PROTEINS:
Structure, Function, and Bioinformatics

Examining the Promiscuous Phosphatase Activity of Pseudomonas aeruginosa Arylsulfatase: A Comparison to Analogous Phosphatases

Journal:	*PROTEINS: Structure, Function, and Bioinformatics*
Manuscript ID:	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Luo, Jinghui; Leiden University, Biophysical and Structural Chemistry; Uppsala University, Cell and Molecular Biology (ICM) van Loo, Bert; University of Cambridge, Biochemistry Kamerlin, Shina; Uppsala University, Department of Cell and Molecular Biology (ICM)
Key Words:	Enzyme catalysis, Catalytic promiscuity, Phosphoryl transfer, Empirical valence bond, substrate-assistend catalysis, protein evolution

Examining the Promiscuous Phosphatase Activity of *Pseudomonas aeruginosa* Arylsulfatase: A Comparison to Analogous Phosphatases

Jinghui Luo[1,2], Bert van Loo[3] and S. C. L. Kamerlin[1,*]

1. Department of Cell and Molecular Biology (ICM), Uppsala University, Uppsala Biomedical Center (BMC), Box 596, S-75124 Uppsala, Sweden.
2. Department of Biophysical and Structural Chemistry, Leiden Insitute of Chemistry, Leiden University, 2300 RA Leiden, The Netherlands.
3. Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, CB2 1GA Cambridge, United Kingdom

Email: kamerlin@icm.uu.se

Telephone: +46 (18) 471 4423

Fax: +46 (18) 471 2000

Short Title: Phosphate hydrolysis by a native sulfatase.

Keywords: Enzyme Catalysis • Catalytic Promiscuity • Phosphoryl Transfer • Empirical Valence Bond • Substrate-Assisted Catalysis • Protein Evolution

ABSTRACT

Pseudomonas aeruginosa arylsulfatase (PAS) is a bacterial sulfatase capable of hydrolyzing a range of sulfate esters. Recently, it has been demonstrated to also show very high proficiency for phosphate ester hydrolysis. Such proficient catalytic promiscuity is significant, as promiscuity has been suggested to play an important role in enzyme evolution. Additionally, a comparative study of the hydrolyses of the *p*-nitrophenyl phosphate and sulfate monoesters in aqueous solution has demonstrated that despite superficial similarities, the two reactions proceed through markedly different transition states with very different solvation effects, indicating that the requirements for the efficient catalysis of the two reactions by an enzyme will also be very different (and yet they are both catalyzed by the same active site). The present work explores the promiscuous phosphomonoesterase activity of PAS. Specifically, we have investigated the identity of the most likely base for the initial activation of the unusual formylglycine hydrate nucleophile (which is common to many sulfatases), and demonstrate that a concerted substrate-as-base mechanism is fully consistent with the experimentally observed data. This is very similar to other related systems, and suggests that, as far as the phosphomonoesterase activity of PAS is concerned, the sulfatase behaves like a "classical" phosphatase, despite the fact that such a mechanism is unlikely to be available to the native substrate (based on pK_a considerations and studies of model systems). Understanding such catalytic versatility can be used to design novel artificial enzymes that are far more proficient than the current generation of designer enzymes.

I. INTRODUCTION

The traditional paradigm for enzyme catalysis, dating back to Emil Fischer's lock-and-key model, has been that enzymes are highly specific, with "one enzyme-one substrate"[1]. However, it is becoming increasingly clear that many enzymes are promiscuous catalysts, with one enzyme being able to facilitate the turnover of a variety of different substrates, thus catalyzing a range of chemically diverse reactions (i.e. substrates with different chemical functionality and requirements for efficient catalysis; for recent reviews, see e.g. [2-5], amongst others). This phenomenon has been referred to as "catalytic promiscuity"[6], and it has been proposed that such promiscuity is important in steering enzyme evolution[2,6,7]. Specifically, in 1976, Jensen[7] hypothesized that, unlike modern enzymes, which were at the time thought to be extraordinarily specific, primordial "ancestor" enzymes were generalists, with a very broad range of activities. This would allow them to react with a wide variety of related substrates, which would be advantageous in that it would maximize the catalytic versatility of primitive cells (that had to out of necessity function with limited enzyme resources). However, ancestor enzymes were then subjected to evolutionary pressure, which allowed them to specialize, resulting in a progressive alteration from broad to narrow specificity with time. This process is shown schematically in Fig. 1. This figure, which is adapted from [2], outlines Jensen's hypothesis, and illustrates how ancestor proteins, which exhibit a range of generalized activities (a, b, c, d) have been subjected to evolutionary pressure to select for specific activities. This is generally thought to proceed *via* a process of gene duplication[8-10], during the course of which the enzyme evolves and gives rise to highly potent and specialized "offspring" (A, B, …). These enzymes may also, to some extent, retain low-levels of their original promiscuous activities (a, b, …). Additionally, as the second gene copy is released from selective pressure, this can in turn accumulate mutations that eventually result in the emergence of previously unseen activities (E, F). However, even these novel offspring can also maintain some level of

promiscuous activity, and indeed it is possible to assume that many (if not most) enzymes are to some degree promiscuous[2].

On the one hand, gene duplication is a slow process (with an estimated time scale of millions of years in eukaryotic organisms)[8], and enzymes have had billions of years to evolve and fine-tune their exquisite catalytic activity. On the other hand, carefully selected mutations (in, for instance, directed evolution or protein redesign studies) can allow this process to be accelerated *in vitro*, and allow for the depression or enhancement of different activities (as well as the evolution of new activities) as desired (see e.g. discussion in [2]). Therefore, a long-term goal would be to be able to reach a stage where enzyme evolution can be accelerated to such a level as to be able to rapidly manipulate enzyme activity in the laboratory (based on insights from both *in vitro* and *in silico* studies). However, in order to be able to engage in efficient protein redesign, it is essential to first understand the molecular basis for enzyme catalysis, and, in fact, being able to design (or redesign) highly efficient artificial enzymes would perhaps be the best proof that we have finally elucidated the rationale for the high proficiency of enzymatic catalysis. It was suggested as early as 1978 that this can be accounted for by the electrostatic preorganization of the active site[11], a proposal that has been reinforced by countless simulation studies (for further discussion, see [12,13] and references cited therein). Therefore perhaps it is unsurprising that the most successful (re)design strategy would appear to be the rational improvement of active site preorganization[14,15].

Following on from the general concept of catalytic promiscuity, it is additionally becoming increasingly clear that many enzymes within a superfamily are capable of cross-catalyzing each other's natural substrates[2,3,5,6], a fact that is particularly observed within members of the alkaline phosphatase superfamily[3,16], which can often readily cross-catalyze phophoryl, sulfuryl and phosphonyl transfer. The present work focuses on a specific member of the alkaline phosphatase superfamily, namely

Pseudomonas aeruginosa arylsulfatase (PAS)[17,18], a sulfate monoester hydrolyase that has been demonstrated to be able to catalyze phosphoryl transfer from both phosphate mono-[19] and diesters[20] in addition to its native sulfatase activity. Interestingly, while the promiscuous activity of many enzymes tends to be quite inefficient compared to that of the native reaction, PAS is able to provide remarkably high rate accelerations for both its phosphate mono- and diesterase activities, the latter of which becomes comparable with its proficiency for its native reaction[19,20]. This is particularly interesting, in light of the fact that theoretical studies of the prototype reference reactions for the hydrolysis of phosphate monoester dianions and sulfate monoester monoanions suggest that, despite the apparent similarities in the measured rate constants for the uncatalyzed hydrolyses of the two monoesters[21,22] (k_{uncat}(*p*-nitrophenyl phosphate)/k_{uncat}(*p*-nitrophenyl sulfate)≈4[19]), the solvation requirements for the efficient catalysis of phosphoryl and sulfuryl transfer are likely to be quite different[23], in turn requiring active sites with quite different electrostatic preorganization. This makes it surprising that PAS is capable of showing such high proficiency for its promiscuous reactions compared to the native reaction, suggesting that, to some extent, this enzyme maintains a primordial "scavenger" tendency, and understanding how the enzyme manages to achieve such proficient catalysis of a reaction with very different catalytic requirements is a question of major importance.

Herein, we take the first step in this direction by focusing on the promiscuous phosphate monoesterase activity of PAS, examining the mechanism by which this enzyme hydrolyzes the *p*-nitrophenyl phosphate dianion (see discussion in [19]). We demonstrate that despite being a sulfatase, PAS behaves much like other analogous native phosphatases, with the reaction most likely proceeding *via* a substrate-as-base mechanism, akin to that observed in systems such as the RAS GTPase[24,25] and elongation factor Tu (EF-Tu)[26,27]. Subsequent work will focus on the native sulfatase reaction, as well as the promiscuous phosphodiesterase activity, in order to understand how PAS is capable of accelerating

such catalytically diverse reactions, and to identify the distinguishing structural and energetic features between the native and promiscuous activities. We believe that a combination of experimental approaches and theoretical tools (that are capable of reproducing all relevant available experimental data) to examine and dissect the catalysis of different substrates by related promiscuous enzymes could hold the key to understanding evolution of function within enzyme superfamilies. This, in turn, would lead to the development of novel systems that come far closer to competing with nature than the current generation of designer enzymes.

II. COMPUTATIONAL BACKGROUND

II.1 Ab Initio Calibration of the Reference Reaction in Solution

Prior to attempting to quantify the catalytic effect of an enzyme, it is essential to also have detailed information on the corresponding background (uncatalyzed) reaction in solution[13,28]. The purpose of this is two-fold. Firstly, in solution, the reaction can proceed through multiple viable mechanisms, whereas the enzyme can restrict the system to follow a specific path. Therefore in order to understand the effect of the enzyme, it is important to first understand the basic chemistry of the reaction in the absence of the catalyst. Secondly, high-level theoretical calculations of the background chemistry in solution are important for the calibration of theoretical approaches used to study the enzyme-catalyzed reaction (though note that for any such comparison to be meaningful, both enzymatic and solution reactions need to be examined at the same level of theory). This is particularly important for reactions involving phosphorus, where there can be multiple equally viable mechanisms in solution, and unambiguously distinguishing between them using experimental probes can be challenging (see e.g. [29-33]). Therefore, in order to calibrate our subsequent EVB calculations of the phosphoryl transfer reaction, we first performed *ab initio* calculations on the corresponding system in solution, using the approach we have

successfully utilized for our previous studies of phosphate and sulfate hydrolysis[23,31,33-36]. Specifically, the reactivity in solution was examined by generating 2-D More O'Ferrall- Jencks plots[37,38], in the space defined by phosphorus-oxygen distances to the incoming nucleophile and the departing leaving group. This approach allows for the direct comparison of all viable mechanisms, if more than one is available to the system. Careful reaction coordinate pushing was then used to generate the surface, mapping bond distances from 1.6 to 3.0Å, in 0.15Å increments. At each point on the plot, the energetics and geometries were examined to ensure that the correct lowest energy free energy surface is being generated for the system. In order to avoid introducing additional bias into the system, no constraints were placed on any protons in the system (though in principle proton transfer could be defined as a separate additional reaction coordinate). This approach is similar to that utilized in previous works where proton transfer has been an issue (e.g. [30,32,33]). However, in order to correctly model the phosphoryl transfer to the geminal diol, it was necessary to place a C-O constraint on the non-nucleophilic oxygen of the geminal diol, to prevent its "decomposition" to the corresponding aldehyde. This is because, as discussed extensively elsewhere (e.g. [13,36]), it is essential to model *exactly the same* mechanism in solution as in the enzyme, even if this is not necessarily the preferred pathway in the absence of the enzyme), and, in this case, the geminal diol is not necessarily the favored species in solution. All *ab initio* calculations were performed using Gaussian03[39], and Becke's three-level hybrid functionals (B3LYP)[40], which utilize a combination of Hartree-Fock exchange and density functional theory. Initial geometries were obtained using the 6-31+G* basis set and the polarizable continuum model (PCM)[41-44] with integral equation formalism, which allows for the calculation of the analytical gradients of the solute solvent surface. However, a subsequent single-point correction was performed on each of the optimized geometries, using the larger 6-311+G** basis set and the COSMO continuum model[45,46]. Finally, entropic contributions were evaluated using the restraint release approach[47], in line

with our previous studies of related systems (e.g. [23,32,34]), and then added on as a correction to the activation barrier. Therefore, the final calculated activation barriers and free energies also take into account the contribution of the configurational (solute) entropy, which can be significant, particularly in solution.

II.2 Empirical Valence Bond Calculations of the Enzymatic Reaction

While the increase in computational power has allowed us reach a stage where it is possible to perform *ab initio* calculations of small systems with high levels of accuracy, we are quite far from being able to treat entire enzymes using only *ab initio* approaches. Alternatives to this have been the use of either small cluster models[48] in which only a small subset of the system is treated quantum mechanically, with the remainder of the system being modeled as a continuum with a user-selected dielectric constant, or, alternately, the use of QM/MM calculations where the relevant subset of interest is treated by means of high-level quantum mechanical calculations, with the remainder of the system being treated by molecular mechanics (for reviews see e.g. [49-51]). However, despite the increasing popularity of such approaches (and their success in some cases), a significant challenge remains due to the need of extensive sampling in order to perform meaningful free energy calculations, which becomes rapidly prohibitively computationally expensive considering the high cost of each QM call when one performs *ab initio* QM/MM. A highly powerful option exists in the form of the empirical valence bond (EVB) approach of Warshel and coworkers[52,53], which is a semi-empirical QM/MM approach that is on the one hand sufficiently computationally robust and carries sufficiently large wealth of chemical information in order to be able to provide a reliable description of bond-breaking and bond-making chemical processes, while at the same time being sufficiently fast to allow for the extensive sampling required for meaningful convergent free energy calculations. This approach, which has been reviewed in detail in e.g. [54], has been repeatedly demonstrated to be a highly powerful tool for studying chemical

reactivity in solution and in enzymes, and is particularly useful if one is interested in obtaining a *quantitative* understanding of the molecular basis for enzyme catalysis (see examples in the reviews mentioned above, and discussion in e.g. [13]). Additionally, the EVB approach has been demonstrated to be a very powerful tool at various stages in computational enzyme design[14,15,55,56]. Since the EVB has already been discussed in such detail elsewhere (e.g. [54]), we will not reiterate the technical details here, and instead direct interested readers to these reviews.

All calculations in this work, with the exception of the *ab initio* calculations discussed in Section II.1, have been performed using the MOLARIS simulation package and the ENZYMIX force field[57-59]. The EVB activation barriers were obtained using the standard free energy perturbation / umbrella sampling (FEP/US) approach outlined in [60]. The initial atomic coordinates for *Pseudomonas aeruginosa* arylsulfatase at 1.3Å resolution[61] were obtained from the Protein Data Bank[62] (accession code 1HDH). The sulfate (SO_4^{2-}) ion bound to the active site was replaced by the *p*-nitrophenyl phosphate dianion, which was docked into the active site using the Hex protein docking package[63]. The resulting complex was then initially relaxed for 30ps at 30K, followed by 100ps at 300K, using a 1fs timestep. The ionization state of relevant titratable groups within ~6-8Å of the bound substrate was determined by means of initial empirical pK_a screening using PROPKA 2.0[64,65], and then taking into account the fact that the pH profile for the phosphomonoesterase activity of PAS shows its maximum activity at a pH of 7.2 or less (see Fig. 7 of [19]). Based on this, the following residues where ionized during the simulations: Asp13, Asp14, Arg55, Glu74, Lys113, Asp317, Glu326 and Lys375, which, when also taking into account the charge of the calcium ion (+2) and the bound substrate (-2) leads to an overall system charge of -2 during the simulations.

For the initial structure relaxation, the simulation system was solvated by a 24Å water sphere, centered on the substrate and surrounded first by a 3Å grid of Langevin dipoles and then by bulk

solvent. Solvent was modeled using the surface constrained all atom solvent model (SCAAS)[57], and long range effects were treated using the local reaction field (LRF) approach[66]. The final orientation of relevant active site residues after 100ps of relaxation at 300K is shown in Fig. 2. For the subsequent EVB FEP/US calculations, a smaller 20Å explicit water sphere was used, which was extended first to 22Å with a grid of Langevin dipoles, and then to infinity with bulk solvent. The FEP mapping was performed in 51 mapping frames of 50ps each in length, giving a total simulation time of 2.55ns. In order to ensure that the system was reliably sampled, all EVB mapping runs were repeated ten times, with different initial conditions, obtained from snapshots taken every 1ps from a 10ps relaxation trajectory after the initial relaxation run, using the smaller sphere size. Finally, the contribution of different residues to the overall activation barrier was evaluated using linear response approximation (LRA), as was also done in e.g. [67].

III. Results and Discussion

III.1 Initial Mechanistic Considerations

Sulfatases tend to have very high sequence identity (~20-60%), with particularly high conservation of the N-terminal third of the polypeptide chain, as well as the active site residues[61]. Interestingly, a common feature to many sulfatases is the presence of a formylglycine (FGly) residue, obtained from the post-translational oxidation of either a cysteine (prokaryotic and eukaryotic sulfatases) or serine (bacterial sulfatases) to an aldehyde, that exist in its hydrated form (geminal diol), in the active site[68-71]. This residue then acts as the nucleophile in the chemical step (see Fig. 3), and mutational studies[19,20] in which the FGly is mutated to either a serine or alanine show that, upon mutation, the activity of PAS towards both its native and promiscuous substrates is significantly depressed. A simplified version of the proposed catalytic mechanism[19] for the promiscuous phosphate monoesterase activity of PAS is shown

in Fig. 3. In this mechanism, the nucleophile is the aforementioned hydrate of a metal coordinated formylglycine (FGly51), which results in the transfer of a phosphoryl group to this residue in order to form a phosphoenzyme intermediate. The second step involves the breaking down of the intermediate, which has been proposed to proceed by cleavage of the C_β-O bond of the hemiacetal, and not the inert S-O[61] (or in this case P-O) bond. The present work focuses on the chemical step, involving the phosphoryl transfer, which is significantly accelerated compared to the corresponding uncatalyzed reaction in solution (see Section III.3). Note that the reactions catalyzed by PAS have been proposed[19,61] to proceed *via* acid-base catalysis, with Asp317 acting as a general base to deprotonate the nucleophile and Lys375 or His211 acting as a general acid to protonate the leaving group during the phosphoryl transfer (step 2→3 of Fig. 3), and His115 acting as the general base during hemiacetal cleavage (step 3→1 of Fig. 3). Therefore, the first issue of relevance would be the identification of the most likely base for the initial deprotonation of the nucleophile. Additionally, as mentioned in Section II.2, the pH rate profiles for the promiscuous phosphatase activity of PAS show that the enzyme has its maximum activity at pH 7.2 or below[19]. Combining this with the fact that *p*-nitrophenol is anyhow a very good leaving group, with a pK_a of ~7, at roughly physiological pH, there is not *necessarily* any impetus for its protonation in order to facilitate bond cleavage (and this was indeed the case in our preliminary examination of the background reactivity in solution[23]).

III.2 Identifying the Correct Base for Deprotonation of the Nucleophile

As illustrated in Fig. 3, an important part of the catalytic mechanism involves the initial deprotonation of FGly51, which brings us to the first question posed in Section III.1, namely what the correct identity of the base responsible for this deprotonation is. From a simple examination of the crystal structure, there appear to be two likely candidates. The first of these is simply the original proposition of e.g. [19,61], namely that a nearby metal-coordinated aspartate (Asp317) acts as the base, and

the second is that the proton is transferred to the phosphate itself in a substrate-as-base mechanism, as is the case in systems such as the RAS GTPase[24,25]. Here, it is important to take pK_a considerations into account. Specifically, while it was not possible to find a precise literature value for the pK_a of the proposed nucleophile in solution, geminal diols tend to have pK_as in a range of 13-14[72], and the pK_a of the closely analogous acetaldehyde hydrate in solution is 13.48[73], so it should be safe to assume that in solution, the pK_a of the nucleophile would be ~13.5. In contrast, the pK_as of aspartic acid and the phosphate in solution are 3.86[73] and 4.79[74] respectively. The free energy of a proton transfer in solution can, to a very good approximation, be obtained simply from the difference in pK_a between the donor and acceptor groups (see e.g. [60,75]), using the following relationship:

$$\Delta G^{w}_{PT} = 2.303RT[pK_a(donor) - pK_a(acceptor)] \quad (1)$$

Based on which the free energies of transferring a proton from FGly to either an Asp or the phosphate itself in solution are 13.1 and 11.9 kcal/mol respectively (note as an aside that, logically, for either alternative to be a viable base, the free energy for the proton transfer cannot exceed the experimentally observed activation free energy for the whole reaction). Therefore, from simply examining the difference in pK_a between the proton donor and potential proton acceptors, it would appear that, in solution, the phosphate itself has a very slight edge over Asp317 as the potential base. However, the situation in the enzyme could potentially be quite different, as PAS has a divalent metal cation (in the case of PAS a Ca^{2+} cation) in the active site which not only positions and activates the FGly for nucleophlic attack, but also most likely reduces its pK_a (though the Ca^{2+} will most likely also lower the pK_a of the metal-coordinated Asp317). There is additionally the issue of whether a metal-coordinated Asp is going to be a good base at all, as one of the purposes of this residue being in that particular position is most likely to simply coordinate the Ca^{2+} and "keep it in place". Additionally, a structural examination of the orientation of the key active site residues after 100ps of MD relaxation (see

Fig. 2) shows that the proton on the proposed nucleophilic oxygen appears to be preferentially hydrogen bonding to one of the oxygen atoms of the phosphate dianion, and seems to be optimally arranged for proton transfer along this hydrogen bond. However, at this stage such arguments are still just speculation, and in order to identify the most likely base, it is important to take proper energy-based considerations into account. Eq. 2 outlined the relationship between the free energy for a proton transfer in solution and the pK_as of the donor and the acceptor groups. The corresponding energetics in a protein active site can be obtained using the following relationship (see discussion in [76]):

$$\Delta G_{PT}^{p} = \Delta G_{PT}^{w} + \Delta\Delta G_{sol}^{w\to p} \quad (2)$$

Where $\Delta\Delta G_{sol}^{w\to p}$ designates the change in the solvation energy upon moving the reacting species, in a given configuration, from solution to the active site of a protein. This relationship is based on the thermodynamic cycle shown in e.g. Fig. 3 of [76], and has been discussed in detail elsewhere[60,75]. The key point here is that even if the proton transfer is seemingly unfavorable in solution, the corresponding energetics in the protein active site depend on the coupled effect of the protein environment, which can in turn either stabilize or destabilize the resulting ion pair. Therefore, following e.g. [76,77] we performed EVB/FEP-US calculations to examine the energetics of a general base mechanism with Asp317 serving as the base, and a substrate-as-base mechanism with the phosphate serving as the base, in both the enzyme and in solution.

The relevant valence bond structures are shown in Fig. 4, where states I and II correspond to the reactant and intermediate structures for the aspartate-as-base mechanism, and state III corresponds to the intermediate structures for the substrate-as-base mechanism in the enzyme. In all cases, the Ca^{2+} ion was included in the reacting part in the enzyme calculations as this appeared to improve the stability of the calculations, however, Asp317 was only included in the EVB reacting part when examining the aspartate-as-base mechanism (both in the enzyme and in solution). The reaction I→II in water was

modeled by proton transfer from FGly51 in its hydrated form to aspartate, and the reaction I→III in water was modeled by proton transfer from FGly51 in its hydrated form to *p*-nitrophenyl phosphate. Additionally, in both cases, OG_1 was taken as the nucleophilic oxygen, as this was the oxygen closest to the phosphorus atom, with the sidechain of FGly51 rotated in such a way that OG_2 was facing away from the phosphate. In the case of the aspartate-as-base mechanism, the proton was transferred to OD_2 of Asp317, and in the case of the substrate-as-base mechanism, the proton was transferred to OP_2 of the phosphate, as these were the oxygen atoms closest to the nucleophilic oxygen atom in the reactant state (for labeling, see Fig. S1). Finally, in all cases (both enzyme and solution), a 2.0 kcal mol^{-1} $Å^{-2}$ distance constraint was applied to the relevant donor and acceptor atoms to keep them within a distance of ~3Å during the course of the simulation, and the EVB parameters used were a pertinent subset of those presented in Tables S1-S8.

Note that here our main focus is on the overall *free energy* of the proton transfer (as estimated from Eq. 2), and not the corresponding activation barrier. This is because evaluating ΔG_0 for the process in solution by means of the EVB approach is fairly straightforward, and involves adjusting the relevant gas-phase shift (α_0) for the process (see e.g. [77]). However, obtaining a reliable value for the activation barrier for such a proton transfer can be quite challenging, not only experimentally, but also theoretically, as gas-phase *ab initio* calculations will typically yield only single minima potential energy surfaces when the overall free energy is > 10 kcal/mol (see footnote in [77]). Additionally, even taking into account the solvent by means of e.g. a continuum approach can be challenging, as the solvent tends to stabilize the charges of the reactant and product states far more than the corresponding transition state, which has a delocalized charge distribution[78]. Finally, for such proton transfer reactions in solution, non-equilibrium solvation effects can make a non-negligible contribution[79] (though such contributions are still likely to be small, as proton transfers between oxygen atoms in solution are only diffusion

limited[80]). Fortunately, however, while the activation barrier is clearly important as well, for our purposes, as long as the proton transfer it not rate-limiting, it is not as directly relevant to the question of what the best base in the enzyme will be for the initial deprotonation of the nucleophile, and this issue can be addressed by examining the relevant ΔG_0 for both potential bases. Therefore, in both cases, the activation barrier for the proton transfer in solution was estimated to be ~2-4 kcal/mol above the overall free energy for the process, following the detailed discussion in [77] for an analogous case. The calculated activation barriers in solution were set to agree with this estimate, by adjusting the H_{12} elements of the EVB Hamiltonian accordingly, and the relevant gas-phase shift and H_{12} elements were then used unchanged to evaluate the energetics of the corresponding proton transfer in the enzyme. The resulting activation and reaction free energies for both proton transfers in enzyme and in solution are shown in Table 1, as well as the corresponding EVB parameters. From this table, it can be seen that in the case of the aspartate-as-base mechanism, the overall free energy for the proton transfer in the enzyme is 4.5 kcal/mol *higher* than the (already high) corresponding free energy in solution, which is not surprising in light of the arguments above against the viability of this residue as a base. In contrast, in the case of the substrate-as-base mechanism, the ion pair is slightly stabilized by 2.2 kcal/mol in the enzyme compared to solution. Note that solely examining the free energy of the initial proton transfer does not necessarily identify the correct base, as it tells us nothing about the activation barrier to the proton transfer, or, even more importantly, the energetics of the subsequent nucleophilic attack. However, while this approach cannot confirm whether something is the preferred base, it can *rule out* unlikely bases, since if the free energy for the proton transfer in the enzyme is almost as high or even higher than the observed activation barrier (as is the case with Asp317, see Table 1), then it is extremely unlikely that this is going to be the base required for nucleophile activation. Therefore, while these calculations alone don't confirm that the phosphate is itself acting as a base, they do illustrate that it is a much more likely

candidate for this role than Asp317, an observation that would also appear to agree with the experimentally measured pH rate profile for *p*-nitrophenyl phosphate hydrolysis by PAS[19], where the slope of $\log(k_{cat}/K_M)$ for the pH profile is greater than 1, leaving open the possibility that this corresponds to the protonation of the phosphate dianion (see discussion in [19]).

III.3 Phosphoryl Tranfer from p-Nitrophenyl Phosphate to Arylsulfatase

Having identified the phosphate as the most likely base in this reaction, the next step is to examine the actual phosphoryl transfer step from *p*-nitrophenyl sulfate to arylsulfatase. In principle, this could occur either *via* a stepwise mechanism involving an intermediate, or *via* a concerted mechanism in which bond cleavage and bond formation occur simultaneously in the transition state. Here, we will be specifically modeling the phosphoryl transfer step as a concerted process, as related studies of phosphate monoester dianions in solution do not suggest the presence of an intermediate of any form (either associative or dissociative), and, in any case, the overall conclusions are independent of whether any penta- or trivalent species formed as the phosphoryl group is in flight is a transition state or a high-energy intermediate[77]. An additional issue is the question of whether the initial deprotonation of the nucleophile occurs in a pre-equilibrium proton-transfer step (as modeled in the previous section), or whether it is concerted to the phosphoryl transfer step. Experimentally, these two mechanistic possibilities are indistinguishable, as they will have identical kinetics[33,81]. There has been some controversy as to whether such a pre-equilibrium proton transfer is actually a viable mechanism for the hydrolysis of phosphate monoesters[82], however, it has been demonstrated that such a mechanism is fully energetically feasible for the case of both phosphate mono- and diesters, not only in solution[33,81], but also in relevant enzymatic reactions[24,25] (even though, despite the overwhelming body of evidence in support of such a mechanism, its relevance is still brought into question[83]). However, as the two possibilities are anyhow kinetically indistinguishable, for simplicity, for the actual phosphoryl transfer

step, we have modeled the system as proceeding *via* a concerted A_ND_N mechanism, in which proton transfer from the nucleophile to the phosphate occurs synchronously to the phosphoryl transfer. It is also worth noting that the *p*-nitrophenyl phosphate *monoanion* is ~70x less susceptible to water attack than the corresponding dianion[84], which would further support the likelihood of "simultaneous" proton transfer. The corresponding valence bond structures used in the EVB calculations are shown in Fig. 5.

Prior to examining the enzyme-catalyzed reaction, however, it is important to have a clear understanding of the corresponding energetics and mechanism in solution. Experimental studies[21,22,85-87] have proposed that this reaction will proceed through a largely dissociative mechanism, with little nucleophile participation, and a largely broken bond to the leaving group in the transition state. However, this has been based on markers such as the measurement of the activation entropy, isotope effect measurements and linear free energy relationships, whereas computational studies have shown that for such complex reactions (i.e. in the case of phosphoryl transfer, for which multiple mechanisms can be equally viable for the same system), such markers need to be approached with care as they cannot uniquely and unambiguously determine a single mechanism[29-33]. Additionally, computational studies[23] of this system which examined the free energy surface for the reaction (and not just isolated transition states) obtained a comparatively compact concerted A_ND_N transition state, while reproducing the experimentally measured activation energy and entropy with reasonable accuracy.

In the case of [23], however, the system examined was *water* attack on *p*-nitrophenyl phosphate, whereas in the present case, the nucleophile is a geminal diol. When examining the *catalytic amplification* of an enzyme (i.e. $\Delta g^{\neq}_{wat} - \Delta g^{\neq}_{enz}$, where $\Delta g^{\neq}_{wat}$ and $\Delta g^{\neq}_{enz}$ are the activation barriers for the reaction in enzyme and in solution respectively), it is important to take the correct reference state into account, the most appropriate reference state being the corresponding uncatalyzed reaction in solution. In doing so, it is important to take into account that the enzymatic reaction and the

experimentally observed reaction in water are not necessarily identical. Examples of this would include the fact that in solution, there can be multiple mechanistic possibilities, whereas the enzyme restricts the reaction in such a way as to follow a specific reaction mechanism (that may even be different from the preferred mechanism in water). Alternately, the reaction in water may proceed with water as a base, while the enzyme can utilize a general base. Or, as in the present case, the uncatalyzed reaction may proceed with a water nucleophile, whereas the enzyme can use an amino acid sidechain as the nucleophile. Such effects are fortunately well understood, and can be considered "chemical" effects. The key issues when examining the catalytic amplification of the reaction, however, are "environmental" effects, which involve comparing the effect of the change in environment upon moving from solution to the enzyme for a reaction that involves *exactly* the same mechanism, binding the same chemical groups, but being conducted in the enzyme instead of water. This complex issue has been discussed in great detail elsewhere (e.g. [13]). For the present case, this would require examining the energetics of the corresponding reaction in solution, using the same nucleophile as in the enzyme-catalyzed reaction. The experimentally measured rate[21] for the aqueous hydrolysis (water attack, k_0) of *p*-nitrophenyl phosphate at 39° C is 1.6 x 10^{-8} s^{-1}, and the rate for the corresponding alkaline hydrolysis (hydroxide attack, k_{OH}) at the same temperature is 3.8x10^{-9} M^{-1} s^{-1} [21]. This corresponds to negligibly different activation free energies of 29.4 and 29.8 kcal/mol respectively from transition state theory. Taking into account the pK_as of water and the hydroxide ion (-1.7 and 15.7 respectively), and extrapolating this to the case where the nucleophile is the geminal diol (pK_a ~ 13.5, see discussion in Section III.2), one would expect a similar activation barrier for the correct reference reaction in solution as for hydroxide attack. *However*, there are two important issues that need to be taken into account: the first is that it is based on a (fairly reliable) assumption for the solution pK_a of the nucleophile, but, more importantly, as the

experimental rate for the alkaline hydrolysis of *p*-nitrophenyl phosphate most likely corresponds to C-O cleavage and not P-O cleavage[21], 8.3×10^{-9} M^{-1} s^{-1} is at best an upper bound for the rate of this reaction.

In order to further explore this issue, we have utilized the same computational approach as in e.g. [23] (outlined in Section II), using acetaldehyde hydrate as a model for the enzymatic nucleophile, and the resulting free energy surface is shown in Fig. 6. As can be seen from this surface, the reaction proceeds over a high-energy ridge separating the reactant and product states, with an approximate transition state being marked on the surface. However, unlike previous studies of the hydrolysis of phosphate dianions, and similarly to the corresponding surface for water attack on the *p*-nitrophenyl phosphate dianion, for this specific substrate, the presence of the proton transfer increases the complexity of the surface such that it is not possible to identify a clear saddle point along the ridge. Therefore, the transition state presented in Fig. 6 is at best approximate (obtained by following the contours along the surface, and by analogy to our previous work[23]), and the purpose of the surface is predominantly to test whether a substrate-as-base mechanism is even viable when changing the nucleophile from water to acetaldehyde hydrate, and, if so, to provide a rough guideline to the energetics (see Table 2), with the experimental data analyzed above providing a much more reliable approximation for the value of $\Delta g^{\neq}_{cage}$.

From this figure and table, it can be seen that once again, the mechanism proceeds *via* a comparatively compact transition state, with P-O distances of 2.05 and 1.9Å in solution, with a free energy surface that is qualitatively similar to that observed for water attack on *p*-nitrophenyl phosphate[23]. The resulting activation free energy of 31.2 kcal/mol is slightly higher than the "experimental" estimate, however, note that this is mostly likely an overestimate by a few kcal/mol, as was also the case in [23] for related systems. Additionally, the correct reference reaction in solution is not the reaction relative to the fragments at infinite separation ($\Delta g^{\neq}_{w}$) but rather the reaction where the

reacting fragments have been brought into the same solvent cage ($\Delta g^{\neq}_{cage}$) (see discussion in [13,77]), which, for biomolecular reactions of 1M reactants in solution are related by the following relationship[60]:

$$\Delta g^{\neq}_{w} \approx \Delta g^{\neq}_{cage} + 2.4 kcal/mol \quad (3)$$

Therefore, we assume $\Delta g^{\neq}_{cage}$ is ~27.0 kcal/mol, based on an approximation of ~29.5 kcal/mol for $\Delta g^{\neq}_{w}$, obtained from comparing the experimentally measured rates for water and hydroxide attack on *p*-nitrophenyl phosphate, as discussed above. Note though that this is still an approximation due to the uncertainty in both the pK_a of the nucleophile in solution, and also the actual relevant reaction rate for the alkaline hydrolysis of *p*-nitrophenyl phosphate. Additionally, the overall ΔG_0 for the reaction was obtained as the difference between the sum energies of the reactant and product species at infinite separation. However, the resulting ΔG_0 is fairly exothermic (-9.8 kcal/mol), which is again likely to be an artifact when dealing with a highly charged system as mentioned above, and therefore our main focus is on correctly reproducing the observed activation barrier rather than the overall endo/exothermicity of the reaction.

The calculated energetics for the phosphoryl transfer step from *p*-nitrophenyl phosphate to the attacking nucleophile in water and in solution (as well as the corresponding reorganization energies) averaged over 10 trajectories are shown in Table 3, and representative free energy profiles in water and in the enzyme are shown in Fig. 7. As discussed above, in modeling the phosphoryl transfer step, the following assumptions were taken: (1) that the best/most likely base is the phosphate itself, and therefore that the nucleophile is deprotonated by proton transfer to the phosphate in a substrate-as-base mechanism, (2) that this proton transfer occurs simultaneously to the phosphoryl transfer step and *not* in a preequilibrium mechanism (for simplicity, as a reminder, the two mechanisms are kinetically identical and modeling this process as a preequilibrium proton transfer will not affect the validity of our results), (3) that the phosphoryl transfer itself occurs *via* a concerted A_ND_N pathway without a clear intermediate

(based on extensive computational examination of the mechanism of the hydrolysis of phosphate monoester dianions in aqueous solution, e.g. [23,30,32], amongst others), and (4) that the leaving group does not require protonation to depart (as was the case for the reaction in solution[23], and is in good agreement with experimentally measured isotope effects for *p*-nitrophenyl phosphate hydrolysis in aqueous solution[86]). The experimentally measured k_{cat} for wild-type PAS at pH 8.0 and 25° C is 0.023 s^{-1} [19], which corresponds to an activation barrier of 19.6 kcal/mol from transition state theory.

As can be seen from Table 3, we obtain a barrier reduction of 7.8 kcal/mol for the enzyme-catalyzed reaction compared to the reaction in solution, to give an overall activation barrier of 19.5 kcal/mol, though this extremely good agreement with experiment is likely to be serendipitous, due to the fact that (1) the precise activation barrier for the uncatalyzed reaction with the geminal diol instead of water as a nucleophile is uncertain and estimated from a combination of *ab initio* calculations and experimentally measured rates for analogous reactions, and (2) the chemical step is not necessarily the rate-limiting step, and therefore, the actual activation barrier for the phosphoryl transfer reaction in the enzyme may be lower than that obtained from examining only the experimentally measured rate constant (i.e. this provides an upper limit). Finally, the overall free energy (ΔG_0) for the reaction in the enzyme is slightly more exothermic than in solution, giving a ΔG_0 of -12.7 kcal/mol in the enzyme (with a standard deviation of 2.2 kcal/mol). Clearly, were this to be real, this would result in significant product trapping, whereas Fig. 2 of [19] clearly demonstrates that PAS achieves multiple turnovers for phosphate hydrolysis. However, it is completely possible that the excessive exothermicity of this reaction is simply a result of problems accurately evaluating the ΔG_0 of the corresponding reference reaction in solution and is thus a simulation artifact, and further studies (such as, for instance, product inhibition assays) would be required in order to verify whether the observed exothermicity is indeed real or only a simulation artifact. The key issue here is that we have been able to reproduce the catalytic activity of the enzyme

with good agreement with experiment (with a standard deviation of ~1 kcal/mol over 10 trajectories), and without the need to protonate the leaving group. Therefore, our overall conclusion is that the most likely mechanism for the catalysis of phosphoryl transfer by PAS is a concerted substrate-as-base mechanism where the phosphate is the ultimate proton acceptor and the leaving group departs in anionic form.

The corresponding transition state for the enzyme-catalyzed reaction from a representative trajectory is shown in Fig. 8. The average P-O distances to the attacking nucleophile and departing leaving group over 10 trajectories are ~2.35 and 2.22Å respectively, compared to 2.46 and 2.29Å for the corresponding EVB simulations in solution. Therefore, the transition state would appear to be very slightly more compact in the enzymatic reaction compared to the reaction in solution. Note also that the EVB transition state in solution is more expansive than the one obtained from the *ab initio* calculations, however, the two are not directly comparable for a variety of reasons (such as the use of explicit solvent *vs.* a continuum or differences due to the parameterization of the forcefield used for the EVB calculations to name a few), and in order for any theoretical comparison between the enzymatic reaction and the corresponding uncatalyzed reaction in solution to be meaningful, this comparison has to be made based on calculations *at the same* level of theory with the *same* parameter set (as is the case with our EVB calculations). Additionally, when examining distances to relevant surrounding residues in the active site at the transition state, it would appear that Lys375 is within hydrogen bonding distance of one of the equatorial oxygen atoms of the PO_3^- in flight, which is in turn coordinating to the Ca^{2+} ion. Finally, the oxyanion of the departing leaving group appears to be within hydrogen bonding distance of His221, which appears to be assisting its departure.

In order to further dissect the contributions of different residues to the overall activation barrier, we have evaluated the electrostatic group contributions of each residue using the LRA approach, as was

done in e.g. [67]. The corresponding results are summarized in Figs. 9a and b, which show the group contributions to (a) the overall phosphoryl transfer step (with concomitant proton transfer from the geminal diol to the phosphate, to generate a phosphate monoanion), and (b) only the isolated initial proton transfer to the phosphate (discussed in Section III.2). From Fig. 9a, it would appear that there are significant stabilizing contributions from Asp13/14 (note that as the contributions from the two residues are fairly similar they appear as almost one bar in the figure) and Asp317, with a very modest contribution from His211 (which appears to be assisting in leaving group departure). Additionally, there appear to be significant destabilizing contributions from both Lys113 and 375 and Arg55, as well as a *very* large destabilizing contribution from the Ca^{2+} ion. This latter observation would seem counterintuitive, not only in light of the fact that the reaction involves nuclephilic attack on a highly charged dianion, but also in light of the large stabilizing role of the Ca^{2+} ions in similar phosphoryl transfer reactions such as that catalyzed by Staphylococcal nuclease[88,89].

However, there are two important issues that need to be taken into account here. The first of these is the fact that, as discussed in[88-90], correctly quantifying the effect of the metal cation is *extremely* challenging, and therefore, in the case of the Ca^{2+} ion, the results presented in Fig. 9 can only provide a qualitative and not quantitative trend (particularly when considering the very large apparent deactivating effect of the Ca^{2+} during the chemical step). The second is that the role of the Ca^{2+} in this reaction is potentially quite complex, as, while the enzyme initially binds a phosphate dianion, the actual chemical step involves hydroxide attack on a phosphate monoanion (following proton transfer to the substrate). Therefore, it is conceptually possible that the Ca^{2+} provides a significant electrostatic contribution to the binding of the phosphate dianion, and assists in the proton transfer from the geminal diol to the phosphate, however, once the phosphate monoanion has been generated, this proton effectively replaces the role of the Ca^{2+} ion in helping neutralize the charge on the phosphate, and thus the electrostatic

contribution of the Ca^{2+} to the subsequent phosphoryl transfer is significantly comparatively diminished, though it is also entirely plausible that the deactivating effect of the Ca^{2+} is down to the loss of electron density on the phosphate due to the departing leaving group. In order to test whether the Ca^{2+} plays a significant role in the initial proton transfer, we have also examined the electrostatic group contributions to the isolated initial proton transfer step (Fig. 9b, see also Section III.2), and, as can be seen from this figure, there is now indeed a stabilizing electrostatic contribution from the Ca^{2+} ion, and it appears to be helping to drive the proton transfer. There also appear to be destabilizing contributions to the proton transfer from both Lys113 and 375, His211 and Asp317, which is conceptually reasonable when one takes into account the reduction in charge on the substrate from dianion to monoanion upon the proton transfer. Thus, the qualitative picture that can be drawn from Figs. 9a and b is that the Ca^{2+} ion appears to be important in assisting the deprotonation of the nucleophile to generate the phosphate monoanion (and, by extension, binding of the dianionic substrate), whereas it destabilizes the subsequent phosphoryl transfer. Additionally, His211 appears to play a modest role in assisting the departure of the leaving group. However, while both Lys113 and Lys375 are most likely important for the binding of a dianionic substrate, they subsequently provide no stabilization to the chemical step (most probably due to the coupled proton transfer to the substrate). Based on the overall data presented above, we would therefore like to propose a revised mechanism for the phosphate monoesterase activity of PAS (Fig. 10), in which the phosphoryl transfer occurs in a concerted A_ND_N mechanism, with the phosphate itself acting as the base, and with no protonation of the departing leaving group.

IV. Conclusions

Despite the classical image of enzyme selectivity, namely one-enzyme one-substrate, it is becoming increasingly clear that a large number of enzymes are capable of "catalytic promiscuity" (i.e. the

catalysis of additional chemical reactions on top of their native reaction). Such promiscuity can manifest itself in the cross-catalysis of very similar substrates, as is the case with members of the alkaline phosphatase superfamily[3,16] that can intercatalyze phosphate, sulfate and phosphonate transfer, or substrates that are very different from each other, as in the case of the phosphotriesterase[91] that has demonstrated both phosphatase and lactonase activity, to name just a few examples. Such catalytic promiscuity is important, as it can potentially play an important role in the evolution of function within enzyme superfamilies[2,6,7]. Here, *Pseudomonas aeruginosa* arylsulfatase (PAS), a member of the alkaline phosphatase superfamily, is of particular interest, as while many members of this superfamily are primarily phosphatases that can also catalyze sulfuryl transfer, PAS is a native sulfate monoester hydrolyase[17,18], capable of hydrolyzing both phosphate mono-[19] and diesters[20] in addition to its native sulfatase activity, with multiple turnovers. In order to explore this phenomenon, we first engaged in a comparative study of the hydrolysis of *p*-nitrophenyl phosphate and sulfate in aqueous solution[23], both of which are prototype systems for examining phosphoryl and sulfuryl transfer respectively. In our previous work, we demonstrated that despite the deceptive similarities in the experimental data for phosphoryl and sulfuryl transfer, the nature of the transition state for the two reactions is quite different (with sulfuryl transfer proceeding through a far more expansive transition state than phosphoryl transfer). Additionally, in solution, there is apparently no need for the protonation of the sulfate in the transition state in contrast to the phosphate, and the solvation effects for the two transition states are very different (which is unsurprising in light of the fact that the two systems differ by a full charge unit). This is interesting, as it suggests that the requirements for efficient catalysis of phosphoryl and sulfuryl transfer are going to be very different, whereas PAS is clearly capable of catalyzing both.

The present work examines the *phosphoryl transfer* reaction catalyzed by PAS, comparing this to the reactivity of native phosphatases, such as the RAS GTPase[24,25]. We demonstrate that, like this system (as

well as related systems such as the elongation factor Tu, EF-Tu[19,27]) the reaction is most likely to proceed *via* a substrate-as-base mechanism with proton transfer to the phosphate itself. The subsequent phosphoryl transfer step appears to involve a concerted A_ND_N mechanism, and the transition state for the phosphoryl transfer in the enzyme is slightly more compact than that for the corresponding reaction in solution. Therefore, there appears to essentially be nothing "exotic" about the phosphatase activity of PAS, and despite being a native sulfatase, it is clearly capable of catalyzing phosphoryl transfer with reasonable efficiency (providing a catalytic amplification of ~8 kcal/mol), though this catalytic amplification is still extremely small compared to that of native phosphatases such as Staphylococcal nuclease[88]. What *is* significant here, however, is that while the promiscuous phosphtase activity is in line with (but poorer than) related enzymes whose primary physiological role is to catalyze phosphoryl transfer, the substrate-as-base mechanism observed here for the phosphoryl transfer catalyzed by PAS is *hugely unlikely* to be a viable mechanism for the analogous sulfuryl transfer reaction, due to the *much* lower pKa (less than -3[92]) of the non-briding oxygens of *p*-nitrophenyl sulfate. Detailed studies of this issue are under way, however, even without knowing the precise mechanism for the sulfuryl transfer, it is clear that in contrast to the phosphoryl transfer reaction, the sulfuryl transfer reaction will require PAS to recruit a base (most likely an enzymatic side-chain the identity of which is at present elusive), meaning that the promiscuous phosphatase activity of PAS is following a "chemically simplified" reaction mechanism in which the phosphate itself is capable of acting as a proton sink with no need for an external base, suggesting quite a large scope for catalytic versatility within the same active site.

A factor to take into account is that there are an unusually large number of ionizable residues within close vicinity of the substrate, namely two lysines, two histidines, two glutamines, and three asparatates (which are coordinating the Ca^{2+} ion in the active site). Additionally, the experimentally measured[19] pH dependence of k_{cat}/K_M is quite different for the native sulfatase activity and the promiscuous phosphate

monoesterase activity, with the particularly notable fact that the apparent pK_a values in the k_{cat}/K_M pH profiles differ by at least 2 pH units (8.3 for the sulfatase and significantly lower for the phosphatase activities, see [19]). This yields the possibility that the electrostatic environment can be quite different for the binding of the sulfate compared to the phosphatase (see also [19]), which is unsurprising in light of the difference in charge. However, it will be of great interest to understand how the enzyme is able to adapt to catalyze the reaction of two substrates that are likely to have quite different catalytic requirements. Further studies are underway to investigate this, and a detailed understanding of this issue will have significant implications for computer assisted enzyme design, where the most successful strategy to date has been the rational improvement of active site preorganization[14,15].

ACKNOWLEDGEMENTS

The authors thank the Swedish Research Council (SCLK, VR, grant 2010-5026) and the BBSRC (BvL) for funding, as well as the PDC center for High Performance Computing at KTH and the National Supercomputer Center (NSC) at Linköping University for access to computational resources (SNIC grants 001/11-12 and 001/11-20 respectively). Finally, thanks also go to Arieh Warshel (University of Southern California), Nicholas Williams (University of Sheffield), Johan Åqvist (Uppsala University) and Paul Wentworth (University of Oxford/The Scripps Research Institute) for insightful and stimulating discussion.

REFERENCES

1. Fischer E. Einfluss der configuration auf der Wirkung der Enzyme. Ber Dtsch Chem Ges 1894;27:2985-2993.
2. Khersonsky O, Tawfik DS. Enzyme promiscuity: A mechanistic and evolutionary perspective. Ann Rev Biochem 2010;79:471-505.

3. Jonas S, Hollfelder F. Mapping catalytic promiscuity in the alkaline phosphatase superfamily. Pure Appl Chem 2009;81:731-742.
4. Bornscheuer UT, Kazlauskas RJ. Catalytic promiscuity in biocatalysis: Using old enzymes to form new bonds and follow new pathways. Angew Chem Int Ed 2004;43:6032-6040.
5. Khersonsky O, Roodveldt C, Tawfik DS. Enzyme promiscuity: evolutionary and mechanistic aspects. Curr Opin Chem Biol 2006;10:498-508.
6. O'Brien PJ, Herschlag D. Catalytic promiscuity and the evolution of new enzymatic activities. Chem Biol 1999;6:R91-R105.
7. Jensen RA. Enzyme recruitment in evolution of new function. Annu Rev Microbiol 1976;30:409-425.
8. Lynch M. Gene duplication and evolution. Science 2002;297:945-947.
9. Chothia C, Gough J, Vogel C, Teichmann SA. Evolution of the protein repertoire. Science 2003;300:1701-1703.
10. Ohno S. Evolution by gene duplication. New York: Springer; 1970.
11. Warshel A. Energetics of enzyme catalysis. Proc Natl Acad Sci USA 1978;75:5250-5254.
12. Kamerlin SCL, Sharma PK, Chu ZT, Warshel A. Ketosteroid isomerase provides further support to the idea that enzymes work by electrostatic preorganization. Proc Natl Acad Sci U S A 2010;107:4075-4080.
13. Warshel A, Sharma PK, Kato M, Xiang Y, Liu H, Olsson MHM. Electrostatic basis for enzyme catalysis. Chem Rev 2006;106:3210-3235.
14. Frushicheva MP, Cao J, Chu ZT, Warshel A. Exploring challenges in rational enzyme design by simulating the catalysis in artificial Kemp eliminase. Proc Natl Acad Sci USA 2010;107:16869-16874.
15. Frushicheva MP, Cao J, Warshel A. Challenges and advances in validating enzyme design proposals: The case of Kemp eliminase catalysis. Biochemistry 2011;50:3849-3858.
16. Babtie A, Tokuriki N, Hollfelder F. What makes an enzyme promiscuous? Curr Opin Chem Biol 2010;14:200-207.
17. Galperin MY, Bairoch A, Koonin EV. A superfamily of metalloenzymes unifies phosphopentomutase and cofactor-independent phosphoglycerate mutase with alkaline phosphatases and sulfatases. Protein Sci 1998;7:1829-1835.

18. Galperin MY, Jedrzejas MJ. Conserved core structure and active site residues in alkaline phosphatase superfamily enzymes. Proteins 2001;45:318-324.
19. Olguin LF, Askew SE, O'Donoghue AC, Hollfelder F. Efficient catalytic promiscuity of an enzyme superfamily: An arylsulfatase shows a rate acceleration of 10^{13} for phosphate monoester hydrolysis. J Am Chem Soc 2008;130:16547-16555.
20. Babtie AC, Bandyopadhyay S, Olguin LF, Hollfelder F. Efficient catalytic promiscuity for chemically distinct reactions. Angew Chem Int Ed 2009;48:3692-3694.
21. Kirby AJ, Jenks WP. The Reactivity of nucleophilic reagents toward the *p*-Nirophenyl phosphate dianions. J Am Chem Soc 1965;87:3209.
22. Hoff RH, Larsen P, Hengge AC. Isotope effects and medium effects on sulfuryl transfer reactions. J Am Chem Soc 2001;123:9338-9344.
23. Kamerlin SCL. A theoretical comparison of *p*-nitrophenyl phosphate and sulfate hydrolysis in aqueous solution: Implications for enzyme catalyzed sulfuryl transfer. In press.
24. Schweins T, Langen R, Warshel A. Why have mutagenesis studies not located the general base in ras p21. Nat Struct Biol 1994;1(7):476-484.
25. Schweins T, Geyer M, Scheffzek K, Warshel A, Kalbitzer HR, Wittinghofer A. substrate-assisted catalysis as a mechanism for GTP Hydrolysis of p21 ras and other GTP-binding proteins. Nat Struct Biol 1995;2(1):36-44.
26. Adamczyk AJ, Warshel A. Converting structural information into an allosteric-energy-based picture for elongation factor Tu activation by the ribosome. Proc Natl Acad Sci U S A In Press;108:9827-9832.
27. Liljas A, Ehrenberg M, Åqvist J. Comment on "The mechanism for the activation of GTP hydrolysis on the ribosome". Science 2011;333:37.
28. Wolfenden R, Snider MJ. The depth of chemical time and the power of enzymes as catalysts. Acc Chem Res 2001;34(12):938-945.
29. Åqvist J, Kolmodin K, Florián J, Warshel A. Mechanistic alternatives in phosphate monoester hydrolysis: what conclusions can be drawn from available experimental data? Chem & Biol 1999;6(3):R71-R80.
30. Klähn M, Rosta E, Warshel A. On the mechanism of hydrolysis of phosphate monoester dianions in solutions and proteins. J Am Chem Soc 2006;128(47):15310-15323.

31. Rosta E, Kamerlin SCL, Warshel A. On the interpretation of the observed linear free energy relationship in phosphate hydrolysis: A thorough computational study of phosphate diester hydrolysis in solution. Biochemistry 2008;47(12):3725-3735.
32. Kamerlin SCL, Florián J, Warshel A. Associative versus dissociative mechanisms of phosphate monoester hydrolysis: On the interpretation of activation entropies. ChemPhysChem 2008;9:1767-1773.
33. Kamerlin SCL, Williams NH, Warshel A. Dineopentyl phosphate hydrolysis: Evidence for stepwise water attack. J Org Chem 2008;73:6960-6969.
34. Alkherraz A, Kamerlin SCL, Feng G, Sheik QI, Warshel A, Williams NH. Phosphate ester analogues as probes for understanding enzyme catalyzed phosphoryl transfer. Faraday Discuss 2010;145:281-299.
35. Kamerlin SCL, Wilkie J. The role of metal ions in phosphate ester hydrolysis. Org Biomol Chem 2007;5:2098-2108.
36. Kamerlin SCL, McKenna CE, Goodman MF, Warshel A. A computational study of the hydrolysis of dGTP analogues with halomethylene-modified leaving groups in solution: Implications for the mechanism of DNA polymerases. Biochemistry 2009;48:5963-5971.
37. Jencks WP. A primer for the Bema Hapothle. An empirical approach to the characterization of changing transition-state structures. Chem Rev 1985;85(6):511-527.
38. More O'Ferrall RA. Relationships between E2 and E1cB mechanisms of β-elimination. J Chem Soc B 1970:274 - 277.
39. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Vreven JT, Kudin KN, Burant JC, Millam JM, Iyengar SS, J. Tomasi, V. Barone BM, M. Cossi, G. Scalmani, N. Rega,, G. A. Petersson HN, M. Hada, M. Ehara, K. Toyota,, R. Fukuda JH, M. Ishida, T. Nakajima, Y. Honda, O. Kitao,, H. Nakai MK, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross,, C. Adamo JJ, R. Gomperts, R. E. Stratmann, O. Yazyev,, A. J. Austin RC, C. Pomelli, J. W. Ochterski, P. Y. Ayala,, K. Morokuma GAV, P. Salvador, J. J. Dannenberg,, V. G. Zakrzewski SD, A. D. Daniels, M. C. Strain,, O. Farkas DKM, A. D. Rabuck, K. Raghavachari,, J. B. Foresman JVO, Q. Cui, A. G. Baboul, S. Clifford,, J. Cioslowski BBS, G. Liu, A. Liashenko, P. Piskorz,, I. Komaromi RLM, D. J. Fox, T. Keith, M. A. Al-Laham,, C. Y. Peng AN, M. Challacombe, P. M. W. Gill,, B. Johnson WC, M. W. Wong, C. Gonzalez, and J. A. Pople. Gaussian 03, Revision D.01; 2004.

40. Becke AD. Density-Functional Thermochemistry. 3. The Role of Exact Exchange. J Chem Phys 1993;98(7):5648-5652.

41. 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.

42. Cancès MT, 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.

43. Mennucci B, Cancès MT, Tomasi J. Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: Theoretical basis, computational implementation and numerical applications. J Phys Chem B 1997;101:10506-10517.

44. 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.

45. Barone V, Cossi M. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. J Phys Chem A 1998;102(11):1995-2001.

46. Klamt A, Schuurmann G. Cosmo - A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. J Chem Soc Perkins Trans 2 1993(5):799-805.

47. Štrajbl M, Sham YY, Villà J, Chu ZT, Warshel A. Calculations of activation entropies of chemical reactions in solution. J Phys Chem B 2000;104:4578-4584.

48. Siegbahn PE, Himo F. Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. J Biol Inorg Chem 2009;14:643-651.

49. Kamerlin SCL, Haranczyk M, Warshel A. Progress in *ab initio* QM/MM free-energy simulations of electrostatic energies in proteins: Accelerated QM/MM studies of pK_a, redox reactions and solvation free energies. J Phys Chem B 2009;113:1253-1272.

50. Shurki A, Warshel A. Structure/function correlations of proteins using MM, QM/MM, and related approaches: Methods, concepts, pitfalls, and current progress. Adv Protein Chem 2003;66:249-313.

51. Senn HM, Thiel W. QM/MM methods for biomolecular systems. Angew Chem Int Ed 2009;48:1198-1229.

52. Warshel A, Weiss RM. An empirical valence bond approach for comparing reactions in solutions and in enzymes. J Am Chem Soc 1980;102(20):6218-6226.

53. Hwang J-K, King G, Creighton S, Warshel A. Simulation of free energy relationships and dynamics of S_N2 reactions in aqueous solution. J Am Chem Soc 1988;110(16):5297-5311.

54. Warshel A, Florián J. The empirical valence bond (EVB) method. In: von Ragué Schleyer P, Allinger NL, Clark T, Gasteiger J, Kollman PA, Schaefer III HF, Schreiner PR, editors. The Encyclopedia of Computational Chemistry. Chichester, UK: John Wiley & Sons; 2004.

55. Roca M, Vardi-Kilshtain A, Warshel A. Toward accurate screening in computer-aided enzyme design. Biochemistry 2009;48:3046-3056.

56. Vardi-Kilshtain A, Roca M, Warshel A. The empirical valence bond as an effective strategy for computer-aided enzyme design. Biotechnol J 2009;4:495-500.

57. Lee FS, Chu ZT, Warshel A. Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. J Comp Chem 1993;14:161-185.

58. Sham YY, Chu ZT, Tao H, Warshel A. Examining methods for calculations of binding free energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA calculations of ligands binding to an HIV protease. Proteins: Struct Funct Genet 2000;39:393-407.

59. Villa J, Warshel A. MOLARIS: Version alpha9.06, Theoretical background and practical examples. 2002.

60. Warshel A. Computer modeling of chemical reactions in enzymes and solutions. New York: John Wiley & Sons; 1991.

61. Boltes I, Czapinska H, Kahnert A, von Bülow R, Dierks T, Schmidt B, von Figura K, Kertesz MA, Usón I. 1.3Å structure of arylsulfatase from *Pesudomonas aeruginosa* establishes the catalytic mechanism of sulfate ester cleavage in the sulfatase family. Structure 2001;9:483-491.

62. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucl Acids Res 2000;28:235-242.

63. Ritchie DW. Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. PROTEINS: Struct Func Genet 2003;52:98-106.

64. Li H, Robertson AD, Jensen JH. Very fast empirical prediction and interpretation of protein pK_a values. Proteins 2005;61:704-721.

65. Bas DC, Rogers DM, Jensen JH. Very fast prediction and rationalization of pK_a values for protein-ligand complexes. Proteins 2008;73:765-783.

66. Lee FS, Warshel A. A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations. J Chem Phys 1992;97:3100-3107.

67. Adamczyk AJ, Cao J, Kamerlin SCL, Warshel A. The catalytic power of dihydrofolate reductase and other enzymes arises from electrostatic preorganization, not conformational motions. Proc Natl Acad Sci U S A In Press.

68. Schmidt S, Arlt T, Hamm P, Huber H, Naegele T, Wachtveitl J, Meyer M, Scheer H, Zinth W. Energetics of the primary electron transfer reaction revealed by ultrafast spectroscopy on modified bacterial reaction centers. Chem Phys Lett 1994;223:116-120.

69. Selmer T, Hallmann A, Schmidt B, Sumper M, von Figura K. The evolutionary conservation of a novel protein modification, the conservation of cysteine to serinesemialdehyde in arylsulfatase from *Volvox carteri*. Eur J Biochem 1996;238:341-345.

70. Dierks T, Schmidt B, von Figura K. Conversion of cysteine to formylglycine: a protein modification in the endoplasmic reticulum. Proc Natl Acad Sci U S A 1997;94:11963-11968.

71. Meich C, Dierks T, Selmer T, von Figura K, Schmidt B. Arylsulfatase from *Klebsiella pneumoniae* carries a formylglycine generated from a serine. J Biol Chem 1998;27:4835-4837.

72. Anslyn EV, Dougherty DA. Modern physical organic chemistry. USA: University Science Books; 2006.

73. pKa data compiled by R. Williams.

74. Zalatan JG, Herschlag D. Alkaline phosphatase mono- and diesterase reactions: Comparative transition state analysis. J Am Chem Soc 2006;128(4):1293-1303.

75. Warshel A. Calculations of enzymic reactions: Calculations of pK_a, proton transfer reactions, and general acid catalysis reactions in enzymes. Biochemistry 1981;20:3167-3177.

76. Langen R, Schweins T, Warshel A. On the mechanism of guanosine triphosphate hydrolysis in ras p21 proteins. Biochemistry 1992;31:8691-8696.

77. Florián J, Goodman MF, Warshel A. Computer simulation of the chemical catalysis of DNA polymerases: Discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase. J Am Chem Soc 2003;125:8163-8177.

78. Scheiner S, Kar T. The nonexistance of specially stabilized hydrogen bonds in enzymes. J Am Chem Soc 1995;117:6970-6975.

79. Villà J, Warshel A. Energetics and dynamics of enzymatic reactions. J Phys Chem B 2001;105(33):7887-7907.
80. Eigen M. Proton transfer, acid-base catalysis, and enzymatic hydrolysis. Part I: Elementary processes. Angew Chem Int Ed 1964;3(1):1-19.
81. Florián J, Warshel A. A fundamental assumption about OH^- attack in phosphate hydrolysis is not fully justified. J Am Chem Soc 1997;119:5473-5474.
82. Admiraal SJ, Herschlag D. The substrate-assisted general base catalysis model for phosphate monoester hydrolysis: Evaluation using reactivity comparisons. J Am Chem Soc 2000;122(10):2145-2148.
83. Lassila JK, Zalatan JG, Herschlag D. Biological phosphoryl transfer reactions: Understanding mechanism and catalysis. Annu Rev Biochem 2011;80:669-702.
84. Kirby AJ, Varvoglis AG. The reactivity of phosphate esters. Monoester hydrolysis. J Am Chem Soc 1967;89:415-423.
85. Hoff RH, Hengge AC. Entropy and enthalpy contributions to solvent effects on phosphate monoester solvolysis. The importance of entropy effects in the dissociative transition state. J Org Chem 1998;63(19):6680-6688.
86. Hengge AC. Isotope effects in the study of phosphoryl and sulfuryl transfer reactions. Acc Chem Res 2002;35:105-112.
87. Lima FS, Chaimovich H, Cuccovia IM. Kinetics and product distribution of p-nitrophenyl phosphate dianion solvolysis in ternary DMSO/alcohol/water mixtures are compatible with metaphosphate formation. J Phys Org Chem 2011;In Press.
88. Aqvist J, Warshel A. Calculations of free energy profiles for the Staphylococcal Nuclease catalyzed reaction. Biochemistry 1989;28:4680-4689.
89. Aqvist J, Warshel A. Free energy relationships in metalloenzyme-catalyzed reactions. Calculations of the effects of metal ion substitutions in Staphylococcal Nuclease. J Am Chem Soc 1990;112:2860-2868.
90. Xiang Y, Oelschlaeger P, Florian J, Goodman MF, Warshel A. Simulating the effect of DNA polymerase mutations on transition-state energetics and fidelity: Evaluating amino acid group contribution and allosteric coupling for ionized residues in human pol β. Biochemistry 2006;45(23):7036-7048.

91. Afriat L, Roodveldt C, Manco G, Tawfik DS. The latent promiscuity of newly identified microbial lactonases is linked to a recently diverged phosphotriesterase. Biochemistry 2006;45:13677-13686.

92. Gibby SG, Younker JM, Hengge AC. Investigation of the sulfuryl transfer step from substrate to enzyme by arylsulfatases. J Phys Org Chem 2004;17:541-547.

FIGURE CAPTIONS

Figure 1: Schematic illustration of the evolution of function from broadly generalized primordial "ancestor" enzymes to (comparatively) specialized modern enzymes. This figure is based on Jensen's hypothesis[7], and is adapted from [2,5].

Figure 2: The orientation of key active site residues in the ground state after 100ps of MD relaxation of PAS at 300K. In this figure, "pNPP" and "DDZ" denote *p*-nitrophenyl phosphate and FGly51 respectively. Here, all relevant ionizable residues within ~6-8Å of the reactive region have also been marked for clarity. Finally, note that with the exception of the Ca^{2+} ion, partial charges are not shown here.

Figure 3: Proposed mechanism for the promiscuous phosphatase activity of *Pseudomonas aeruginosa* arylsulfatase (PAS). This figure, which was adapted from [19,20], illustrates (a) the attack of a water molecule on an aldehyde to form the corresponding nucleophilic aldehydate hydrate (1→2), (b) nucleophilic attack on the phosphate with concomitant leaving group departure (2→3), and (c) hemiacetal cleavage to release the phosphate and regenerate the aldehyde (3→1). For clarity, the proposed acid and base are not shown in this figure; however, their identity is discussed in the main text. Additionally, there is a Ca^{2+} cation in the active site of PAS which is also not shown here for clarity, in line with the representation used in [20].

Figure 4: Valence bond structures representing reactants and intermediates for the initial deprotonation of the nucleophile in either an aspartate-as-base (I→II) or a substrate-as-base (I→III) mechanism.

Figure 5: Valence bond structures representing reactants (IV) and products (V) for the concerted (A_ND_N) phosphoryl transfer step, with concomitant proton transfer from the nucleophile to the phosphate, in a concerted substrate-as-base mechanism.

Figure 6: *Ab initio* free energy surface (kcal mol^{-1}) and transition state geometry for phosphoryl transfer from *p*-nitrophenyl phosphate to acetaldehyde hydrate.

Figure 7: Empirical valence bond (EVB) free energy profiles for the phosphoryl transfer step in PAS (blue) and in solution (black).

Figure 8: Representative transition state geometry for nucleophlic attack on *p*-nitrophenyl phosphate in *Pseudomonas aeruginosa* arylsulfatase. This corresponds to an A_ND_N substrate-as-base mechanism in which P-O formation to the nucleophile and cleavage to the leaving group occur in a single transition state, with concomitant proton transfer from the nucleophile to the phosphate (the proton can be seen in flight in both transition states). Note that the actual overall activation and free energies presented in Table 3 are obtained by averaging over 10 discrete trajectories.

Figure 9: The linear response approximation (LRA) group contributions to the overall activation barrier in the enzyme for all residues during (a) the overall S_N2 reaction with concomitant proton transfer to the substrate, and (b) the isolated initial proton transfer step in the substrate-as-base mechanism.

Figure 10: Revised mechanism for the promiscuous phosphomonoesterase activity of *Pseudomonas aeruginosa* arylsulfatase (PAS) (see also Fig. 3). The phosphoryl transfer reaction (step 2→3), which is the focus of the present work, is proposed to proceed *via* a concerted substrate-as-base mechanism,

where a proton from the attacking nucleophile is donated to the phosphate (resulting in the attack of an anionic nucleophile on a phosphate monoanion), and without corresponding protonation of the departing *p*-nitrophenyl leaving group.

Table 1: Activation and free energies and EVB parameters for the initial deprotonation of the nucleophile, in water and in the enzyme.

Base[a]	System	H_{12}[b] (kcal/mol)	α_0[b] (kcal/mol)	$\Delta g^{\neq}_{calc}$[c] (kcal/mol)	$\Delta G_{0,calc}$[c] (kcal/mol)	$\Delta G_{0,exp}$[d] (kcal/mol)
Asp317 (I→II)	Water	28.5	27.0	15.9(0.4)	13.2(0.3)	13.1
	Enzyme			19.0(0.6)	17.7(0.8)	
Phosphate (I→III)	Water	36.0	64.0	14.3(0.6)	12.1(0.9)	11.9
	Enzyme			10.9(0.4)	9.9(0.4)	

[a] Shown here are the relevant energetics of both the aspartate-as-base and substrate-as-base possibilities. The notation is according to Fig. 4. [b] H_{12} and α_0 denote the off-diagonal element of the EVB Hamiltonian and the gas-phase shift of the product state respectively (which appears in the H_{22} element of the EVB Hamiltonian). The gas-phase shift in the reactant state (H_{11}) is zero in all cases. These parameters were adjusted to reproduce the "experimental" energetics in solution, and then transferred unchanged to the enzymatic reaction. [c] Calculated activation energies and reaction free energies, in kcal/mol. The values in parentheses represent the standard deviation over 10 trajectories. [d] The "experimental" value is estimated from simple pK_a considerations, using the relationship shown in Eq. 1. This estimate is based on pK_as of 13.5, 3.9 and 4.8 for the nucleophile, Asp317, and the phosphate respectively (see also discussion in the main text). All calculated values given are an average over 10 trajectories.

Table 2: Energy decomposition of $\Delta g^{\neq}_{calc}$ and $\Delta G°_{calc}$ for phosphoryl transfer from *p*-nitrophenyl phosphate to acetaldehyde hydrate in solution. All energies are given in kcal/mol.

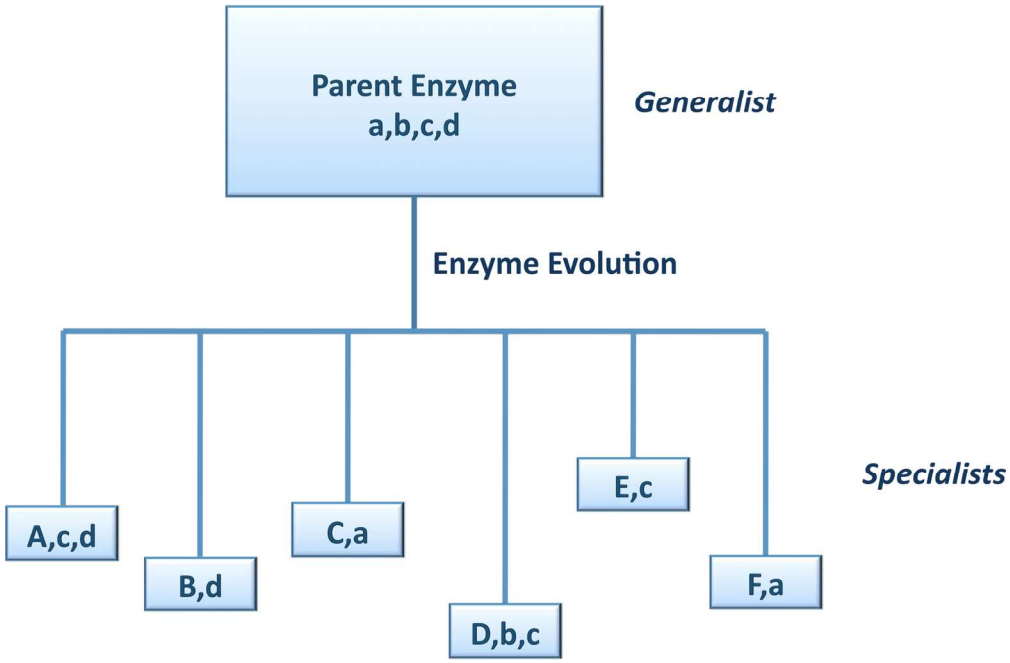
Species	ΔE_{gas}	ΔE_{pol}	$\Delta\Delta G_{solv}$	ZPE	$-T\Delta S_{conf}$	Total
$\Delta g^{\neq}_{calc}$	-5.3	-1.2	30.9	0.3	6.5	**31.2**
$\Delta G°_{calc}$	-93.0	-6.2	89.9	0.4	-0.8	**-9.8**

1

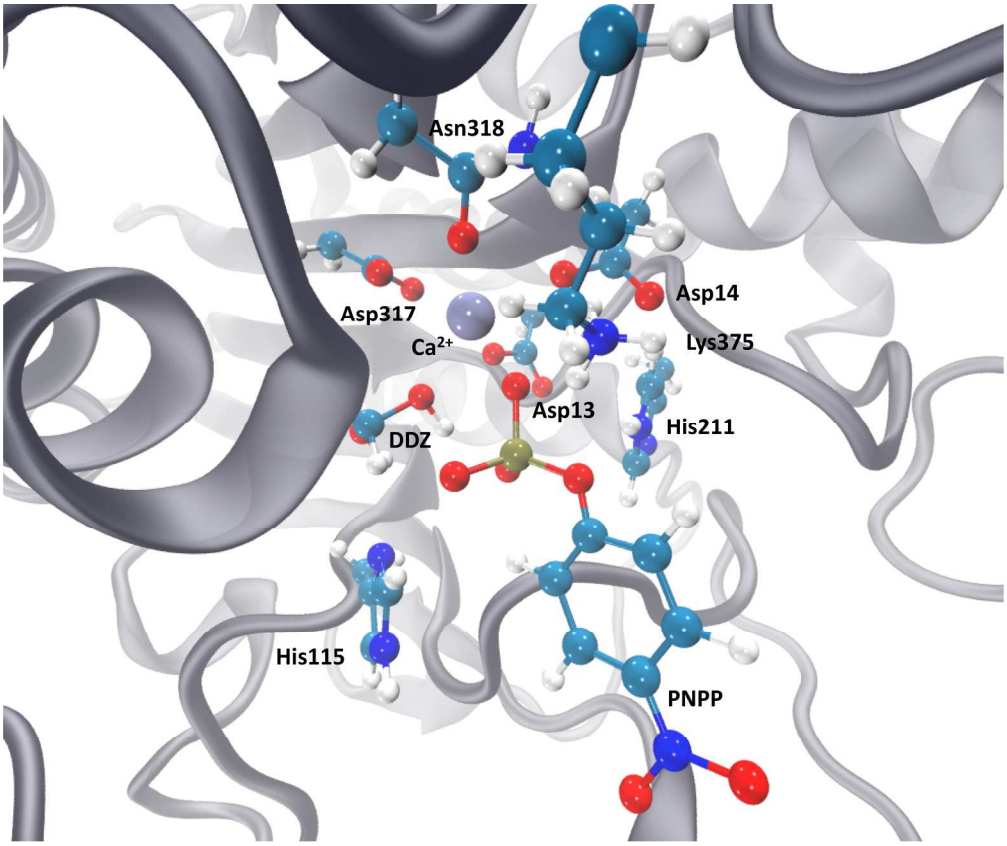
Table 3: EVB parameters, calculated activation and free energies and reorganization energies for the phosphoryl transfer step in the chemical reaction catalyzed by *Pseudomonas aeruginosa* arylsulfatase (corresponding to the valence bond structures IV→V of Fig. 5).

System	H_{12}[a] (kcal/mol)	α_0[a] (kcal/mol)	$\Delta g^{\neq}_{calc}$[b] (kcal/mol)	$\Delta G_{0,calc}$[b] (kcal/mol)	$\Delta g^{\neq}_{exp}$[c] (kcal/mol)	λ_{tot}[d] (kcal/mol)	λ_{elec}[d] (kcal/mol)
Water	310.0	-40.0	27.3(0.3)	-10.4(1.4)	~27.0	807.1(6.8)	38.2(2.4)
Enzyme			19.5(1.1)	-12.7(2.2)	19.6	726.9(2.5)	24.6(0.7)

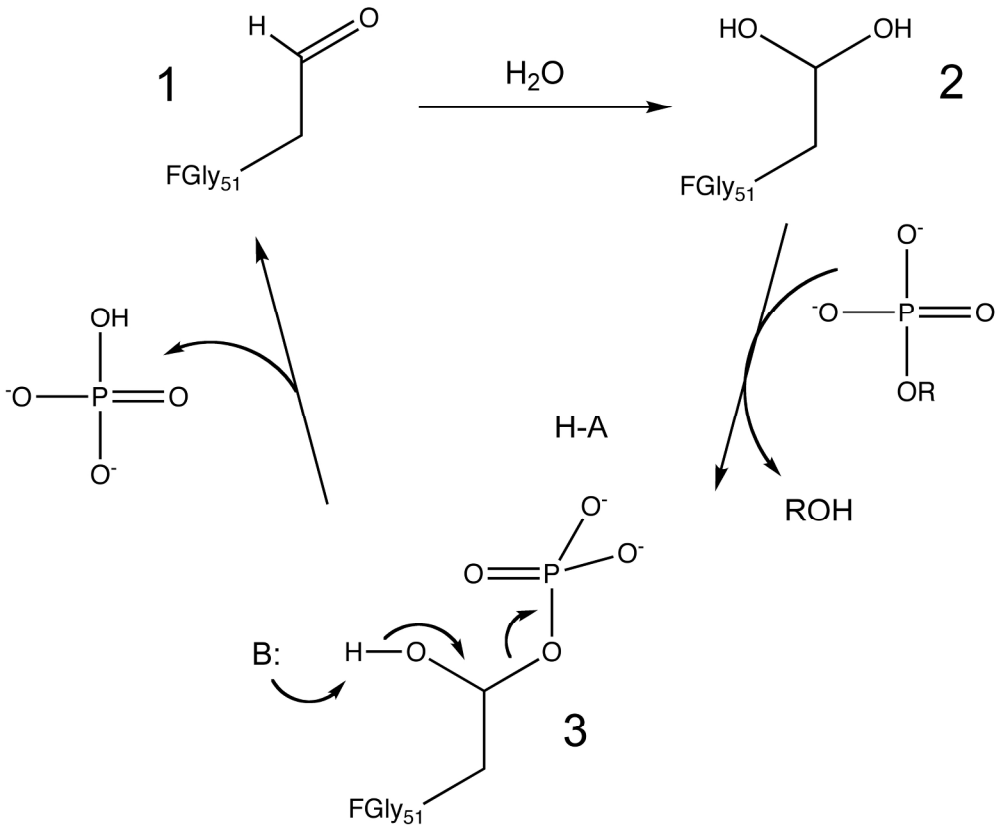
[a] As in Table 1, H_{12} and α_0 denote the off-diagonal element of the EVB Hamiltonian and the gas-phase shift of the product state respectively (which appears in the H_{22} element of the EVB Hamiltonian). The gas-phase shift in the reactant state (H_{11}) is zero in all cases. These parameters were adjusted to reproduce the *ab initio* and experimental energetics in solution, and then transferred unchanged to the enzymatic reaction. [b] Calculated activation energies and reaction free energies, in kcal/mol. The values in parenthesis represent the standard deviation over 10 trajectories. [c] The experimentally observed activation free energies in kcal/mol. The activation free energy in solution is estimated based on the experimentally measured rates for hydroxide and water attack on *p*-nitrophenyl phosphate, taking into account the energetic cost associated with bringing the reacting fragments from infinite separation into the reaction cage (see the discussion in the main text). [d] The total (λ_{total}) reorganization energy, as well as the electrostatic contribution (λ_{elec}) to λ_{tot}, in solution and in the enzyme, in kcal/mol.



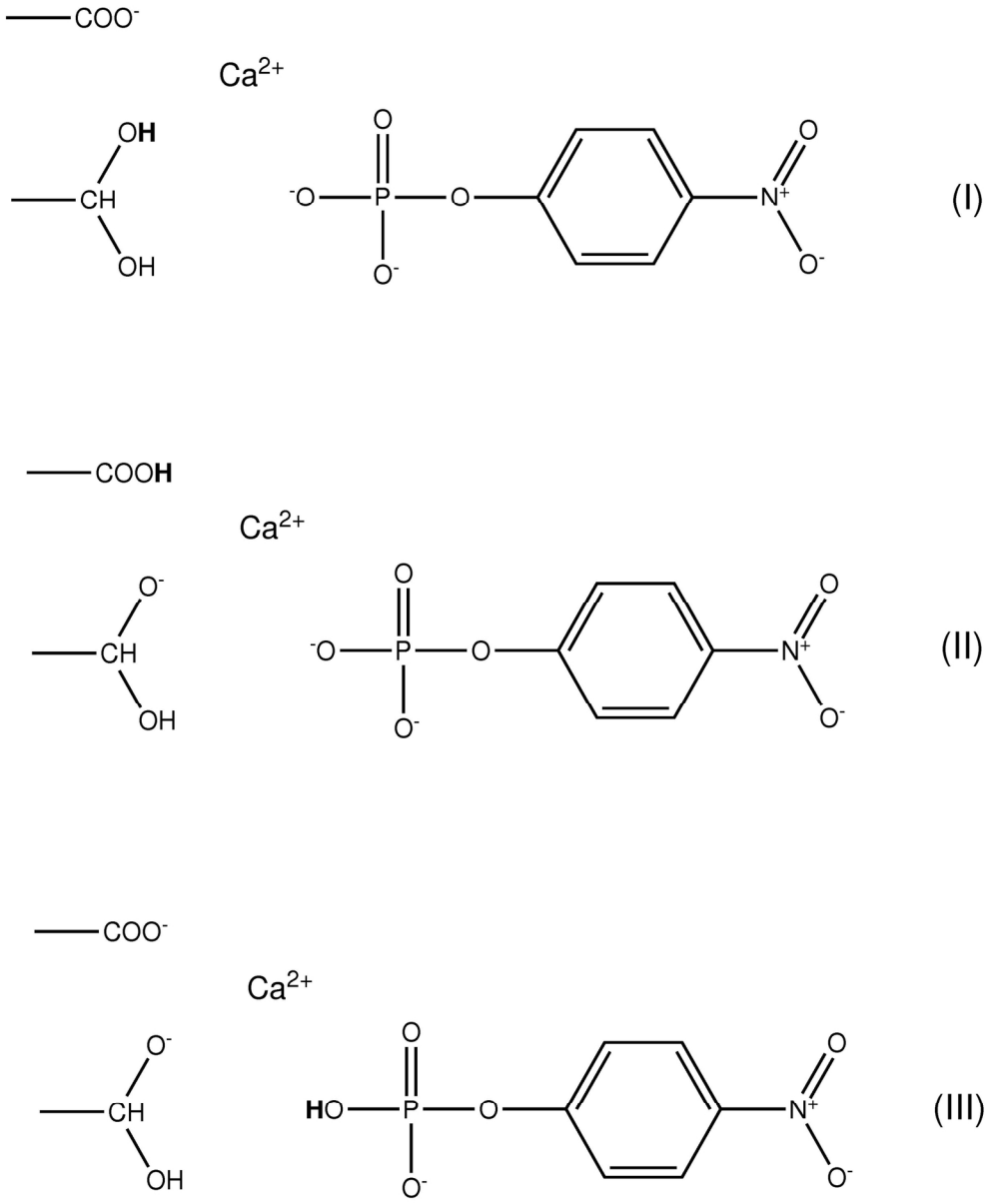
Schematic illustration of the evolution of function from broadly generalized primordial "ancestor" enzymes to (comparatively) specialized modern enzymes. This figure is based on Jensen's hypothesis[7], and is adapted from [2,5].
77x51mm (600 x 600 DPI)



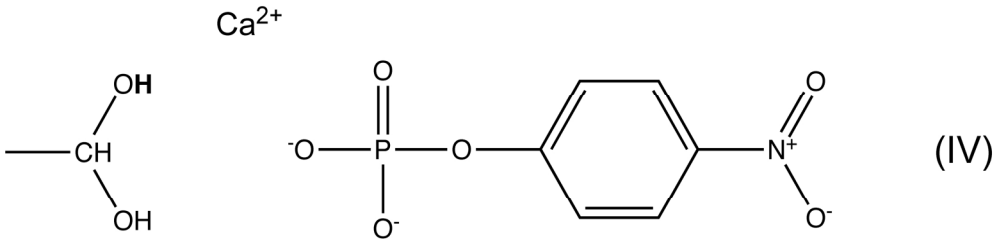
The orientation of key active site residues in the ground state after 100ps of MD relaxation of PAS at 300K. In this figure, "pNPP" and "DDZ" denote *p*-nitrophenyl phosphate and FGly51 respectively. Here, all relevant ionizable residues within ~6-8Å of the reactive region have also been marked for clarity. Finally, note that with the exception of the Ca^{2+} ion, partial charges are not shown here.
119x99mm (600 x 600 DPI)



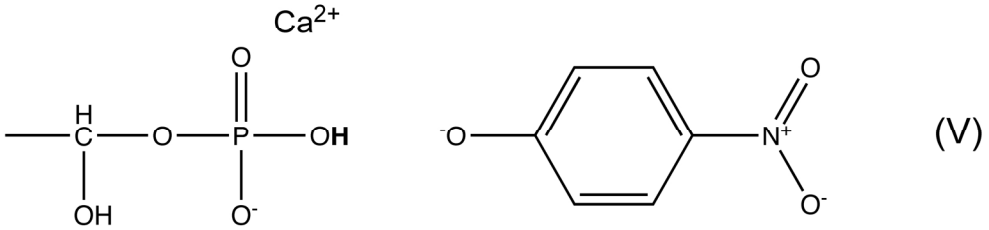
Proposed mechanism for the promiscuous phosphatase activity of Pseudomonas aeruginosa arylsulfatase (PAS). This figure, which was adapted from [19,20], illustrates (a) the attack of a water molecule on an aldehyde to form the corresponding nucleophilic aldehydate hydrate (1→2), (b) nucleophilic attack on the phosphate with concomitant leaving group departure (2→3), and (c) hemiacetal cleavage to release the phosphate and regenerate the aldehyde (3→1). For clarity, the proposed acid and base are not shown in this figure; however, their identity is discussed in the main text. Additionally, there is a Ca^{2+} cation in the active site of PAS which is also not shown here for clarity, in line with the representation used in [20].
111x92mm (600 x 600 DPI)



Valence bond structures representing reactants and intermediates for the initial deprotonation of the nucleophile in either an aspartate-as-base (I→II) or a substrate-as-base (I→III) mechanism.
163x198mm (300 x 300 DPI)

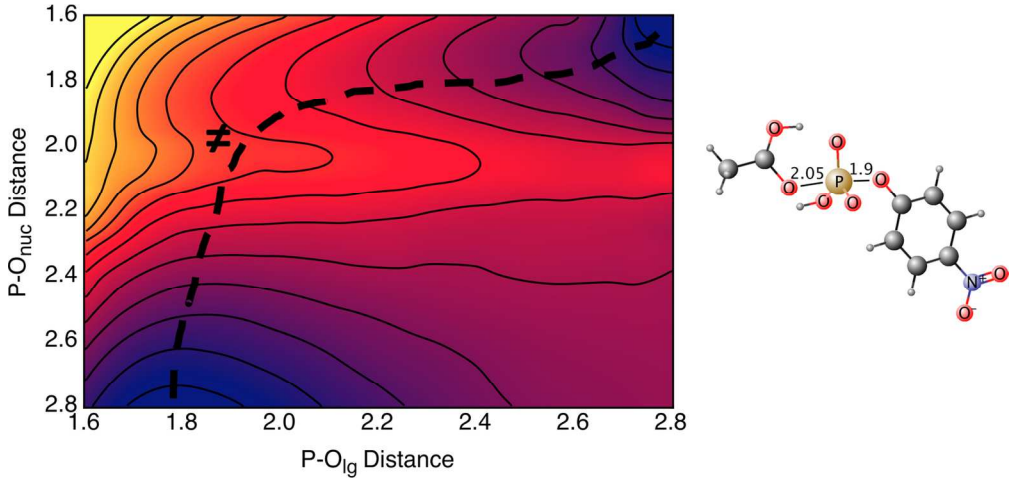


(IV)

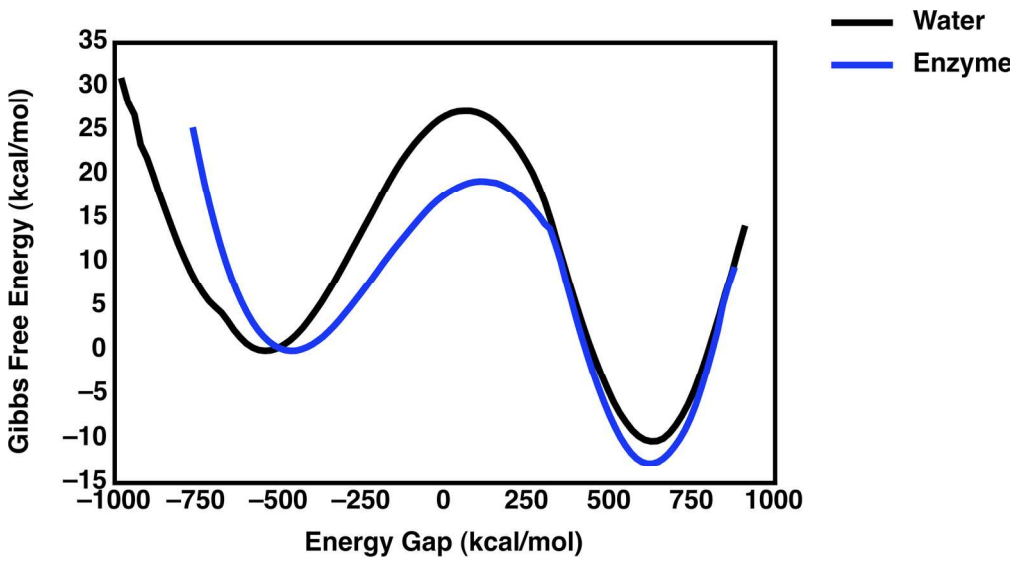


(V)

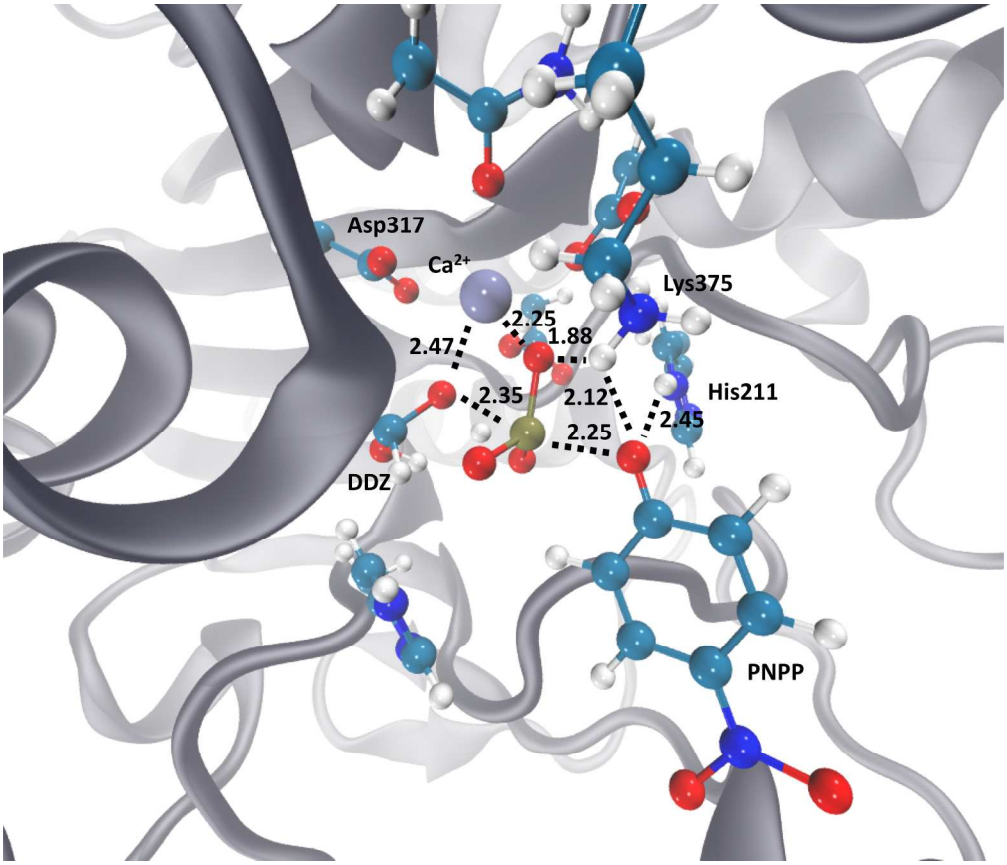
Valence bond structures representing reactants (IV) and products (V) for the concerted (A_ND_N) phosphoryl transfer step, with concomitant proton transfer from the nucleophile to the phosphate, in a concerted substrate-as-base mechanism.
95x68mm (600 x 600 DPI)



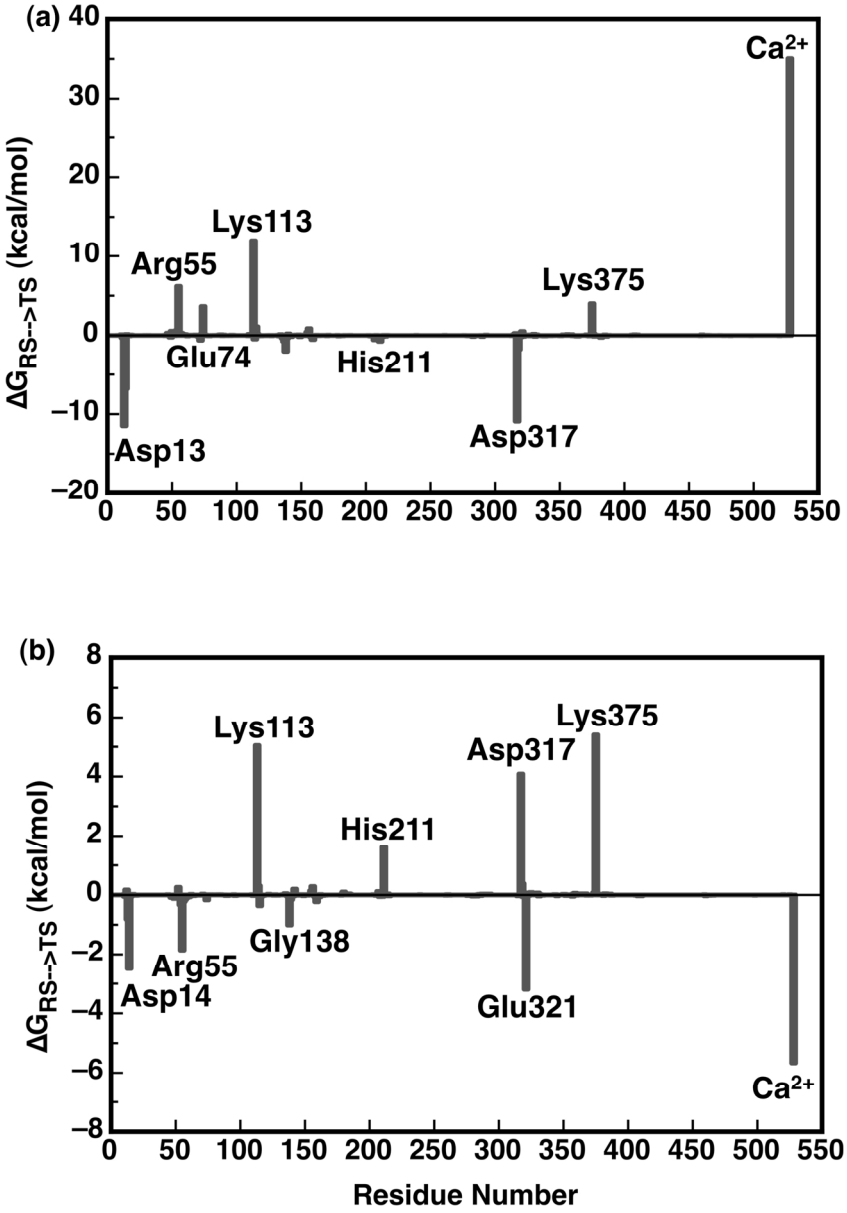
Ab initio free energy surface (kcal mol^{-1}) and transition state geometry for phosphoryl transfer from *p*-nitrophenyl phosphate to acetaldehyde hydrate.
72x34mm (600 x 600 DPI)



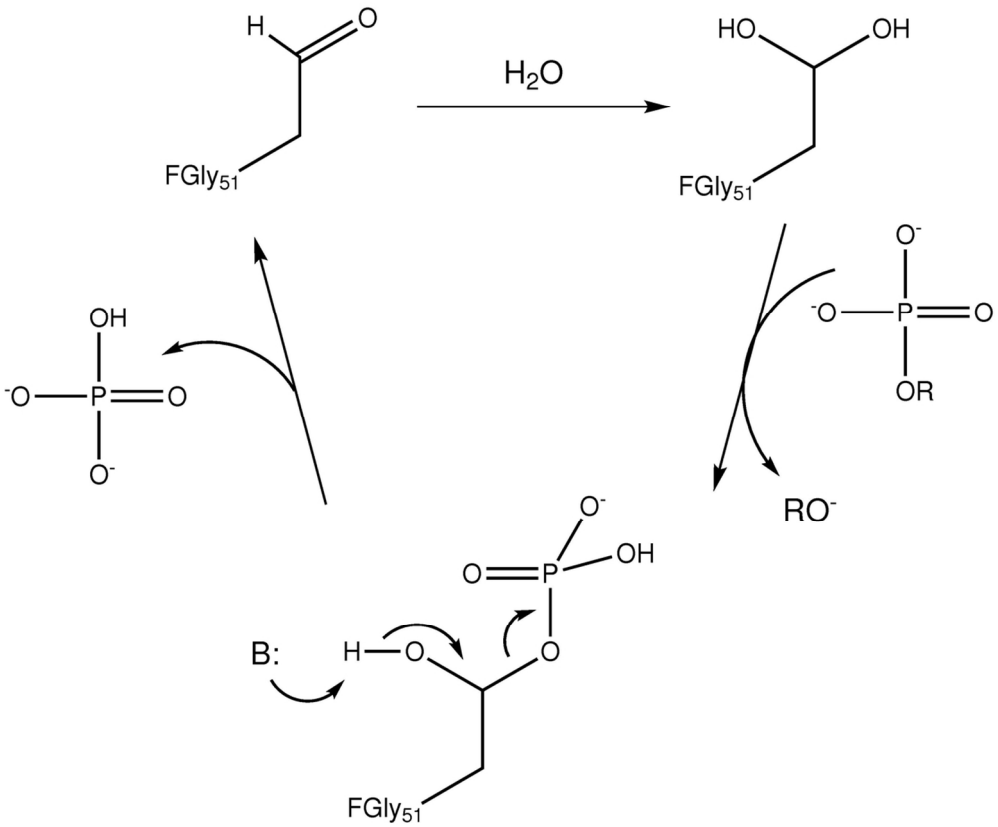
Empirical valence bond (EVB) free energy profiles for the phosphoryl transfer step in PAS (blue) and in solution (black).
84x46mm (600 x 600 DPI)



Representative transition state geometry for nucleophlic attack on *p*-nitrophenyl phosphate in Pseudomonas aeruginosa arylsulfatase. This corresponds to an A_ND_N substrate-as-base mechanism in which P-O formation to the nucleophile and cleavage to the leaving group occur in a single transition state, with concomitant proton transfer from the nucleophile to the phosphate (the proton can be seen in flight in both transition states). Note that the actual overall activation and free energies presented in Table 3 are obtained by averaging over 10 discrete trajectories.
119x101mm (600 x 600 DPI)



The linear response approximation (LRA) group contributions to the overall activation barrier in the enzyme for all residues during (a) the overall S_N2 reaction with concomitant proton transfer to the substrate, and (b) the isolated initial proton transfer step in the substrate-as-base mechanism.
148x213mm (300 x 300 DPI)



Revised mechanism for the promiscuous phosphomonoesterase activity of Pseudomonas aeruginosa arylsulfatase (PAS) (see also Fig. 3). The phosphoryl transfer reaction (step 2→3), which is the focus of the present work, is proposed to proceed via a concerted substrate-as-base mechanism, where a proton from the attacking nucleophile is donated to the phosphate (resulting in the attack of an anionic nucleophile on a phosphate monoanion), and without corresponding protonation of the departing *p*-nitrophenyl leaving group.
112x92mm (300 x 300 DPI)