate hydrolysis · reaction mechanism · mixed solvation model · implicit solvation model · QM/MM

Introduction

Phosphate ester hydrolysis is a crucially important reaction for many biological processes, being involved in amongst, other things, DNA and RNA synthesis, protein synthesis, as well as several energy and signal transduction processes^[1-7]. Therefore, it is unsurprising that in recent years, many phosphatases have been identified as key drug targets for a wide range of diseases including (but by no means limited to) bipolar disorder, diabetes, obesity, various neurodegenerative diseases and even for cancer therapy^[8-11]. Thus, these enzymes have become the focus of intensive research. Simultaneously, several research groups have been involved in the design of artificial catalysts for phosphate hydrolysis, often based on known enzyme structures^[12-17].

Clearly, understanding the mechanisms of enzymatic phosphate ester hydrolysis is a problem of major therapeutic relevance. However, the study of enzymatic reaction mechanisms still presents a major computational challenge. Additionally, prior to undertaking a study of the enzymatic reaction mechanism, it is essential to have a clear understanding of the corresponding reaction in solution. Over the past few decades, significant research effort has been invested into the study of phosphate hydrolysis in solution (e.g. Refs.^[4, 6, 18-32]). However, despite frequent implications that many of the mechanistic issues have already been resolved experimentally^[33, 34], the precise nature of both the solution and the enzyme-catalyzed hydrolysis reactions remains controversial^[18, 23, 25, 27-29, 34-42].

Part of the problem that limits our understanding of the detailed reaction pathways for phosphate ester hydrolysis (whether in enzymes or in solution) arises to a large extent from the fact that the energies of the transition states for these hydrolysis reactions are very high, thus making it quite hard to characterize such transition states by means of key intermediates (as is often done experimentally). There have been several recent computational studies on this reaction class^[32, 39-46]. However, these have often been gas-phase studies (for a review, see Ref. ^[43]), which obtained barriers of for instance 80 kcal/mol (where the possible difference between having a minimum of a few kcal/mol and not having it is questionable). Additionally, none of these studies directly compared associative and dissociative reaction pathways for the same nucleophile and substrate. In fact, the few computational studies that *do* directly compare associative and dissociative pathways (by means of mapping full free energy surfaces) have demonstrated that the interpretation of key experimental markers such as linear free energy effects and activation entropies for the elucidation of reaction mechanisms are ambiguous, and cannot be used to define a unique reaction pathway^[26, 28, 47], thus highlighting the need for experimental work to be accompanied by careful computational studies (that of course reproduce the available experimental data) in order to be able to obtain conclusive mechanistic information.

There are several computational techniques available for modeling reaction mechanisms in solution, one of which (that is slowly starting to gain popularity in studies of phosphate hydrolysis in solution) is the use of a so-called "mixed continuum model", in which one or more water molecules are explicitly included in the *ab initio* description of the reacting system, and the remainder of the solvent is modeled implicitly as a continuum (see e.g. Refs^[40, 48, 49] for examples of the use of such an approach). This is a cause of concern, as there are several potential pitfalls with such an approach, even if one were to ignore the inherent problems associated with generating adequate free-energy surfaces for phosphate hydrolysis in the presence of explicit water molecules (and with subsequent transition state identification)^[27, 50]. Specifically, an issue of great importance when examining a solvated system is to have correct boundary conditions between the solute and the bulk, and it is highly questionable whether solvation models which mix continuum and explicit water accurately reproduce these boundary conditions. The issue of proper polarization boundary conditions between the explicit and continuum regions has been explored and emphasized in many of our former works^[28, 51-53]. It has also received recognition by workers other than us^[54]. In short, the problem with such an approach is that the water molecules in the first solvation shells will be over-polarized if they are not subjected to polarization constraints that properly represent the effect of the rest of the system in the true infinite system. This situation becomes in turn more serious when one only uses a few water molecules immersed in a continuum model as was done in Refs. ^[40, 48, 49]. Here, the explicit water molecules are not likely to have the correct orientation of the corresponding molecules in an explicit infinite system. This issue can be easily tested by simply taking four water molecules in a continuum, and checking whether the minimized structure has any similarity to the corresponding simulated structure of water in a large simulation system. In a similar problem, three water molecules near an ion will always be over-polarized relative to the polarization of the nearest neighbors in a large water sphere.

Perhaps a more serious problem is the fact that the inclusion of several explicit water molecules in a continuum model combined with an energy minimization treatment would require the evaluation of the entropic effect associated with the explicit solvent models (this effect is included implicitly in the continuum model), an issue which has been effectively been ignored by previous studies of phosphate hydrolysis that utilize such an approach^[40, 48, 49]. That is, unless one is specifically interested in the chemistry (e.g. proton transfer) associated with the presence of a few explicit water molecules, having a mixed model is problematic, and at least logically, it would be much better to use either a well calibrated simple model, or a QM/MM model with proper free energy treatment.

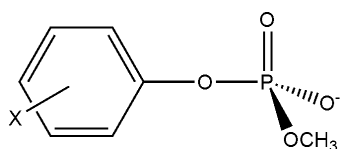


Figure 1. Prototype for the series of substituted methyl phenyl phosphate diesters examined in Refs.^[27, 55]. In this work, we have re-examined hydroxide attack on the parent compound as well as the 4-nitro and 4-cyano substituted compounds, by use of different solvation models.

Ultimately, a proper theoretical study of any issue should be judged by its consistency, and by its ability to reproduce experimental benchmarks. We are not aware of any studies that have carefully evaluated and demonstrated the advantages of mixed solvation models which include several explicit water molecules with regards to these two requirements. In a recent computational study^[27], we reproduced the experimental LFER for hydroxide attack on a series of homologous substituted methyl phenyl phosphate diesters (Fig. 1) in solution with high accuracy by use of a continuum model. Additionally, we demonstrated that despite the potential shortcomings of a continuum model relative to full explicit solvation, the continuum model used is actually quite reliable (see Supplementary Material of Ref. ^[27]). In this work, we have taken three of these compounds (i.e. the parent compound as well as the 4-nitro and 4-cyano substituted compounds), and generated 1-D free energy profiles for the hydrolysis of these compounds using three different solvation models. Specifically, we have compared calculated and experimental barriers when solvation is simulated by means of a simple continuum model, by means of hybrid quantum mechanical / molecular mechanical free energy perturbation approaches (QM/MM-FEP) and by means of mixed solvation models in which one, two or three water molecules are explicitly included in the system. In all cases, we have also evaluated the configurational entropy of each transition state. We demonstrate that both the simple continuum model and QM/MM-FEP reproduce ΔG_{exp} to within 2 kcal/mol. Similarly, in the case of the mixed solvation models, it is possible to obtain reasonable results *provided* that the explicit water molecules have been carefully selected from short QM/MM-FEP runs with full explicit water molecules. However, if the explicit water molecule is chosen at random and merely allowed to optimize with the remainder of the reacting system (as was done in Refs. ^[40, 48, 49]), the reaction barrier is over-estimated by ~15kcal/mol compared to the experimental barrier when the entropic contribution of this extra water molecule is taken into account, making this approach highly unreliable for modelling phosphate hydrolysis in solution.

Computational Methods

One dimensional free energy surfaces were generated by careful reaction coordinate pushing while constraining the distance between the phosphorus atom and the attacking nucleophile (P-O_{nuc}). Geometries were optimized at P-O_{nuc} distances of 3.5, 3.0 and then 2.7 to 1.8Å in 0.1Å increments. Also, as the transition state was approached, it was necessary to impose a constraint that keeps the distance between O_{nuc}-O_{lg} constant to prevent the leaving group from flying away (previous studies on the full 2-D free energy surface have shown that this region of the surface is very flat^[27] and thus the imposition of this constraint is not expected to have a dramatic impact on the results). Structures for the relevant phosphate and attacking hydroxide were also optimized, in the absence of any constraints, and all energies throughout are given relative to the sum energies of these two moieties at infinite separation. All geometry optimizations were performed in the gas phase, using the 6-31+G* basis set and the MPW1PW91 hybrid functional^[56]. Solvation was then simulated in one of three ways: by means of a single point correction to the SCF energy using the COSMO continuum model^[57, 58] (in this case, the larger 6-311++G basis set was used), by means of QM/MM-FEP perturbation, and finally, by means of a mixed solvation model in which one, two or three water molecules were explicitly included in the gas-phase optimization using the smaller 6-31+G* basis set, and then a single point correction to the SCF energy of the entire system was performed using COSMO/6-31++G**. In the final case, we generated 1-D reaction profiles both where the water molecules were either carefully selected from short QM/MM-FEP runs, and when the conformation of the explicit water molecule was taken at random for comparison. All *ab initio* calculations were performed using the G03 software package^[59].

QM/MM-FEP calculations were performed by means of our accelerated QM/MM approach, which was originally presented and thoroughly discussed in Ref.^[47]. Also, the applications of this approach towards predicting the solvation free energies of small biomolecules in water and calculating the pK_as of protein side chains have been demonstrated in Refs. ^[60] and ^[61] respectively. In this approach, the system is divided into two different regions: of main interest to us is the part of the system which changes (i.e. by chemical reaction), and this part of the system is designated as the solute. The remainder of the system is designated as the solvent. These two regions are then described by QM and MM approaches respectively.

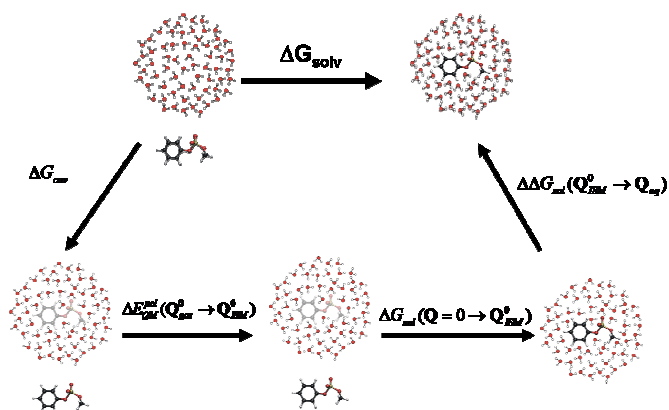


Figure 2. Thermodynamic cycle for the evaluation of the solvation free energy by means of QM/MM-FEP. This scheme was originally introduced in Refs. ^[47, 60].

In brief, our QM/MM-FEP approach follows the thermodynamic cycle shown in Fig. 2, where the free energy of charging the solvated solute to a given charge distribution (which represents the solute polarized by the solvent partial charges that are obtained with a given implicit solvation model (ISM), which may be COSMO^[57, 58], the polarized continuum model (PCM)^[62-66] or the Langevin Dipoles model^[67]) is obtained by the classical free energy perturbation adiabatic charging (FEP-AC)^[68, 69], and the free energy change associated with allowing these charges to equilibrate with the solvent potential is obtained by our hybrid QM/MM approach and is subsequently added as a correction. That is, from this cycle, the free energy can be written as:

$$\Delta G_{\text{sol}} = \Delta G_{\text{cav}} + \Delta E_{\text{QM}}^{\text{pol}}(Q_g^0 \rightarrow Q_{\text{ISM}}^0) + \Delta G_{\text{sol}}(Q = 0 \rightarrow Q_{\text{ISM}}^0) + \Delta \Delta G_{\text{sol}}(Q_{\text{ISM}}^0 \rightarrow Q_{\text{eq}}) \quad (1)$$

Here, Q_{eq} denotes the vector of equilibrated QM/MM residual atomic charges of the solute, and Q_g and Q_{ISM} denote gas-phase charges and charges obtained using the given ISM respectively. In all cases, the subscripts g and s denote gas-phase and solution states respectively, and the superscript 0 designates a constant value to the charges. In Eq. 1, first term, ΔG_{cav} , represents the solvation free energy of the non-polar neutral form of the solute. This term consists of two parts that respectively describe the hydrophobic and van der Waals free energy of the cavity. These have been implemented in the ChemSol^[70] program and are described in detail in Ref. ^[70]. The second term ($\Delta E_{\text{QM}}^{\text{pol}}(Q_g^0 \rightarrow Q_{\text{ISM}}^0)$) is the solute polarization energy predicted by an implicit solvation model. In this work, we have used the COSMO solvation model. The third term ($\Delta G_{\text{sol}}(Q = 0 \rightarrow Q_{\text{ISM}}^0)$) represents the solvation free energy of the solute, the atomic charges of which are obtained from the COSMO solvation model (not to be confused with the COSMO solvation energies). This value can be obtained by the classical FEP-AC. Finally, the last term of Eq. 1 (i.e. $\Delta \Delta G_{\text{sol}}(Q_{\text{ISM}}^0 \rightarrow Q_{\text{eq}})$) represents the free energy change associated with allowing the partial charges obtained by the given ISM (in this case COSMO) to “equilibrate”. This term can be evaluated using the LRA approach, as has been outlined in Refs. ^[47] and ^[61].

It should be noted that in this approach, we evaluate the polarization energy by adding the averaged potential from the solvent to the solute Hamiltonian. That is, we constrain the atoms of the QM region, evaluate the QM charges of all atoms in this region ($Q^{(1)}$, where (1) represents the first step of an MM/MD run) and then proceeds to run m MM/MD steps. Throughout this process, we allow the solvent atoms to move in the potential ($E_{\text{QM/MM}}^{\text{el}}(Q^{(1)}) + E_{\text{vdW}} + E_{\text{MM}}$). Thus, we obtain m snapshots of the solvent coordinates from m MM/MD steps. Finally, at step m , we scale the charge of each solvent atom by $1/m$, and then send $m \times N$ solvent atoms with the scaled solvent charges to the QM program. At this point, it is expedient to divide the solvent into two regions. In the first region (Region I), the N_{ext} solvent charges are converted to $m \times N_{\text{ext}}$ external charges (which are all scaled by $1/m$), whereas Region II represents the average solvent field coming from $N - N_{\text{ext}}$ solvent molecules by two point charges (q and q') using the following relationship:

$$E_{\text{O}} = -\frac{2q}{|r_{\text{OR}}|^3} r_{\text{OR}} \quad (2)$$

where E_{O} is the electric field at point O (the geometrical centre of the QM system) and r_{OR} is a vector pointing along E_{O} to charge q . This approach has been described in great detail in Ref. ^[47] and has successfully been applied to both the evaluation of solvation free energies of small molecules in solution^[47, 60], as well as the evaluation of sidechain pK_{a} s in proteins^[61].

In our QM/MM-FEP calculations, all structures were solvated by a sphere of explicit water molecules with a radius of 16Å, as well as two charges (a dipole) representing the rest of the system. All solvent molecules have been represented by the ENZYMIK force field^[71]. Additionally, in our simulation model, the sphere of explicit water molecules is surrounded by a surface region, the average polarisation and radial distribution of which are determined by the surface-constrained all-atom solvent (SCAAS) model^[52, 69, 72]. This surface region is then embedded in a bulk continuum region with a dielectric, ϵ , of 80 (the appropriate value of ϵ depends on the electrostatic treatment and should be determined by using benchmarks that can, in addition, weed out unreliable treatments of electrostatics^[73]). Finally, the long-range interactions are treated by the local reaction field (LRF) approach^[74]. All MD simulations presented in this work were performed by use the MOLARIS simulation package^[71]. In each case, we first relaxed the system in a 50ps (protein) long simulation, using a 1 fs time step. We subsequently performed classical adiabatic charging FEP calculations in 26 steps of 50 each, in both the forward ($\Delta G_{\text{sol}}(Q = 0 \rightarrow Q_{\text{COSMO}}^0)$) and backward ($\Delta G_{\text{sol}}(Q_{\text{COSMO}}^0 \rightarrow Q = 0)$) directions. The average of the forward and backward charging processes was then used as the appropriate value for $\Delta G_{\text{sol}}(Q = 0 \rightarrow Q_{\text{COSMO}}^0)$ in the cycle shown in Fig. 1. Finally, for each compound, we again first relaxed the system in a 50ps long simulation, using a 1fs time step, and then ran 65ps QM/MM simulations to evaluate the remaining terms of Eq. 1. When performing our QM calculations, we used the mean solvent potential, averaged over 200MD steps (i.e. $m = 200$). All QM calculations were again performed using the G03 software package^[59], the 6-31+G* 5D basis set and the MPW1PW91 hybrid functional^[56]. The Merz-Kollman scheme^[75] with default atomic radii was used to determine charges on atoms to be used later in the MD simulations, and the combined QM/MM calculations were performed using a specially adapted version of MOLARIS (as outlined in Ref.^[47]), with the Gaussian03-MOLARIS communication being facilitated by Perl scripts based on the Gaussian Output Tools package^[76].

When studying phosphate hydrolysis, entropic contributions are clearly very important, particularly when using a mixed solvation model, as ignoring the entropic effect of the explicit water molecules could potentially create a false impression that the presence of these water molecules creates some sort of catalytic advantage for the reaction. The total activation entropy can be broken down into two components: the solvent contribution (termed the solvation entropy, $-T\Delta S_{\text{sol}}$), and the solute contribution (termed the configurational entropy, $-T\Delta S_{\text{conf}}$). The solvation entropy is already implicitly accounted for by the continuum model (though the precise magnitude of its contribution to the total free energy is not quantifiable^[77]). However, the configurational entropy has to be calculated explicitly, which we have done by means of a combination of the quasiharmonic^[78, 79] (QH) and restraint release^[80, 81] (RR) approaches,

which evaluate the configurational entropy using sampling from molecular dynamics runs performed in the presence of explicit water molecules. This approach has been previously successfully applied to not only ribosome^[81], but also (more relevantly to this work), several studies of phosphate hydrolysis in solution^[27-29].

In brief, the restraint release approach imposes strong harmonic Cartesian restraints on the structures of both the reactant (RS) and transition states (TS), and subsequently utilizes a free energy perturbation approach to evaluate the free energy of releasing these restraints (as has been discussed in detail in Ref. ^[80]). The resulting configurational entropy can then be given by Eq. 3:

$$-T\Delta S_{conf}^{\ddagger} = \min(\Delta G_{RR}^{TS}) - \min(\Delta G_{RR}^{RS}) + T\Delta S_{cage}^{RR} \quad (3)$$

An issue that cannot be overlooked when evaluating RR entropies is the fact that by definition, all RR free energies contain enthalpic contributions^[80]. However, the restraint coordinates that give the lowest RR energy also contain the smallest enthalpic contributions (as was demonstrated in Ref. ^[80]). Thus, the procedure above was repeated several times with different sets of constraints in order to identify the restraint coordinates that minimized ΔG_{RR} as far as possible in both RS and TS (denoted by $\min(\Delta G_{RR})$ in Eq. 3), and thus minimized the enthalpic contribution to $-T\Delta S_{conf}^{\ddagger}$. The final term of Eq. 3, $-T\Delta S_{cage}^{RR}$, is simply the entropy associated with applying a special restraint ($K_{cage} = 0.3 \text{ mol}^{-1} \text{ \AA}^{-2}$) to a pair of reacting atoms in the system in order to keep them at a pre-defined contact distance, thus bringing the reacting fragments from a molar volume into the same solvent cage^[80]. Finally, in order to save computational time, it is expedient to replace Eq. 3 with Eq. 4:

$$-T\Delta S_{conf}^{\ddagger} = -T\Delta S(K = K_1)_{QH} + \min[\Delta G_{RR}(K = K_1 \rightarrow K = 0)] \quad (4)$$

Eq. 4 combines the RR entropy with the entropy computed by the QH approximation (denoted by $-T\Delta S(K = K_1)_{QH}$), where K_1 is the initial value of the restraint. This approach exploits the fact that the QH approximation in general tends to be accurate when restraints are significant, and only starts being problematic as these restraints become small (as this results in a range of very shallow and anharmonic potential energy surfaces). Since it is unfortunately at present not possible to carry out the FEP calculations using an *ab initio* surface, we have instead represented the reaction system by an EVB potential that approximates the *ab initio* surface. All RR calculations were performed on both the RS and the TS, using a 16Å sphere of explicit water molecules that were subjected to the SCAAS boundary conditions^[71]. RR-FEP calculations were then performed while changing the value of K_1 from 10 kcal mol⁻¹ Å⁻¹ to 0.0 kcal mol⁻¹ Å⁻² in 41 mapping steps, each of which were 3ps in length (using a 1fs time step), at 300K. $\min(\Delta G_{RR})$ was then evaluated based on 10 series of runs. All configurational entropies were calculated by use of the MOLARIS software package and the ENZY MIX force field^[71, 82].

Results and Discussion

QM/MM Free Energy Perturbation vs. the COSMO Continuum Model

There have been several computational studies which utilize continuum solvation models in order to study phosphate hydrolysis in solution^[27-31, 40, 42, 83], and it has been demonstrated that despite cause for reservation, at least the COSMO continuum model screens charge-charge interactions effectively, and thus can reproduce consistent and accurate solvation energies (see Supporting Material of Ref. ^[27]). Also, in a recent work, we introduced a protocol for accelerated QM/MM-FEP calculations^[47] which we have successfully applied not only to the study of small molecules in solution^[60, 76], but also to the evaluation of sidechain pK_as in proteins^[61]. Here, we will be comparing the efficacy of the two approaches at reproducing experimental data for phosphate diester hydrolysis.

Fig. 3 shows 1-D energy profiles for the hydrolysis of substituted methyl phenyl phosphate diesters shown in Fig. 1, where X = H (Fig. 3a), 4-cyano (Fig. 3b) and 4-nitro (Fig. 3c). In each case, the energy profile was obtained both by means of the COSMO solvation model (shown in red), and by means of QM/MM free energy perturbation (shown in blue). It should be noted here that the entropy and zero point energy (ZPE) corrections were only added at the transition state, and all energy profiles in this work show energies prior to this correction. From this figure, it can be seen that even though in each case there are some initial differences between the reaction profiles obtained by COSMO and by QM/MM-FEP, as the transition state is reached, the two surfaces become increasingly similar, and in each case, the transition states are quite close to each other in energy. This fact is highlighted in Tables 1 and 2, which show the different contributions to ΔG_{calc} for the hydrolysis of each compound studied in this work based on results obtained from the COSMO solvation model and by QM/MM-FEP respectively. As can be seen, in all cases, not only are the ΔG_{calc} obtained by the two approaches quite similar, but also, we reproduced ΔG_{exp} with chemical accuracy by means of QM/MM-FEP. Finally, Table 3 shows phosphorus – oxygen distances to the nucleophile (P–O_{nuc}) and leaving group (P–O_{lg}) respectively. Here, as the leaving group pK_a is increased, the P–O_{nuc} distance is slightly increased by 0.1Å, which is consistent with our previous findings^[27]. However, this increase is accompanied by a slight decrease in the distance of P–O_{lg}. It should be noted that whilst P–O distances differ by 0.1Å for the two approaches for the same transition state, such a minor deviation is not unexpected when changing the solvation model and the overall trend between the three compounds remains consistent. More significantly, we have demonstrated that our QM/MM-FEP approach can be used to reliably reproduce experimental data for phosphate hydrolysis with chemical accuracy.

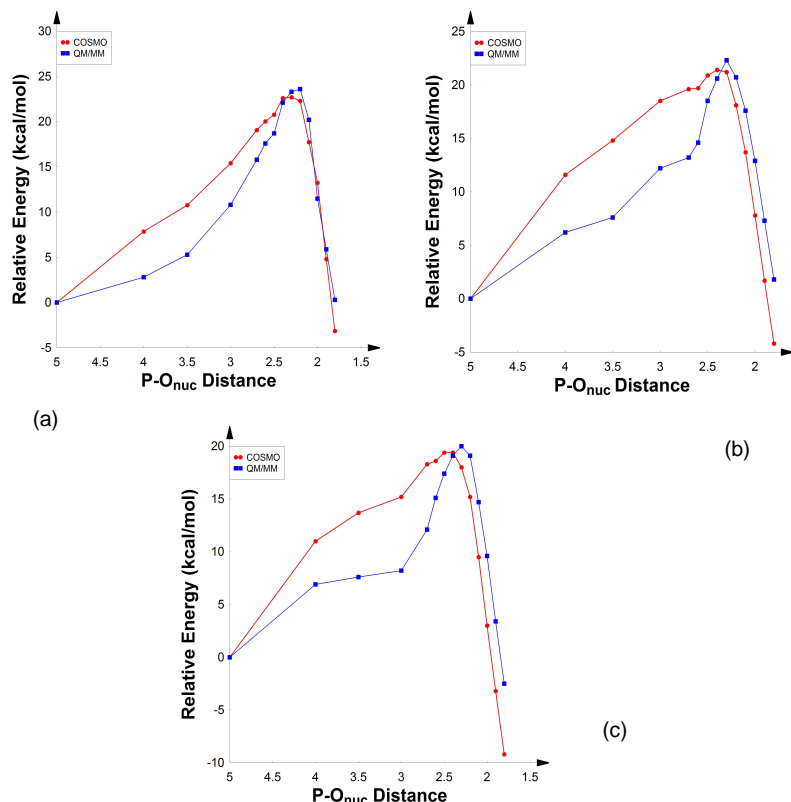


Figure 3. A comparison of 1-D energy profiles obtained by COSMO (red) and by QM/MM-FEP (blue) for the hydrolysis of the methyl phenyl phosphate diesters shown in Fig. 1, where (a) X=H, (b) X=4-cyano and (c) X=4-nitro. All distances are in Å and all energies are in kcal/mol.

X	pK _a	ΔE_{gas}	ΔE_{pol}	$\Delta\Delta G_{\text{solv}}$	$-T\Delta S_{\text{conf}}$	ZPE	ΔG_{calc}	ΔG_{exp}
Parent	9.95	63.3	4.3	-44.8	4.5	0.4	27.7	28.6
4-cyano	7.95	50.3	2.5	-32.3	2.6	1.0	24.1	26.1
4-nitro	7.14	46.4	1.7	-29.7	4.6	1.2	24.2	25.7

[a] ΔE_{gas} denotes the difference in energy upon going from RS to TS in the gas phase, ΔE_{pol} and $\Delta\Delta G_{\text{solv}}$ denote the polarization and solvation energies respectively, $-T\Delta S_{\text{conf}}$ denotes the configurational entropy and ZPE denotes the zero point energy. All values are given in kcal/mol. Experimental values and leaving group pK_as were obtained from Ref. [55].

X	pK _a	ΔE_{gas}	ΔE_{cav}	QM/MM	ΔG_{AC}	$-T\Delta S_{\text{conf}}$	ZPE	ΔG_{calc}	ΔG_{exp}
Parent	9.95	59.7	-0.9	3.0	-38.2	4.5	0.4	28.5	28.6
4-cyano	7.95	46.5	-0.9	-0.5	-22.8	2.6	1.0	25.9	26.1
4-nitro	7.14	41.7	-0.9	-1.4	-19.4	4.6	1.2	25.8	25.7

[b] ΔE_{gas} denotes the difference in energy upon going from RS to TS in the gas phase, ΔE_{cav} denotes the cavitation energy, "QM/MM" denotes the contribution from the LRA terms of Eq. 1, ΔG_{AC} denote the solvation from a classical AC calculation, $-T\Delta S_{\text{conf}}$ denotes the configurational entropy and ZPE denotes the zero point energy. All values are given in kcal/mol. Experimental values and leaving group pK_as were obtained from Ref. [55].

X	pK _a	COSMO		QM/MM-FEP	
		P-O _{nuc}	P-O _{lg}	P-O _{nuc}	P-O _{lg}
Parent	9.95	2.30	2.12	2.20	2.23
4-cyano	7.95	2.40	2.03	2.30	2.13
4-nitro	7.14	2.40	2.01	2.30	2.12

[c] All distances are given in Å. Leaving group pK_as were obtained from Ref. [55].

Fig. 3 shows 1-D energy profiles for the hydrolysis of substituted methyl phenyl phosphate diesters shown in Fig. 1, where X = H (Fig. 3a), 4-cyano (Fig. 3b) and 4-nitro (Fig. 3c). In each case, the energy profile was obtained both by means of the COSMO solvation model (shown in red), and by means of QM/MM free energy perturbation (shown in blue). It should be noted here that the entropy and zero point energy (ZPE) corrections were only added at the transition state, and all energy profiles in this work show energies prior to this correction. From this figure, it can be seen that even though in each case there are some initial differences between the reaction profiles obtained by COSMO and by QM/MM-FEP, as the transition state is reached, the two surfaces become increasingly similar, and in each case, the transition states are quite close to each other in energy. This fact is highlighted in Tables 1 and 2, which show the different contributions to ΔG_{calc} for the hydrolysis of each compound studied in this work based on results obtained from the COSMO solvation model and by QM/MM-FEP respectively. As can be seen, in all cases, not only are the ΔG_{calc} obtained by the two approaches quite similar, but also, we reproduced ΔG_{exp} with chemical accuracy by means of QM/MM-FEP. Finally, Table 3 shows phosphorus – oxygen distances to the nucleophile (P–O_{nuc}) and leaving group (P–O_{lg}) respectively. Here, as the leaving group pK_a is increased, the P–O_{nuc} distance is slightly increased by 0.1 Å, which is consistent with our previous findings^[27]. However, this increase is accompanied by a slight decrease in the distance of P–O_{lg}. It should be noted that whilst P–O distances differ by 0.1 Å for the two approaches for the same transition state, such a minor deviation is not unexpected when changing the solvation model and the overall trend between the three compounds remains consistent. More significantly, we have demonstrated that our QM/MM-FEP approach can be used to reliably reproduce experimental data for phosphate hydrolysis with chemical accuracy.

3.2 Mixed Implicit/Explicit Solvation Models

Having verified that both implicit (COSMO) and explicit (QM/MM-FEP) solvation models give reliable results for phosphate hydrolysis, it becomes important to test whether mixed implicit/explicit solvation models in which a few water molecules are defined explicitly with the remainder of the solvent being modeled as a continuum can give reliable results, and, if this is the case, whether these results are a significant improvement over using a continuum model to which no explicit water molecules have been added. To this end, we modeled the hydrolysis of the parent compound (X=H) in the presence of N explicit water molecules (where N = 0 – 3). For each system, we initially solvated our gas-phase structures in a box of explicit water molecules, ran short QM/MM simulations in which the solute was kept fixed, and for each point on the 1-D surface carefully selected the appropriate number of explicit water molecules before finally re-solvating the complex of the ligand and the selected water molecules using COSMO (thus creating a mixed continuum model).

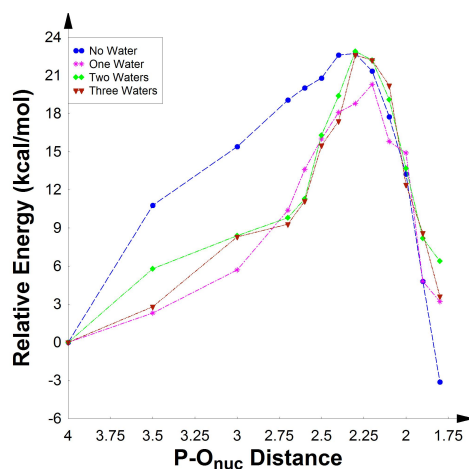


Figure 4. A comparison of the 1-D energy profiles for the hydrolysis of the parent methyl phenyl diester (X=H) shown in Fig. 1 in the presence of N explicit water molecules, where N = 0 to 3. All distances are in Å and all energies are in kcal/mol.

Table 4. The contributions to ΔG_{calc} for the hydrolysis of the parent methyl phenyl phosphate diester studied in this work, obtained from the COSMO solvation model.							
N	ΔE_{gas}	ΔE_{pot}	$\Delta \Delta G_{\text{solv}}$	$-T\Delta S_{\text{conf}}$	ZPE	ΔG_{calc}	ΔG_{exp}
0	63.3	4.3	-44.8	4.5	0.4	27.7	28.6
1	52.4	3.2	-36.4	8.8	0.5	28.5	
2	60.9	1.2	-39.8	4.5	2.3	29.1	
3	55.4	0.4	-33.2	10.3	2.7	35.6	

[d] N denotes the number of explicit water molecules in the simulation system, ΔE_{gas} denotes the difference in energy upon going from RS to TS in the gas phase, ΔE_{pot} and $\Delta \Delta G_{\text{solv}}$ denote the polarization and solvation energies respectively, $-T\Delta S_{\text{conf}}$ denotes the configurational entropy and ZPE denotes the zero point energy. All values are given in kcal/mol. The experimental value was obtained from Ref. [55].

Fig. 4 shows a comparison of the energy profiles obtained in each case, and the different contributions to ΔG_{calc} for each transition state are shown in Table 4. It should be reminded here that the energy profiles of Fig. 4 depict the energy before corrections for the entropy and for the zero-point energy, which were done only at the transition state. Prior to these corrections, we obtain barriers of 22.8 (N=0), 19.2 (N=1), 22.3 (N=2) and 22.6 (N=3). Thus, it would appear at first glance that the presence of the explicit water molecules either has little effect (N= 2 or 3) or even slightly lowers the barrier (N=1) relative to using only a continuum model with no explicit water molecules. However, as was discussed in the introduction, once adding explicit water molecules into the system, it is essential to take the entropy of adding these water molecules into account (which amongst other things can be seen by the fact that the obtained barrier is otherwise far lower than the experimental barrier), something that has not been done by existing studies of phosphate hydrolysis using mixed solvation models (see e.g. Refs. ^[40, 48, 49]).

Once we add the correction for the activation entropy, it can be seen that the presence of each additional water molecule actually slightly increases the barrier by up to 8 kcal/mol (N=3) relative to not using any explicit water molecules (see Table 4 for the energy contributions). Despite this, from Table 4 it can be seen, we still obtain reasonable agreement between ΔG_{calc} and ΔG_{exp} as long as we do not add more than 2 explicit water molecules to the system (even though, as was seen from Table 1, quite high accuracy can already be obtained by using only a continuum without explicit water molecules thus making the explicit water molecules unnecessary). However, an important issue is the conformation of the explicit water molecule(s): so far, all data presented has corresponded to systems in which each explicit water molecule was carefully selected from short QM/MM runs. Whilst this appears to give reasonable results when $N < 2$, there is no guarantee that this is also the case when the orientation of the explicit water molecule(s) is chosen at random, as was done in Refs. ^[40, 48, 49]. Therefore, for comparison, we have also generated the 1-D energy profile for hydroxide attack on the parent diester in the presence of one explicit water molecule, where the conformation of this water molecule was chosen at random in the starting structure ($P-O_{\text{nuc}} = 3.5\text{\AA}$), and the remainder of the energy profile was generated by reaction coordinate pushing from this structure.

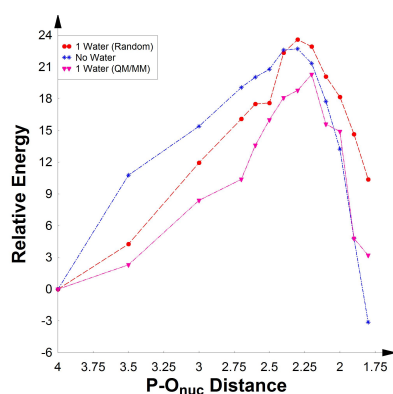


Figure 5. A comparison of the 1-D energy profiles for the hydrolysis of the parent methyl phenyl diester (X=H) shown in Fig. 1 in the presence of no explicit water molecules (red), one explicit water chosen carefully from QM/MM (pink) or one explicit water molecule chosen at random (blue). All distances are in Å and all energies are in kcal/mol.

Fig. 5 shows a comparison of the energy profiles in the presence of no explicit water molecules (blue), one explicit water molecule carefully selected from QM/MM runs (red), and one water molecule chosen at random (pink). As was already seen in Fig. 1, the energy profiles obtained for the hydrolysis of the methyl phenyl phosphate diester from both a “pure” continuum model and from a mixed solvation model with one explicit water molecule are very similar around the transition state, provided that the conformation of the explicit water molecule has been carefully selected from a QM/MM run. From Fig. 5 it can be seen that if the conformation of the explicit water molecule has not been carefully selected but rather allowed to optimize at random, this results to a significant increase of energy compared to selecting the water from a QM/MM run, even before correcting for ZPE and entropy. The energy breakdown for the three cases illustrated in Fig. 5 is given in Table 5. From this table, it can be seen that when the conformation of the explicit water molecule is not carefully selected from a short QM/MM run, the obtained value of 34.7 kcal/mol for ΔG_{calc} is 6 kcal/mol higher than the experimental value. This is consistent with the fact that to the best of our knowledge, there exist no computational studies where the presence of the explicit water molecules in the mixed solvation model has reduced the barrier compared to that obtained from the continuum (when also taking the entropy of the explicit water molecules into account).

Table 5. The contributions to ΔG_{calc} for the hydrolysis of the parent methyl phenyl phosphate diester studied in this work, obtained from the COSMO solvation model.

N	ΔE_{gas}	ΔE_{pol}	$\Delta \Delta G_{\text{solv}}$	$-T\Delta S_{\text{conf}}$	ZPE	ΔG_{calc}	ΔG_{exp}
0	63.3	4.3	-44.8	4.5	0.4	27.7	28.6
1 (from QM)	52.4	3.2	-36.4	8.8	0.5	28.5	

1 (random)	63.8	1.8	-42.1	8.8	2.4	34.7
------------	------	-----	-------	-----	-----	------

[e] N denotes the number of explicit water molecules in the simulation system, ΔE_{gas} denotes the difference in energy upon going from RS to TS in the gas phase, ΔE_{pol} and $\Delta \Delta G_{\text{solv}}$ denote the polarization and solvation energies respectively, $-T\Delta S_{\text{conf}}$ denotes the configurational entropy and ZPE denotes the zero point energy. All values are given in kcal/mol. The experimental value was obtained from Ref. [55].

Conclusion

Phosphate ester hydrolysis is ubiquitous in biology, and is central to the regulation of a wide range of enzymatic reactions with therapeutic relevance. However, prior to undertaking a study of enzymatic phosphate hydrolysis, it is essential to have a clear understanding of the relevant reference reaction in solution, and, as such, it is important to have a relevant computational model for studying phosphate hydrolysis in solution. In recent years, several approaches have become popular for studying phosphate hydrolysis in solution, including the use of mixed implicit/explicit solvation models where solvation is simulated by the addition of one or more explicit water molecules to a continuum (e.g. Refs^[40, 48, 49]).

In this work, we have performed a comparative study of the hydrolysis of three substituted methyl phenyl phosphate phosphate diesters (Fig. 1), in which solvation has been simulated by either a continuum model (COSMO), full explicit solvation (QM/MM-FEP) or a mixed continuum model in which we have added between one and three explicit water molecules to our continuum model. In all cases, hydrolysis proceeds *via* a concerted $A_N D_N$ pathway, with concomitant bond formation to the nucleophile and bond cleavage to the leaving group, as was demonstrated in an earlier computational study^[27].

Here, we have successfully reproduced the experimental barrier with high accuracy by means of both the COSMO solvation model and by means of QM/MM-FEP (in the latter case, we have obtained chemical accuracy for each hydrolysis reaction studied), demonstrating that both the continuum model and our accelerated QM/MM-FEP approach are reliable for studies of phosphate hydrolysis in solution, and that even the continuum model gives quite reasonable results (i.e. a deviation of no greater than 2 kcal/mol from ΔG_{exp}) for this reaction.

In the case of the mixed solvation models, we find that reasonable results can be obtained *only* if no more than 2 water molecules are added to the system, and if the conformation of these water molecules has been carefully selected from prior QM/MM runs in a system with full explicit water molecules (as there is no guarantee that the explicit water molecules have the correct orientation of the corresponding molecules in an explicit infinite system). Also, it is essential to take into account the entropy associated with the presence of the explicit water molecules, as not doing so may give a result that is deceptively similar to the experimental value but this is not actually the true barrier to the hydrolysis of this reaction and the actual barrier including the entropy will be much higher. Finally, we show that when the conformation of the water molecule is chosen at random (as was done in Refs^[40, 48, 49]), ΔG_{calc} is 6 kcal/mol higher than ΔG_{exp} (even though if one had ignored the entropy it might appear that a reasonable value was obtained for ΔG_{calc} , which is a deceptive result). Thus, it would appear that, as we expected, mixing explicit water molecules into a continuum does not accurately reproduce proper polarization boundary conditions between the solute and the bulk, and this situation is aggravated by increasing the number of water molecules, even though it seems that this problem can be slightly mitigated by carefully selecting the conformation of the explicit water molecules from e.g. short QM/MM runs (thus mimicking the conformation these water molecules would have relative to the true infinite system) and also taking the entropy associated with these water molecules, provided one does not add more than two explicit water molecules to the system. However, one can ask oneself whether this is really a sensible approach, as adding explicit water molecules merely increases the computational time of the simulation relative to using a continuum model with no discernible benefit in terms of accuracy, and also, on the other hand, if one has already gone to the effort of setting up the QM/MM simulation in order to select the explicit water molecules one may as well obtain the entire reaction profile using full explicit water molecules which is more reliable than using either a continuum or a mixed solvation model. Thus we believe that the extra cost in terms of both computational time and manpower (i.e. setting up QM/MM runs and selecting appropriate and consistent water molecules for each point on the reaction profile, which is a non-trivial and highly time consuming endeavor) involved in simulating solvation using a mixed solvation model with any degree of reliability makes this an inefficient and undesirable approach, and that one should only include explicit water molecules in the system if one is specifically interested in the chemistry associated with adding these explicit water molecules (for instance in the case of proton transfer reactions).

Acknowledgements

This work was supported by NIH grant 5U19CA105010 and NSF grant MCB-0342276. All computational work was supported by the University of Southern California High Performance Computing and Communication Centre (HPCC). Maciej Haranczyk is a 2008 Seaborg Fellow at Lawrence Berkeley National Laboratory. This research was supported in part (to M. H.) by the U. S. Department of Energy under contract DE-AC02-05CH11231.

- [1] A. S. Mildvan, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1979**, 49, 103.
- [2] F. H. Westheimer, *Chem. Rev.* **1981**, 81, 313.
- [3] I. R. Vetter, A. Wittinghofer, *Q. Rev. Biophys.* **1999**, 32, 1.
- [4] W. W. Cleland, A. C. Hengge, *Chem. Rev.* **2006**, 106, 3252.

- [5] S. J. Benkovic, K. J. Schray, *Chemical basis of biological phosphoryl transfer*, Academic Press, New York, **1973**.
- [6] J. R. Cox, Jr., O. B. Ramsey, *Chem. Rev.* **1964**, *64*, 317.
- [7] J. A. Kirby, S. G. Warren, *The organic chemistry of phosphorous.*, Elsevier, Amsterdam, **1967**.
- [8] R. H. van Huijsduijnen, A. Bombrun, D. Swinnen, *Drug Discov. Today* **2002**, *7*, 1013.
- [9] D. Easty, W. Gallagher, D. C. Bennet, *Curr. Cancer Drug Targets* **2006**, *6*, 519.
- [10] A. Mora, G. Sabio, F. Centeno, *Current Medicinal Chemistry* **2003**, *3*, 335.
- [11] Drug and Market Development Publishing, *Phosphatases: Emerging role in signal transduction, human therapeutics and as drug targets*, **2004**.
- [12] F. Aguilar-Perez, P. Gomez-Tagle, E. Collado-Fregoso, A. K. Yatsimirsky, *Inorg. Chem.*, **2006**, *45*, 9502–9517.
- [13] S. Parimala, M. Kandaswamy, *Transition Met. Chem.*, **2004**, *29*, 35.
- [14] M. Padovani, N. H. Williams, P. Wyman, *J. Phys. Org. Chem* **2004**, *17*, 472–477.
- [15] F. Verge, C. Lebrun, M. Fontecave, S. Menage, *Inorg. Chem.* **2003**, *42*, 499.
- [16] C. Vichard, T. A. Kaden, *Inorg. Chim. Acta* **2004**, *357*, 2285–2293.
- [17] T. Humphrey, M. Forconi, N. H. Williams, A. C. Hengge, *J. Am. Chem. Soc.* **2004**, *126*, 11864.
- [18] J. A. Kirby, W. P. Jencks, *J. Am. Chem. Soc.* **1965**, *87*, 3209.
- [19] R. H. Hoff, A. Hengge, *J. Org. Chem.* **1998**, *63*, 6680.
- [20] J. P. Guthrie, *J. Am. Chem. Soc.* **1977**, *1977*, 99.
- [21] J. A. Barnes, J. Wilkie, I. H. Williams, *J. Chem. Soc. Faraday Trans.* **1994**, *90*, 1709.
- [22] Y. N. Wang, I. A. Topol, J. R. Collins, S. K. Burt, *J. Am. Chem. Soc.* **2003**, *125*, 13265.
- [23] J. Florian, J. Aqvist, A. Warshel, *J. Am. Chem. Soc.* **1998**, *120*, 11524.
- [24] J. Florian, A. Warshel, *J. Am. Chem. Soc.* **1997**, *119*, 4458.
- [25] J. Florian, A. Warshel, *J. Phys. Chem. B* **1998**, *102*, 719.
- [26] M. Klahn, E. Rosta, A. Warshel, *J. Am. Chem. Soc.* **2006**, *128*, 15310.
- [27] E. Rosta, S. C. L. Kamerlin, A. Warshel, *Biochemistry* **2008**, *47*, 3725.
- [28] S. C. L. Kamerlin, J. Florian, A. Warshel, *Chem. Phys. Chem.* **2008**, *9*, 1767.
- [29] S. C. L. Kamerlin, N. H. Williams, A. Warshel, *J. Org. Chem.* **2008**, *73*, 6960.
- [30] N. Iché-Tarrat, J.-C. Barthelat, D. Rinaldi, A. Vigroux, *J. Phys. Chem. B.* **2005**, *109*, 22570.
- [31] N. Iché-Tarrat, M. Ruiz-Lopez, J.-C. Barthelat, A. Vigroux, *Chem. Eur. J.* **2007**, *13*, 3617.
- [32] A. Dejaegere, X. Liang, M. Karplus, *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 1763.
- [33] S. J. Admiraal, D. Herschlag, *Chemistry and Biology* **1995**, *2*, 729.
- [34] S. J. Admiraal, D. Herschlag, *J. Am. Chem. Soc.* **2000**, *122*, 2145.
- [35] J. M. Friedman, S. Freeman, J. R. Knowles, *J. Am. Chem. Soc.* **1988**, *110*, 1268.
- [36] J. A. Kirby, A. G. Varvoglis, *J. Am. Chem. Soc.* **1967**, *89*, 415.
- [37] J. Aqvist, K. Kolmodin, J. Florian, A. Warshel, *Chemistry and Biology* **1999**, *6*, R71.
- [38] A. Hengge, *Adv. Phys. Org. Chem.* **2005**, *40*, 49.
- [39] J. M. Mercero, P. Barrett, C. W. Lam, J. E. Fowler, J. M. Ugalde, L. G. Pedersen, *J. Comput. Chem.* **2000**, *21*, 43.
- [40] Y. Liu, A. Gregersen, A. Hengge, D. M. York, *Biochemistry* **2006**, *45*, 10043.
- [41] Y. Liu, B. A. Gregersen, X. Lopez, D. M. York, *J. Phys. Chem. B.* **2006**, *109*, 19987.
- [42] X. Lopez, A. Dejaegere, F. Leclerc, D. M. York, M. Karplus, *J. Phys. Chem. B.* **2006**, *110*, 11525.
- [43] D.-M. Zhou, K. Taira, *Chem. Rev.* **1998**, *98*, 991.
- [44] M. A. Bochove, F. M. Bickelhaupt, *Eur. J. Org. Chem.* **2008**, 649.
- [45] M. A. Bochove, M. Swart, F. M. Bickelhaupt, *J. Am. Chem. Soc.* **2006**, *128*, 10738.
- [46] M. A. Bochove, M. Swart, F. M. Bickelhaupt, *ChemPhysChem* **2008**, *8*, 2452.
- [47] E. Rosta, M. Haranczyk, Z. T. Chu, A. Warshel, *J. Phys. Chem. B.* **2008**, *112*, 5680.
- [48] L. Zhang, D. Xie, D. Xu, H. Guo, *Chem. Chommun.* **2007**, 1638.
- [49] C.-H. Hu, T. Brinck, *J. Phys. Chem. A.* **1999**, *103*, 5379.
- [50] S. C. L. Kamerlin, J. Wilkie, *Org. Biomol. Chem.* **2007**, *5*, 2098.
- [51] A. Warshel, *J. Phys. Chem.* **1979**, *83*, 1640.
- [52] G. King, A. Warshel, *J. Chem. Phys.* **1989**, *91*, 3647.
- [53] A. Warshel, P. K. Sharma, M. Kato, W. Parson, *BBA* **2006**, *1764*, 1647.
- [54] W. Im, B. Roux, *J. Chem. Phys.* **2001**, *115*, 4850.
- [55] J. G. Zalatan, D. Herschlag, *J. Am. Chem. Soc.* **2006**, *128*, 1293.
- [56] C. Adamo, V. Barone, *J. Chem. Phys.* **1998**, *108*, 664.
- [57] V. Barone, M. Cossi, *J. Phys. Chem. A* **1998**, *102*, 1995.
- [58] A. Klamt, G. J. Schüürmann, *J. Chem. Soc., Perkin Trans. 2* **1993**, *5*, 799.
- [59] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. Montgomery, J. A., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, M. Nakatsuji, M. Hada, K. Ehara, R. Toyota, J. Fukuda, M. Hasegawa, T. Ishida, Y. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adacmo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, Ö. Farkas, D. K. Malick, A. D. Rabuck, K. Clifford, J. Cioslowki, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, *GAUSSIAN 03 (Revision C.02)*, **2004**.
- [60] M. Haranczyk, M. Gutowski, A. Warshel, *Phys. Chem. Chem. Phys.* **2008**, *10*, 4442.
- [61] S. C. L. Kamerlin, M. Haranczyk, A. Warshel, *J. Phys. Chem. B. In Press*.
- [62] J. Tomasi, B. Mennucci, M. T. Cancès, *J. Mol. Struct. (Theochem)* **1991**, *464*, 464.
- [63] M. T. Cancès, B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 3032.
- [64] B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *106*, 5151.
- [65] B. Mennucci, M. T. Cancès, J. Tomasi, *J. Phys. Chem. B* **1997**, *101*, 10506.
- [66] M. Cossi, G. Scalmani, N. Rega, V. Barone, *J. Chem. Phys.* **2002**, *117*, 43.
- [67] J. Florián, A. Warshel, *J. Phys. Chem. B.* **1999**, *103*, 10282.
- [68] A. Warshel, *Computer modeling of chemical reactions in enzymes and solutions*, John Wiley and Sons, New York, **1991**.
- [69] A. Warshel, F. Sussman, G. King, *Biochemistry* **1986**, *25*, 8368.
- [70] J. Florián, A. Warshel, *J. Phys. Chem.* **1997**, *101*, 5538.

- [71] F. S. Lee, Z. T. Chu, A. Warshel, *J. Comp. Chem.* **1993**, *14*, 161.
- [72] A. Warshel, S. T. Russel, *Q. Rev. Biophys.* **1984**, *17*, 283.
- [73] C. N. Schutz, A. Warshel, *PROTEINS: Structure, Function, Genetics* **2001**, *44*, 400.
- [74] F. S. Lee, A. Warshel, *J. Chem. Phys.* **1992**, *97*, 3100.
- [75] B. H. Besler, K. M. Merz, P. A. Kollman, *J. Comput. Chem.* **1990**, *11*, 431.
- [76] M. Haranczyk, M. Gutowski, *J. Chem. Inf. Model.* **2007**, *47*, 686.
- [77] J. Borden, D. C. Crans, J. Florián, *J. Phys. Chem. B.* **2006**, *110*, 14988.
- [78] R. Levy, M. Karplus, J. Kuschick, D. Perahia, *Macromolecules* **1984**, *17*, 1370.
- [79] M. Karplus, J. N. Kuschick, *Macromolecules* **1981**, *14*, 325.
- [80] M. Strajbl, Y. Y. Sham, J. Villa, Z. T. Chu, A. Warshel, *J. Phys. Chem. B.* **2000**, *104*, 4578.
- [81] P. K. Sharma, Y. Xiang, M. Kato, A. Warshel, *Biochemistry* **2005**, *44*, 11307.