

# **Leaving Group Ability Observably Affects Transition State Structure in a Single Enzyme Active Site**

Daniel Roston, Darren Demapan, and Qiang Cui\*

*Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin,  
Madison, WI 53706*

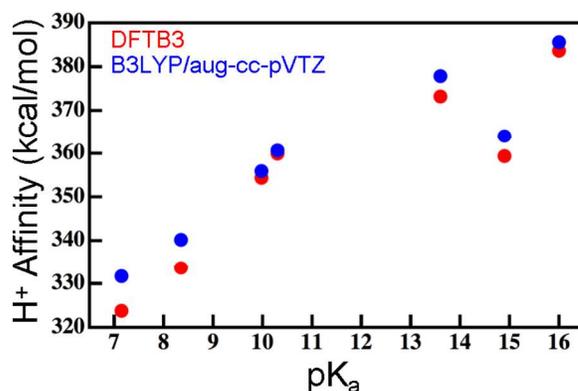
Email: [cui@chem.wisc.edu](mailto:cui@chem.wisc.edu)

## **Supporting Information**

## Text S1

### Computational Benchmarks

Since we sought to understand the effect of leaving group (LG) on phosphoryl transfer, we first test the ability of DFTB3<sup>1-3</sup> to reproduce the results of higher level density functional theory (DFT)<sup>4,5</sup> for relevant properties of the LGs; our previous studies<sup>1-3</sup> indicated that B3LYP/aug-cc-pVTZ serves as an adequate reference method in this regard. DFTB3 calculations were done with CHARMM<sup>6,7</sup> and B3LYP calculations were done with Gaussian09.<sup>8</sup> We calculated gas phase proton affinities for the LGs studied here using DFTB3 and compared those with the proton affinities calculated at the B3LYP/aug-cc-pVTZ level and found good agreement in the trend among the LGs (Figure S1), suggesting that DFTB3 is likely able to capture trends in reactivity of phosphoryl transfer for a series of LGs.



**Figure S1:** Gas phase proton affinities for the alkyl- and aryl-oxide leaving groups studied here as a function of experimental pK<sub>a</sub>.

To test the ability of DFTB3 to capture trends in LG properties within a QM/MM framework, we calculated relative pK<sub>a</sub>s for the conjugate acids of the LGs studied here. To do so, we followed the alchemical free energy perturbation methods described in ref. <sup>9</sup>. Formally, the free energy of aqueous deprotonation ( $\Delta G_{\text{AH}/\text{A}}$ ) is calculated in multiple steps

using the thermodynamic cycle in scheme S1. In this process, one mutates the acidic proton to a dummy atom ( $\Delta G_1$ ); transfers the dummy atom to the gas phase ( $\Delta G_2$ ); mutates it back to a proton in the gas phase ( $\Delta G_3$ ); and dissolves the proton in water ( $\Delta G_4$ ). One must also include the difference in vibrational zero point energy between the alcohol and alkoxide ( $\Delta ZPE$ ). This gives  $\Delta G_{AH/A}$  as:

$$\Delta G_{AH/A} = \Delta G_1 + \Delta G_2 + \Delta G_3 + \Delta G_4 + \Delta ZPE \quad (1)$$

$\Delta G_3$  is equal to zero and ref. <sup>9</sup> showed that  $\Delta G_2$  is negligible. Furthermore, since  $\Delta G_4$  (the solvation free energy of the proton) is invariant with respect to the acid, one can calculate relative free energies of deprotonation between two LGs (LG<sub>1</sub> and LG<sub>2</sub>) as

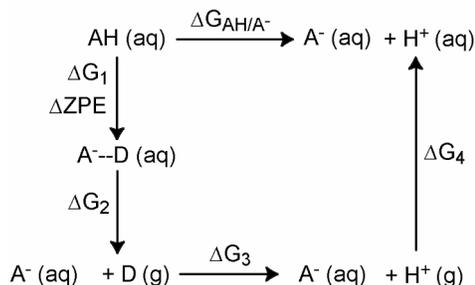
$$\Delta \Delta G_{AH/A}^{LG1-LG2} = (\Delta G_1^{LG1} + \Delta ZPE^{LG1}) - (\Delta G_1^{LG2} + \Delta ZPE^{LG2}) \quad (2)$$

and the relative  $pK_a$  as

$$\Delta pK_a^{LG1-LG2} = \Delta \Delta G_{AH/A}^{LG1-LG2} / (2.303RT) \quad (3)$$

Since we are interested in trends in the LGs, the relative  $pK_a$  is our primary interest. Thus, for each LG, we needed only to calculate  $\Delta G_1$  and  $\Delta ZPE$ . The mutation for  $\Delta G_1$  was accomplished using a DFTB3/MM simulation scheme similar to that used for the ground state of the KIE calculations in the main text. The alcohol/alkoxide plus its first solvation shell around the acidic oxygen were treated with DFTB3, and the rest of the waters were TIP3P;<sup>10</sup> the simulation boundary is defined by the GSBP protocol.<sup>11</sup> The alchemical mutation of the proton to the dummy atom occurred in 6 steps with the coupling parameter  $\lambda$  being varied from 0 to 1 at a constant interval of 0.2. The relevant energy gap (see ref. <sup>9</sup>) was averaged during 500 ps simulations at 298 K and  $\Delta G_1$  was determined by thermodynamic integration. Vibrational analysis in the gas phase determined  $\Delta ZPE$  for each LG. The results (Table S1) demonstrate that the DFTB3/MM scheme adequately

reproduces trends in  $pK_a$  of the LGs studied here and is therefore likely to reproduce trends in reactivity in the enzyme.



**Scheme S1:** Thermodynamic cycle used to calculate  $pK_a$ .

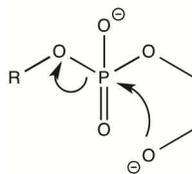
**Table S1:** Experimental and calculated<sup>a</sup>  $pK_a$ s for the conjugate acids of the LGs.

Acid	Experimental	Calculated
<i>p</i> -Nitrophenol	7.2	6.2
<i>m</i> -Nitrophenol	8.4	8.7
Phenol	10.0	11.7
<i>p</i> -Aminophenol	10.3	13.1
Propargyl alcohol	13.6	14.3
Ethanol	16.0	16.0

<sup>a</sup>Since the calculations were for relative  $pK_a$  values, these values were relative to benzyl alcohol, which was assigned the experimental  $pK_a$  of 15.4.

To test DFTB3's ability to reproduce trends in reactivity and TS structure for phosphoryl transfer, we employed the model transesterification reaction recently studied by York and co-workers (Scheme S2), which shows important differences in TS structure for different LGs.<sup>12</sup> We calculated TSs in the gas phase for a series of LGs and found that DFTB3 captures the trend and many details in TS structure that are obtained by B3LYP/aug-cc-pVTZ (Table S2). In particular, worse LGs have TSs that are later in the reaction coordinate (the same one defined in the main text for the AP reaction). DFTB3 captures the trend well, although the precise location of the breaking point from early TS to late TS differs slightly between DFTB3 and DFT. The propargyl oxide LG ( $pK_a=13.6$ ), for example, shows an early TS in DFTB3, but a late TS in DFT. Nonetheless, the results suggest

that DFTB3 overall captures trends in TS structure of phosphoryl transfers (see additional results below).



**Scheme S2:** The transesterification model reaction studied here (from ref. <sup>12</sup>).

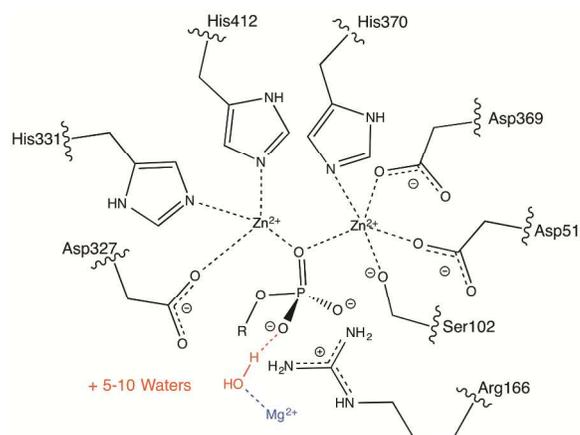
**Table S2:** TS geometries for the transesterification of scheme S1.

R-group	pKa	$\xi$		Tightness		P-O <sub>lg</sub>		P-O <sub>nuc</sub>	
		DFTB3	DFT <sup>a</sup>	DFTB3	DFT	DFTB3	DFT	DFTB3	DFT
acetyl	4.6	-0.45	-0.42	4.27	4.40	1.91	1.99	2.36	2.41
p-nitrophenyl	7.15	-1.00	-0.71	4.64	4.75	1.82	2.02	2.82	2.73
m-nitrophenyl	8.4	-0.38	-0.57	4.18	4.47	1.90	1.95	2.28	2.52
phenyl	10	-0.42	-0.43	4.26	4.37	1.92	1.97	2.34	2.40
p-aminophenyl	10.3	-0.42	-0.41	4.24	4.33	1.91	1.96	2.33	2.37
propargyl	13.6	-0.42	0.38	4.10	4.08	1.84	2.23	2.26	1.85
methyl	15.5	0.46	0.53	4.14	4.21	2.30	2.37	1.84	1.84
allyl	15.5	0.43	0.40	4.09	4.10	2.26	2.25	1.83	1.85
ethyl	16	0.44	0.40	4.12	4.10	2.28	2.25	1.84	1.85

<sup>a</sup>B3LYP/aug-cc-pVTZ

Next we sought to test the method within an active site model of AP, which better stabilizes charges of the reactive groups. The active site model was similar to the one used in ref. <sup>3</sup>, and included the substrate and the active site residues in Figure S2 (water molecules are excluded to avoid multi-minima issues during geometry optimization). The enzymatic residues were truncated at their beta carbons, which were turned into methyl groups and held frozen to their crystal structure positions. Using a range of monoester and diester substrates (Table S3), we performed relaxed potential energy surface scans along the reaction coordinate using DFTB3. The highest energy point along the reaction coordinate was chosen as the TS and for each substrate it had a RMS gradient of less than 0.03 kcal/mol•Å. This structure was used as the starting point for an optimization to a TS (using Gaussian09) at the B3LYP/6-31+G\* level, using an effective core potential<sup>13</sup> for the Zn<sup>2+</sup> ions. Table S3 indicates that DFTB3 captures the trends in TS structure for both

monoesters and diesters, including the case with the propargyl oxide LG. As with the transesterification above, worse LGs have later TSs. We note that the monoesters show significantly less TS variation than the diesters or the model reaction above; the TSs for monoesters are all early. The shift in TS structure from early to late indicates a shift in the relative energetic cost of forming the P-O<sub>nuc</sub> bond vs. breaking the P-O<sub>lg</sub> bond. That is, the early part of the reaction pathway is P-O<sub>nuc</sub> bond formation, while the latter is P-O<sub>lg</sub> bond cleavage. For early TSs, the former is a costlier process and for late TSs, the latter is costlier. Thus, the fact that there is no large shift in the TS of the monoesters in the active site model indicates that the active site model does not stabilize P-O<sub>nuc</sub> bond formation relative to P-O<sub>lg</sub> bond cleavage as well as it does for diesters. The formation of the P-O<sub>nuc</sub> bond requires the nucleophilic oxygen anion to approach the negatively charged phosphate group. The additional charge on the phosphate group of the monoester makes the process particularly costly for that series, so despite the fact that worse LGs still destabilize the second process, the first part of the reaction coordinate remains costlier. The observation that significant variation occurs in the TS structure in AP can be interpreted to reflect the enzyme's ability to stabilize both the P-O<sub>nuc</sub> bond formation and P-O<sub>lg</sub> bond cleavage.



**Figure S2:** Key active site motifs of AP with a phosphate monoester substrate. The residues in black constituted the active site model. The additional waters, in red, were included in the QM region of the DFTB3/MM simulations. The  $Mg^{2+}$  (and the rest of the protein and solvent) were treated with the CHARMM force field,<sup>14,15</sup> within the GSBP framework,<sup>11</sup> as described in the main text. The QM region includes ca. 120 atoms (depending on substrate) and an overall charge of -1 (for WT/monoesters; it differs for diesters and R166S). The number of QM waters varied slightly among the substrates and mutants studied here, and generally included the first solvation shell surrounding the phosphate moiety. In the R166S mutant, the S166 sidechain was not included in the QM region. Test calculations that included S166 in the QM region confirmed that it made negligible differences in the calculated TS structure (cf. Figure S9).

**Table S3:** TS geometries in the active site model of AP.

Substrate	LG pK <sub>a</sub>	O <sub>lg</sub> -P		O <sub>nuc</sub> -P		ξ		Tightness	
		DFTB3	DFT <sup>a</sup>	DFTB3	DFT	DFTB3	DFT	DFTB3	DFT
<i>Monoesters</i>									
mNPP	8.4	1.89	1.89	2.23	2.35	-0.35	-0.46	4.12	4.24
PhOP	10.0	1.83	1.83	2.18	2.32	-0.35	-0.49	4.01	4.15
pAPP	10.3	1.82	1.82	2.17	2.36	-0.35	-0.54	3.99	4.18
PrAP	13.6	1.79	1.80	2.13	2.28	-0.34	-0.48	3.92	4.08
mNBP	14.9	1.80	1.80	2.15	2.30	-0.35	-0.50	3.95	4.10
EtOP	16.0	1.78	1.80	2.11	2.26	-0.33	-0.46	3.89	4.06
<i>Diesters<sup>b</sup></i>									
mpNPP	7.2	1.91	1.99	2.21	2.22	-0.3	-0.23	4.12	4.21
mmNPP	8.4	1.86	1.96	2.14	2.08	-0.28	-0.12	4.00	4.04
mPhOP	10.0	2.14	2.06	1.86	1.99	0.28	0.07	4.00	4.05

<sup>a</sup>B3LYP/6-31+G\*. <sup>b</sup>From ref. 3.

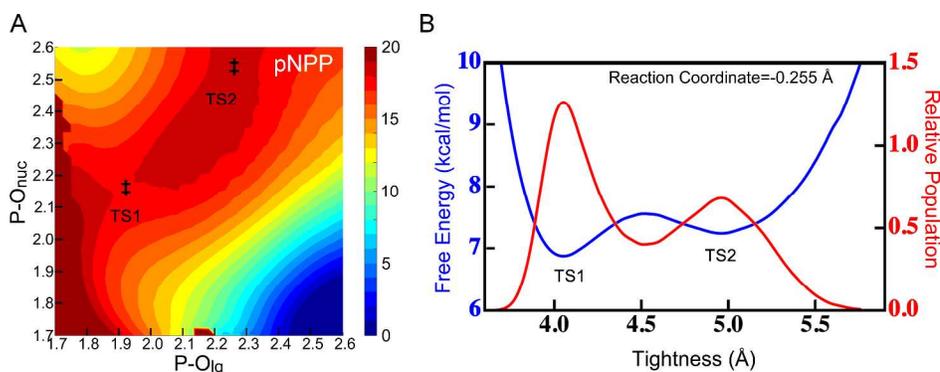
To examine the upper limits of KIEs on O<sub>lg</sub> for phosphate ester hydrolysis, we calculated the relevant equilibrium isotope effect (EIE) for ethyl phosphate hydrolysis. EIEs represent the situation of complete bond cleavage and therefore serve as an upper limit for the corresponding KIE.<sup>16</sup> We calculated vibrational frequencies for ethyl phosphate and ethyl alkoxide in implicit solvent<sup>17</sup> at the B3LYP/6-31+G\* level and computed the EIEs according to the Bigeleisen equation<sup>18,19</sup> using the program ISOEFF.<sup>20</sup> The calculated EIE is 1.022, which is somewhat smaller than previous estimates of this limit.<sup>21-23</sup> The

discrepancy appears to stem from the fact that the previous estimates were based solely on the P-O stretching frequency. In our calculations, too, if we assume that the loss of the P-O stretching frequency ( $707\text{ cm}^{-1}$ ) is all that matters, we obtain an upper limit around 1.07. But this ignores other contributing modes, such as the C-O stretching mode, which *increases* in frequency from  $1052\text{ cm}^{-1}$  to  $1330\text{ cm}^{-1}$  in moving from ethyl phosphate to ethoxide. This finding emphasizes the importance of including all the vibrational modes in computing KIEs; the effect of the “primary” mode can be severely muted by other contributing modes.

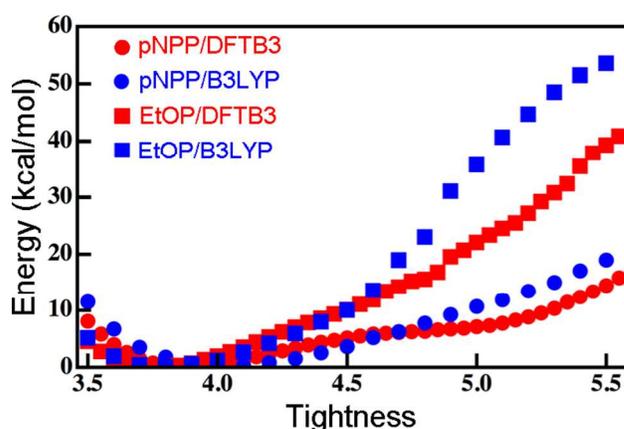
## **Text S2**

### *p-Nitrophenyl Phosphate*

Since many experimental and theoretical studies of phosphate ester hydrolysis in AP and in other enzymes<sup>24,25</sup> use p-nitrophenyl phosphate (pNPP), that substrate would make a logical choice to include in our study. In preliminary examinations of that substrate, though, we obtained results that were not consistent with the trends we observed among the other substrates. In particular, the free energy along the tightness coordinate in the TS region was very flat, leading to both a tight and a loose TS of similar free energies (Figure S3). We examined this effect in our active site model, where we could compare DFTB3 with DFT and the results (Figure S4) suggest that although DFTB3 share qualitatively similar features as DFT, it underestimates the energetic cost of stretching the TS in the loose direction vs. DFT and this leads to the looser TS structure observed in the simulations. Therefore, we leave more detailed analysis of pNPP using DFT/MM free energy path simulations and refinement of the DFTB3 model for future work.



**Figure S3:** A) The free energy surface for pNPP is very flat in the TS region, leading to two TSs with similar free energies. B) A 1-D slice of the free energy surface along the tightness coordinate shows that both a tight and loose TS are significantly populated at 298 K, based on a Boltzmann population distribution.



**Figure S4:** PES along the tightness coordinate in the TS region ( $\xi=0.0$  Å) of the active site model of pNPP and EtOP. The red points correspond to a relaxed PES scan using DFTB3 and the blue points are DFT calculations on the DFTB3 geometries. For both substrates, DFTB3 underestimates the energy at the loose end, but since the energy for pNPP is closer to the range of thermally accessible energies, that effect has more important implications for the structures sampled during a simulation at 298 K.

## Text S3

### *A Model of Curved FERs*

Curved free energy relationships (FERs) are believed to be rare<sup>26,27</sup> and classically result from either a change in mechanism or a change in rate-limiting step. The fact that  $^{18}(\text{V}/\text{K})_{\text{lg}}$  is significantly greater than unity across the range of LGs in the R166S mutant<sup>23</sup> argues against both of those hypotheses, but the two FERs in that mutant (Figure 4) show an obvious curvature.<sup>28-30</sup> In the main text, we have argued that this results from a change in TS structure for the same chemical step and that such a change is a reasonable

expectation from Hammond's Postulate.<sup>31</sup> Some have even pointed out that in principle, all FERs should exhibit curvature due to changes in the TS for each substrate. Why, then, are conspicuously curved FERs seemingly rare without a change in mechanism or rate-limiting step?

As mentioned in the text, one possibility is that they are not, in fact, rare; they have simply been misattributed to other effects or otherwise ignored. To investigate this question further, we have explored reactivity and TS structure as a function of reaction thermodynamics using a generalized LEPS potential energy surface.<sup>32</sup> This surface can be thought of as three coupled Morse potentials (one for the P-O<sub>lg</sub> distance, one for P-O<sub>nuc</sub>, and one for O<sub>lg</sub>-O<sub>nuc</sub>) representing a linear O<sub>lg</sub>-P-O<sub>nuc</sub>. The overall potential (V) is:

$$V = Q_{P-O_{lg}} + Q_{P-O_{nuc}} + Q_{O_{lg}-O_{nuc}} - \sqrt{\frac{1}{2} \left( \left( \frac{J_{P-O_{lg}}}{1+s_{P-O_{lg}}} - \frac{J_{P-O_{nuc}}}{1+s_{P-O_{nuc}}} \right)^2 + \left( \frac{J_{P-O_{nuc}}}{1+s_{P-O_{nuc}}} - \frac{J_{O_{lg}-O_{nuc}}}{1+s_{O_{lg}-O_{nuc}}} \right)^2 + \left( \frac{J_{O_{lg}-O_{nuc}}}{1+s_{O_{lg}-O_{nuc}}} - \frac{J_{P-O_{lg}}}{1+s_{P-O_{lg}}} \right)^2 \right)} \quad (4)$$

where

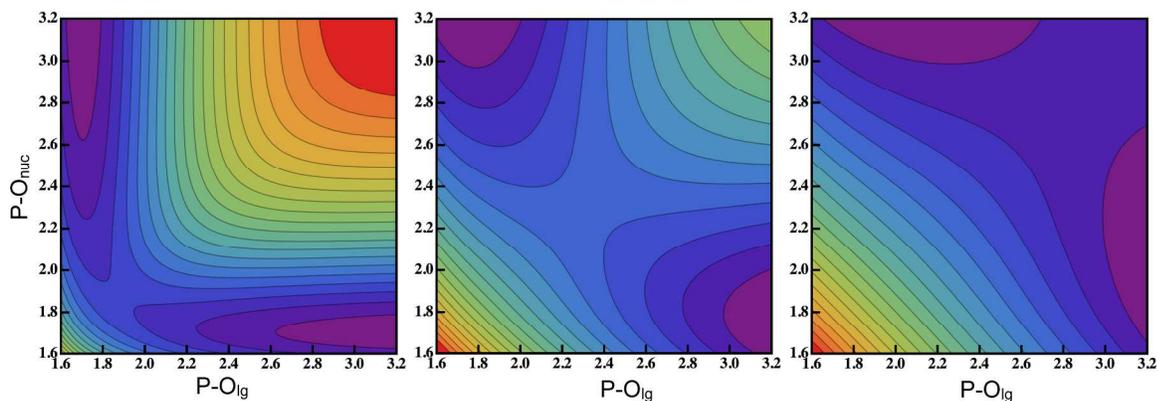
$$Q = \frac{1}{2} d \left( \frac{3}{2} e^{-2(\beta(r-r_0))} - e^{-\beta(r-r_0)} \right) \quad (5)$$

and

$$J = \frac{1}{2} d \left( \frac{1}{2} e^{-2(\beta(r-r_0))} - 3e^{-\beta(r-r_0)} \right) \quad (6)$$

Thus, each of the three interactions is described by 4 parameters:  $r_0$ , the position of the well minimum;  $\beta$ , the Morse parameter determining the width of the well;  $d$ , the dissociation energy; and  $s$ , the Sato parameter, which controls the barrier height. We can control the tightness of the reaction pathway by modulating  $\beta$  while holding other parameters constant (Figure S5). We have explored TS structure as a function of  $\Delta E_{\text{reaction}}$  (one-half the difference in  $d$  for the two P-O bonds) on three surfaces—tight, synchronous,

and loose—by changing the dissociation energy of the product well, while holding that of the reactant well constant at 10 kcal/mol. This model assumes that a change in well depth of the P-O<sub>nuc</sub> bond does not cause any changes in the other parameters defining the PES. Changing the depth of the P-O<sub>nuc</sub> well models the changing stability of the product that occurs across a series of LGs in an FER measurement. As indicated in Figure S6, the bond length to the LG at the TS may change when the dissociation energy of the product changes. For exothermic reactions, P-O<sub>lg</sub> at the TS is generally longer than for endothermic reactions. What is necessary to observe curved FERs is for the change in product well depth to result in a change from a still-formed P-O<sub>lg</sub> bond at the TS to a broken P-O<sub>lg</sub> bond. Our explorations of these surfaces indicate that such a scenario is possible with thermoneutral or exothermic reactions, depending on the level of tightness.

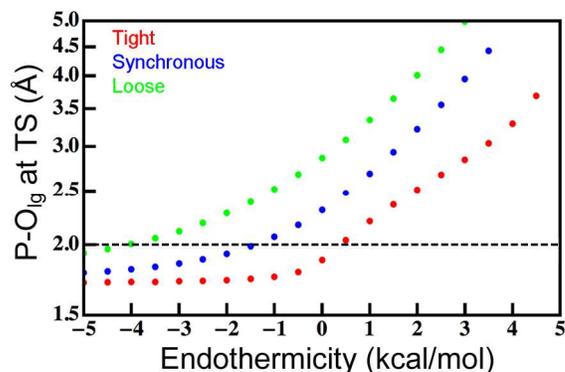


**Figure S5:** Generalized LEPS potential energy surfaces for symmetric ( $\Delta E_{\text{reaction}}=0$ ) reactions corresponding to tight ( $\beta=2 \text{ \AA}^{-1}$ , left), synchronous ( $\beta=0.8 \text{ \AA}^{-1}$ , middle), and loose ( $\beta=0.5 \text{ \AA}^{-1}$ , right) pathways. Other parameters are the same for all three surfaces and summarized in Table S4. The position of the TS along the reaction pathway changes when  $d$  is changed for the product well (the stability of the product changes).

**Table S4:** Summary of parameters used for the LEPS analysis<sup>a</sup>

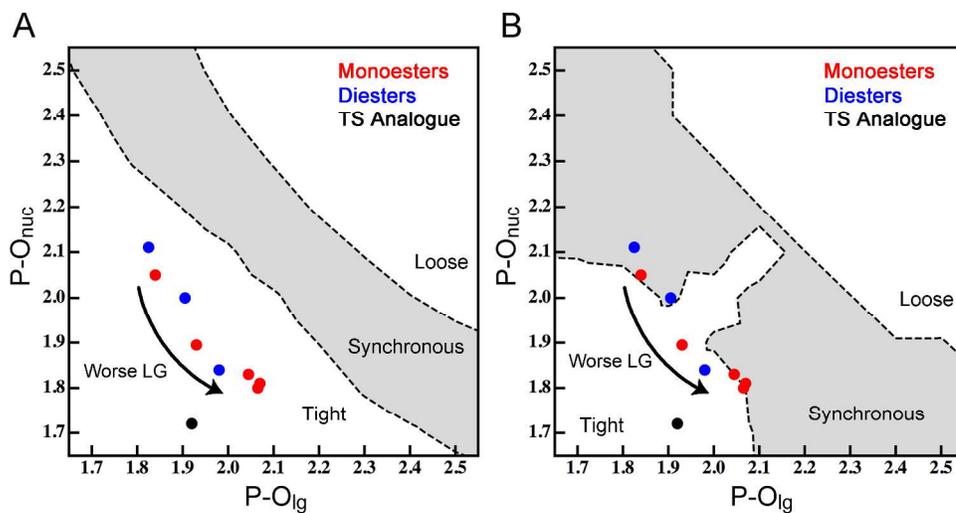
Parameter	Tight	Synchronous	Loose
$d$ (kcal/mol)	10 <sup>b</sup>	10 <sup>b</sup>	10 <sup>b</sup>
$\beta$ ( $\text{\AA}^{-1}$ )	2.0	0.8	0.5
$r_0$ ( $\text{\AA}$ )	1.7	1.7	1.7
$s$	1	1	1

<sup>a</sup>Except where noted, the parameters were equal for all three interactions in the model. <sup>b</sup>This value refers to the dissociation energy for the P-O<sub>lg</sub> bond and the O<sub>lg</sub>-O. That for the P-O<sub>nuc</sub> bond was varied around 10 to alter the reaction's exothermicity.

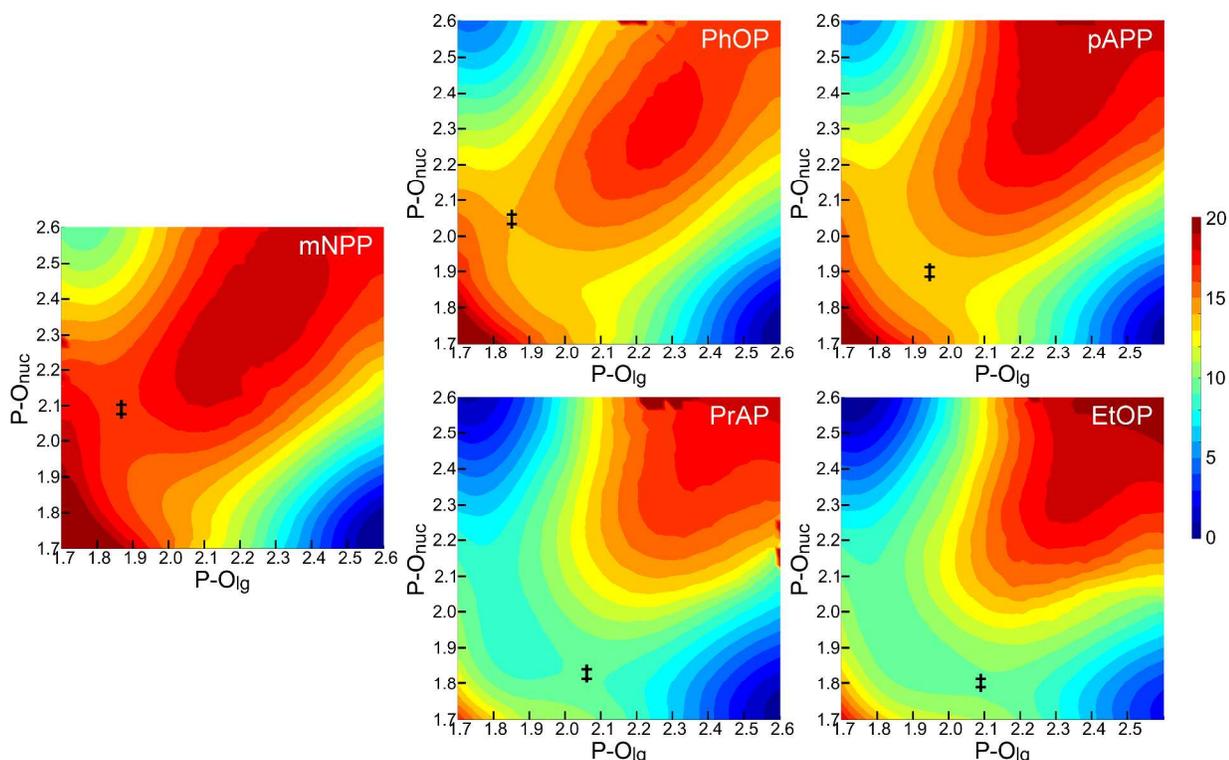


**Figure S6:** The length of the P-O<sub>lg</sub> bond at the TS as a function of a reaction's driving force for reaction pathways with three levels of tightness. These data result from a simple model of reactivity based on a generalized LEPS potential (see Text S3). The dashed line indicates the approximate P-O<sub>lg</sub> length at which one might consider that bond to be broken (loss of 50% bond order; cf. Table S5). To observe a curved FER and corresponding changes in KIEs, the LGs must span a range of endothermicities where the TS for some LGs has a broken P-O<sub>lg</sub> bond, but the TS for other LGs does not. Depending on tightness, this appears to be possible so long as the reaction is not significantly endothermic.

## Additional Figures and Tables



**Figure S7:** The TS region of the phosphate ester hydrolysis reaction, showing the various TS structures found in this work, as well as the experimentally determined structure of a TS analogue, vanadate (VO<sub>4</sub>).<sup>33</sup> The guides as to the tight, synchronous, and loose regions are determined from the sum of bond order using either Wiberg Bond Order (A) or Natural Bond Order (B) of the P-O<sub>lg</sub> and P-O<sub>nuc</sub> bonds. Regions where that sum is less than 0.9 are loose, between 0.9 and 1.1 are synchronous, and greater than 1.1 are tight. One can see that NBO is particularly sensitive to coupling between the two bonds and bond order is particularly high for the region of symmetry, where the bonds are equal length. Bond order was calculated for the symmetric phosphoryl transfer between two methoxide moieties in implicit solvent{Marenich, 2009 #185} at the B3LYP/aug-cc-pVTZ level. P-O<sub>lg</sub> and P-O<sub>nuc</sub> distances were constrained to points within a grid to cover the space shown here.

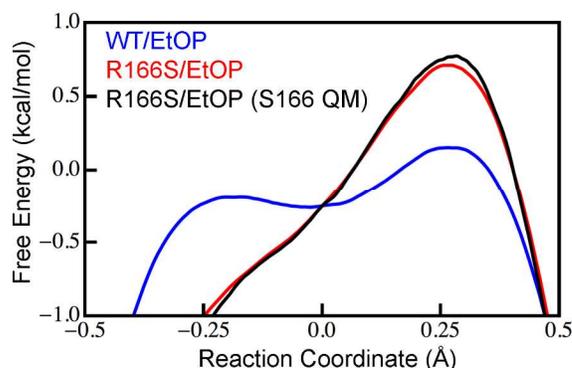


**Figure S8:** PMFs for the hydrolysis of a series of phosphate monoesters (phenyl phosphate, PhOP; p-aminophenyl phosphate, pAPP; propargyl phosphate, PrAP; and ethyl phosphate, EtOP) in AP. The reaction follows a slightly tight pathway and the position of the TS (‡) along that pathway changes as a function of LG ability. Worse LGs (higher  $pK_a$  of the conjugate acid) have later (more product-like) TSs. Free energies are in kcal/mol and the same scale corresponds to all substrates.

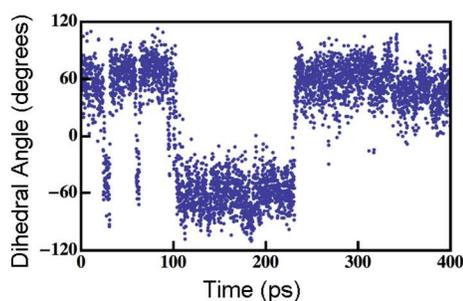
**Table S5:** Calculated TS geometries<sup>a</sup> and bond orders

Substrate	LG $pK_a$	$\xi^b$	P-O <sub>lg</sub>			P-O <sub>nuc</sub>			Tightness					
			Length	WBO <sup>c</sup>	NBO <sup>d</sup>	WBO <sup>e</sup>	Length	WBO <sup>c</sup>	NBO <sup>d</sup>	WBO <sup>e</sup>	Length	WBO <sup>c</sup>	NBO <sup>d</sup>	WBO <sup>e</sup>
<i>Monoesters</i>														
mNPP	8.4	-0.22	1.87	0.75	0.95	0.73	2.09	0.47	0.04	0.45	3.96	1.22	0.99	1.18
PhOP	10.0	-0.18	1.87	0.69	0.96	0.73	2.05	0.47	0.50	0.49	3.91	1.16	1.46	1.22
pAPP	10.3	0.03	1.94	0.60	0.67	0.62	1.91	0.65	0.65	0.66	3.85	1.26	1.32	1.28
PrAP	13.6	0.22	2.05	0.49	0.04	0.49	1.83	0.79	0.96	0.76	3.89	1.29	1.00	1.25
mNBP	14.9	0.26	2.07	0.46	0.04	0.47	1.81	0.77	0.96	0.79	3.88	1.26	1.00	1.25
EtOP	16.0	0.27	2.08	0.43	0.04	0.46	1.81	0.78	0.96	0.79	3.89	1.21	1.00	1.24
<i>Diesters</i>														
mpNPP	7.2	-0.30	1.82	0.70	0.83	0.85	2.11	0.53	0.06	0.50	3.94	1.22	0.89	1.35
mmNPP	8.4	-0.16	1.87	0.59	0.85	0.79	2.02	0.58	0.69	0.61	3.89	1.22	1.53	1.40
mPhOP	10.0	0.14	1.98	0.48	0.73	0.64	1.84	0.80	0.84	0.82	3.83	1.28	1.58	1.46

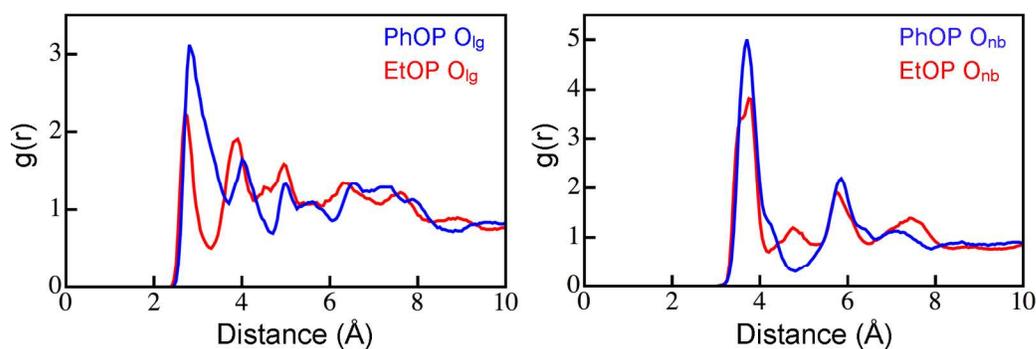
<sup>a</sup>Bond lengths are averages during at least 400 ps trajectories constrained to the listed value of  $\xi$  using a harmonic force constant of 215 kcal/mol $\cdot\text{\AA}^2$ . <sup>b</sup>Reaction coordinate at the TS, determined from the calculated PMFs. <sup>c</sup>Wiberg bond order, given as a fraction of the bond order in the reactant or product for P-O<sub>lg</sub> and P-O<sub>nuc</sub>, respectively. The values were calculated at the B3LYP/6-31+G\*/LANL2DZ level in implicit solvent<sup>17</sup> for the QM region. Each value is the average of 10 snapshots from simulations separated by 50 ps. <sup>d</sup>Natural Bond Order, calculated for the symmetric transfer of a phosphoryl (or methyl phosphoryl) group between two methoxide moieties at the B3LYP/aug-cc-pVTZ level in implicit solvent<sup>17</sup> with the P-O<sub>lg</sub> and P-O<sub>nuc</sub> bonds restrained at the given distances. <sup>e</sup>Wiberg Bond Order calculated using the same geometries and level of theory as the NBO calculations.



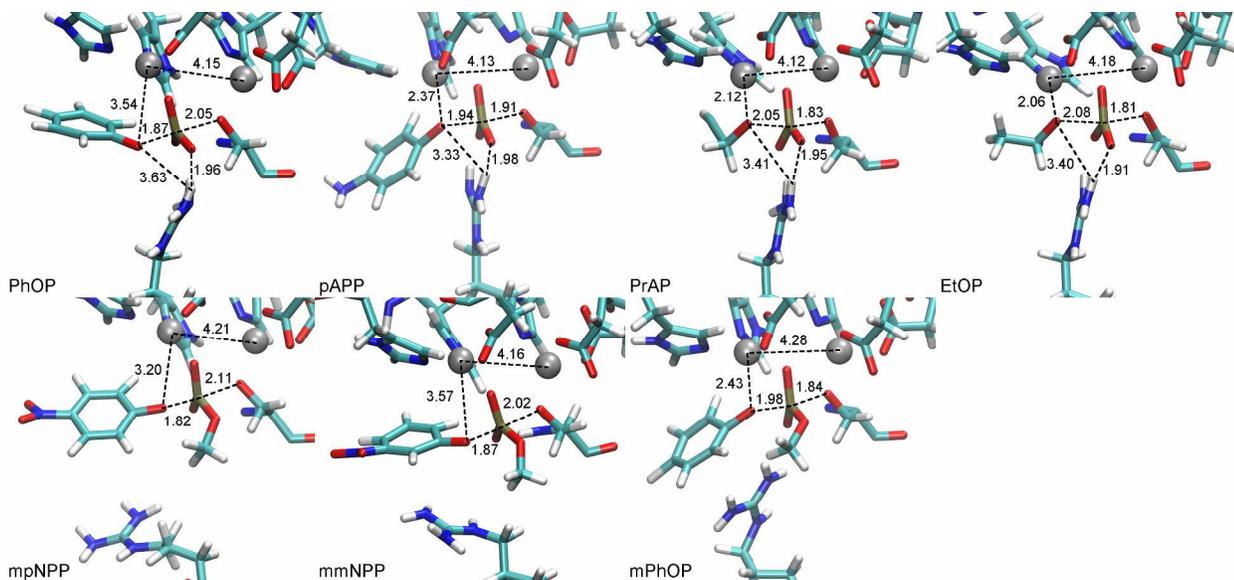
**Figure S9:** 1D PMFs of the TS region in WT and R166S with EtOP as the substrate. The R166S mutation has little effect on the position of the TS, but it does affect the overall shape of the PMF in the TS region. Poor substrates in the WT enzyme traverse a relatively flat plateau in the TS region, where the peak of that plateau is late in the reaction coordinate (cf. Figures 1 and S8 and S13). The TS is more narrowly defined in R166S. It is intriguing that the mutation affects the shape of the PMF without ultimately affecting the location of the TS. The positions of the R166 sidechain are occupied by active site water molecules in the mutant, which are able to provide similar hydrogen bonding partners to the substrate at the TS. Apparently these interactions do not preserve the overall shape of the PMF, but generally preserve the position of the TS along that PMF. We note that the overall similarity of the TS in the WT and mutant has been observed experimentally, as well.<sup>29</sup> One can also see from this figure that excluding S166 from the QM region of simulations of the mutant (Red includes it, black excludes it) has negligible effects on the position of the TS or the shape of the PMF in this region.



**Figure S10:** The dihedral angle ( $O_{lg}-P-O2-C_{CH3}$ ) describing conformational sampling of the methyl substituent in mPhOP during a 400 ps trajectory. In solution this dihedral would presumably sample a third conformation, at  $180^\circ$ , but steric hindrance from S102 destabilizes that conformation in the enzyme.



**Figure S11:** Radial distribution functions (RDFs) from the specified atom to water at the TS of the reactions of PhOP and EtOP.  $O_{lg}$  is somewhat better solvated at the TS for PhOP than for EtOP. In EtOP, and other substrates with poor LGs,  $O_{lg}$  ligates one of the catalytic  $Zn^{2+}$  ions; for good LGs, ligation to the  $Zn^{2+}$  is less important (cf. Figures 2 and S12). Water molecules also stabilize the non-bridging phosphoryl oxygens ( $O_{nb}$ ), which also interact directly with R166.

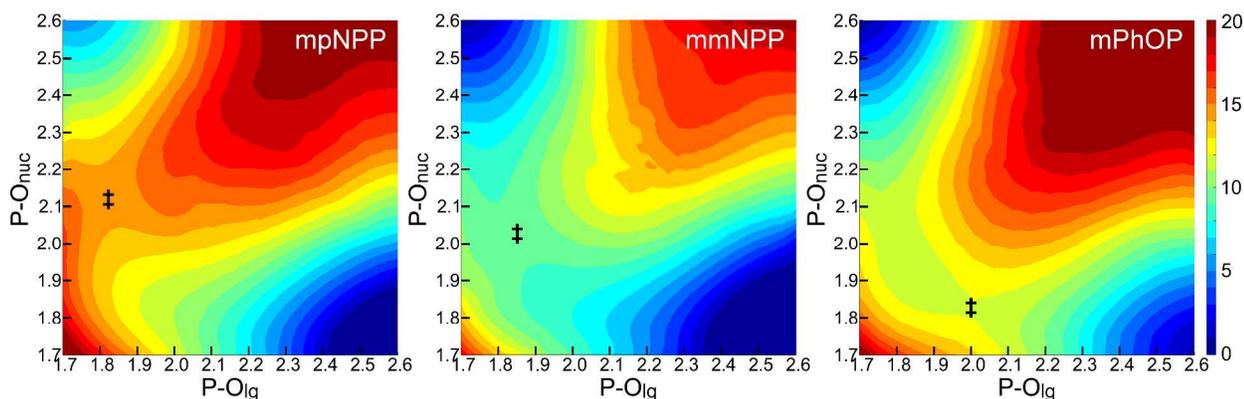


**Figure S12:** Representative snapshots from the TS region for the substrates studied here. Distances are in Å and are averages during at least 400 ps trajectories with the reaction coordinate constrained to the TS value using a harmonic force constant of 215 kcal/mol·Å<sup>2</sup>.

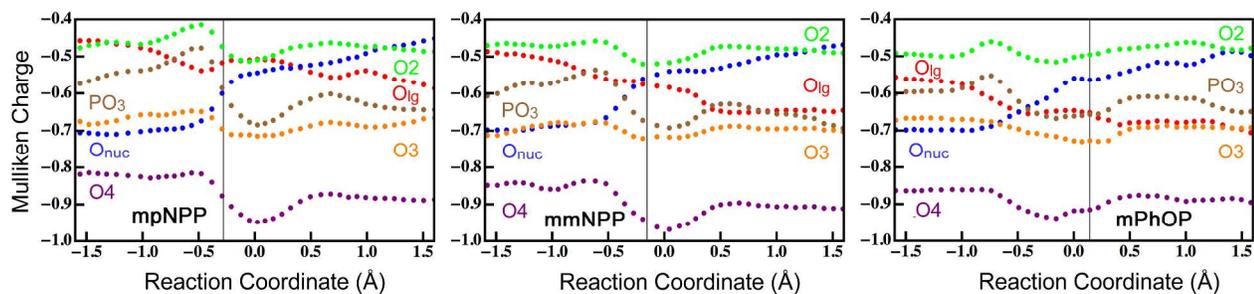
**Table S6:** Difference in natural atomic charges and Mulliken charges between the reactant and TS for monoesters

Substrate	O <sub>nuc</sub>		O <sub>lg</sub>		PO <sub>3</sub>	
	Natural <sup>a</sup>	Mulliken <sup>b</sup>	Natural <sup>a</sup>	Mulliken <sup>b</sup>	Natural <sup>a</sup>	Mulliken <sup>b</sup>
mNPP	+0.05	+0.11	-0.01	-0.08	-0.02	-0.04
PhOP	+0.02	+0.09	+0.02	-0.09	-0.05	-0.04
pAPP	+0.06	+0.11	-0.04	-0.03	-0.05	-0.06
PrAP	+0.08	+0.10	-0.05	-0.05	-0.06	-0.06
EtOP	+0.08	+0.12	-0.07	-0.01	-0.04	-0.05

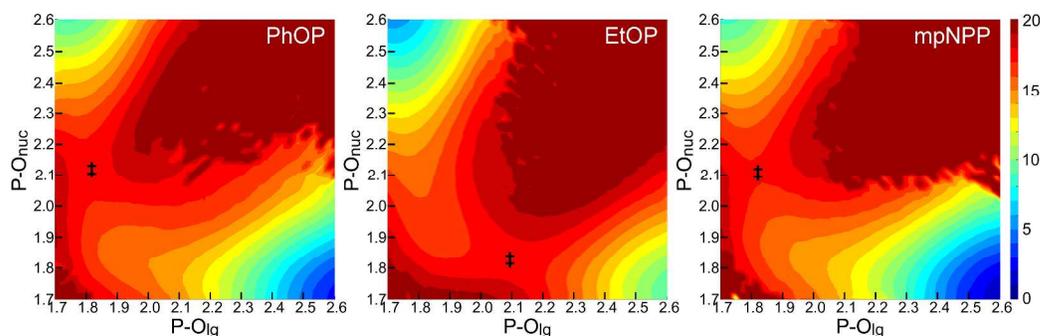
<sup>a</sup>Each entry is the difference between the reactant ( $\xi=-1.500$ ) and the TS ( $\xi$  listed in Table 1 of the main text) in the average calculated for 10 snapshots, each separated by 50 ps. NBO analysis was done for the QM region using the same B3LYP method and basis set as for the active site model. <sup>b</sup>From the analysis in Figure 3 of the main text using DFTB3/MM.



**Figure S13:** PMFs for the hydrolysis of a series of phosphate diesters (methyl p-nitrophenyl phosphate, mpNPP; methyl m-nitrophenyl phosphate, mmNPP; and methyl phenyl phosphate, mPhOP) in wt AP. As with the monoesters, the reaction follows a slightly tight pathway and the position of the TS (‡) along that pathway changes as a function of LG ability. Worse LGs (higher pK<sub>a</sub> of the conjugate acid) have later (more product-like) TSs. Free energies are in kcal/mol and the same scale corresponds to all substrates.



**Figure S14:** Changes in Mulliken charges during the reaction of phosphate diesters in WT AP. Each point represents the average of 1000 frames from multiple trajectories, binned according to reaction coordinate. O2, O3, and O4 are the phosphoryl oxygens; PO<sub>3</sub> (the phosphoryl group) is the sum of the charges on P, O2, O3, and O4. The vertical line indicates the position of the TS for each substrate.



**S15:** PMFs for the hydrolysis of phosphate esters in R166S AP. The positions of the TSs in the mutant are nearly identical to their positions in WT (cf. Figures S8 and S13). In contrast to other PMFs shown in the main text and SI, enhanced sampling along the tightness coordinate was not employed here thus high energy regions (colored dark red) were not sampled during the simulations; this does not have any impact on the key regions of interest or the location of TS.

## References

- (1) Gaus, M.; Cui, Q.; Elstner, M. *J. Chem. Theory Comput.* **2011**, *7*, 931.
- (2) Gaus, M.; Lu, X. Y.; Elstner, M.; Cui, Q. *J. Chem. Theory Comput.* **2014**, *10*, 1518.
- (3) Lu, X. Y.; Gaus, M.; Elstner, M.; Cui, Q. *J. Phys. Chem. B* **2015**, *119*, 1062.
- (4) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (5) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (6) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoseck, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
- (7) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
- (8) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Revision D.01 ed.; Gaussian, Inc.: Wallingford, CT., 2009.
- (9) Li, G. H.; Cui, Q. *J. Phys. Chem. B* **2003**, *107*, 14521.
- (10) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (11) Im, W.; Berneche, S.; Roux, B. *J. Chem. Phys.* **2001**, *114*, 2924.
- (12) Chen, H. Y.; Giese, T. J.; Huang, M.; Wong, K. Y.; Harris, M. E.; York, D. M. *Chem.-Eur. J.* **2014**, *20*, 14336.
- (13) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270.
- (14) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
- (15) Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400.
- (16) Hengge, A. C. *Accounts Chem. Res.* **2002**, *35*, 105.
- (17) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378.
- (18) Bigeleisen, J. *J. Chem. Phys.* **1949**, *17*, 675.

- (19) Bigeleisen, J.; Mayer, M. G. *J. Chem. Phys.* **1947**, *15*, 261.
- (20) Anisinov, V.; Paneth, P. Lodz, Poland, 2007.
- (21) Sawyer, C. B.; Kirsch, J. F. *J. Am. Chem. Soc.* **1973**, *95*, 7375.
- (22) Gorenstein, D. G.; Lee, Y. G.; Kar, D. *J. Am. Chem. Soc.* **1977**, *99*, 2264.
- (23) Zalatan, J. G.; Catrina, I.; Mitchell, R.; Grzyska, P. K.; O'Brien, P. J.; Herschlag, D.; Hengge, A. *J. Am. Chem. Soc.* **2007**, *129*, 9789.
- (24) Lassila, J. K.; Zalatan, J. G.; Herschlag, D. *Annu. Rev. Biochem.* **2011**, *80*, 669.
- (25) Cleland, W. W.; Hengge, A. C. *Chem. Rev.* **2006**, *106*, 3252.
- (26) Schreck, J. O. *J. Chem. Educ.* **1971**, *48*, 103.
- (27) Williams, A. *Free Energy Relationships in Organic and Bio-Organic Chemistry*; The Royal Society of Chemistry: Cambridge, 2003.
- (28) O'Brien, P. J.; Herschlag, D. *Biochemistry* **2002**, *41*, 3207.
- (29) O'Brien, P. J.; Herschlag, D. *J. Am. Chem. Soc.* **1999**, *121*, 11022.
- (30) Zalatan, J. G.; Herschlag, D. *J. Am. Chem. Soc.* **2006**, *128*, 1293.
- (31) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334.
- (32) Muckerman, J. T. *J. Chem. Phys.* **1972**, *56*, 2997.
- (33) Holtz, K. M.; Stec, B.; Kantrowitz, E. R. *J. Biol. Chem.* **1999**, *274*, 8351.