

Supporting Information: Molecular
mechanisms of phosphoester bond formation
in water using tight-binding ab-initio
molecular dynamics

Zakarya Benayad, Matthias Bova Saint-André, and Guillaume Stirnemann*

*CNRS Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, PSL
University, Université de Paris, 13 rue Pierre et Marie Curie, 75005, Paris, France*

E-mail: stirnemann@ibpc.fr

1. Phosphate monoanion

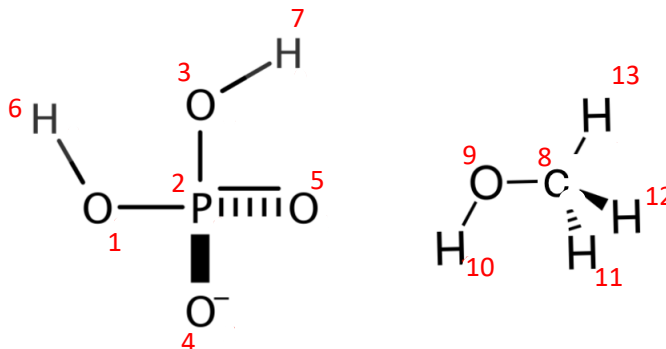


Figure S1: Identification of the atoms of phosphate monoanion and methanol from 1 to 13.

The following constrains were used in all simulations. They ensure that unrealistic molecular geometries do not occur during the metadynamics Gaussian deposition process.

Table S1: Definition of the molecular angles

Angles	α	β_1	γ_1	δ	β_2	γ_2	β_3	γ_3
Atomic index	$O_1P_2O_9$	$O_1P_2O_3$	$O_3P_2O_9$	$C_8O_9P_2$	$O_1P_2O_4$	$O_4P_2O_9$	$O_1P_2O_5$	$O_5P_2O_9$

Table S2: Constraints applied on angles or angle differences during the simulations.

Angles	Values a_i (rad)		Force constant k_i (kcal/mol/rad ²)
	Lower wall	Upper wall	
$\beta_1-\beta_2$	-0.5	0.5	400
$\beta_1-\beta_3$	-0.5	0.5	400
$\beta_3-\beta_2$	-0.5	0.5	400
$\gamma_1-\gamma_2$	-0.5	0.5	400
$\gamma_1-\gamma_3$	-0.5	0.5	400
$\gamma_3-\gamma_2$	-0.5	0.5	400
δ	1.57	/	500

Table S3: Constraints applied on distances or distance differences during the simulations.

Atomic index	Values a_i (Å)		Force constant k_i (kcal/mol/Å ²)
	Lower wall	Upper wall	
$d_{pl}: P_2O_1$	/	4	500
$d_{pm}: P_2O_9$	/	4	500
$d_{pl}-d_{pm}$	-2.4	2.4	400
O_3P_2	/	1.8	500
O_4P_2	/	1.8	500
O_5P_2	/	1.8	500
$H_{11}C_8$	/	1.25	500
$H_{12}C_8$	/	1.25	500
$H_{13}C_8$	/	1.25	500
O_9C_8	1.35	1.6	500
O_1H_6	/	1.2	500

Table S4: Constraints applied on the numbers of coordination between two groups of atoms (A and B) during the simulations.

Group A	Group B	Values a_i		Force constant k_i (kcal/mol)
		Lower wall	Upper wall	
O_1	H_6, H_7, H_{10} and the H of the solvent	0.8	/	500
O_9	H_6, H_7, H_{10} and the H of the solvent	/	1.5	500
O_3	The H of the solvent	/	0.5	500
O_4	The H of the solvent	/	0.5	500
H_7	O_4, O_5, O_9 and the O of the solvent	/	0.5	500
H_{10}	O_1, O_3, O_5 + the O of the solvent	/	0.5	500

1.1. Phosphate monoanion reactivity with enforced substrate-assisted proton transfer

Table S5: Additional constraints related to the proton transfers as defined through the distances to the protons donor and acceptors.

Atomic index	Values a_i (Å)		Force constant k_i (kcal/mol/Å ²)
	Lower wall	Upper wall	
d_{mm}	/	3.5	500
d_{ma}	/	3.5	500
d_{dd}	/	3.5	500
d_{dl}	/	3.5	500
δH_A	-2	2	500
δH_B	-2	2	500

Table S6: Additional constraints to be added to restrain the reaction to go through a mechanism of type $A_N D_N$ or $D_N + A_N$. The constraints are applied on the difference of the distances for transfers A and B.

Type of mechanism	Index	Values (Å) Upper wall	Force constant k_i (kcal/mol/Å ²)
$A_N D_N$	$\delta_B - \delta_A$	0.2	500
$D_N + A_N$	$\delta_A - \delta_B$	0.2	500

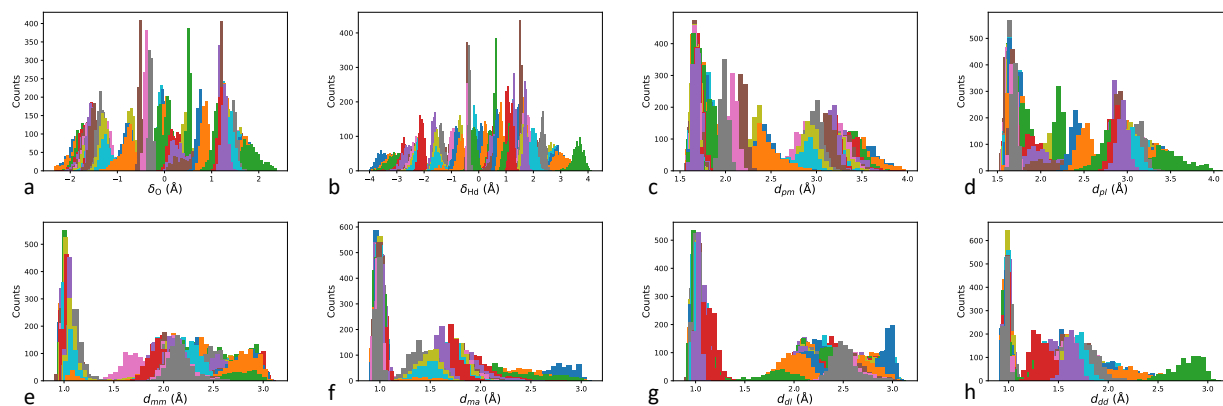


Figure S2: Overlap between umbrella sampling (US) windows, as represented along the different molecular distances as defined in the main text, for the associative mechanism with enforced substrate-assisted proton transfer. Each color corresponds to one US replica.

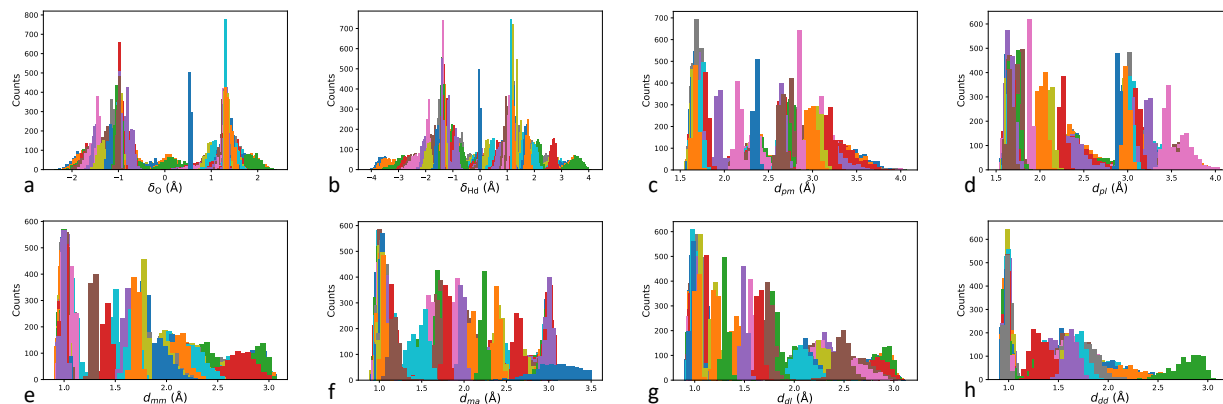


Figure S3: Overlap between umbrella sampling (US) windows, as represented along the different molecular distances as defined in the main text, for the dissociative mechanism with enforced substrate-assisted proton transfer. Each color corresponds to one US replica.

1.2. Phosphate monoanion reactivity with possibility of proton transfer with the solvent

Table S7: New constraints related to the possibility of proton transfer with the solvent. In that case, the constraints of Tab. 4, 5 and 6 are removed.

Name	Group A	Group B	Values		Force constant k_i (kcal/mol)
			Lower wall	Upper wall	
C_1	O_1	H_6, H_7, H_{10} and the H of the solvent	0.8	2.5	500
C_2	O_9	H_6, H_7, H_{10} and the H of the solvent	/	1.5	500

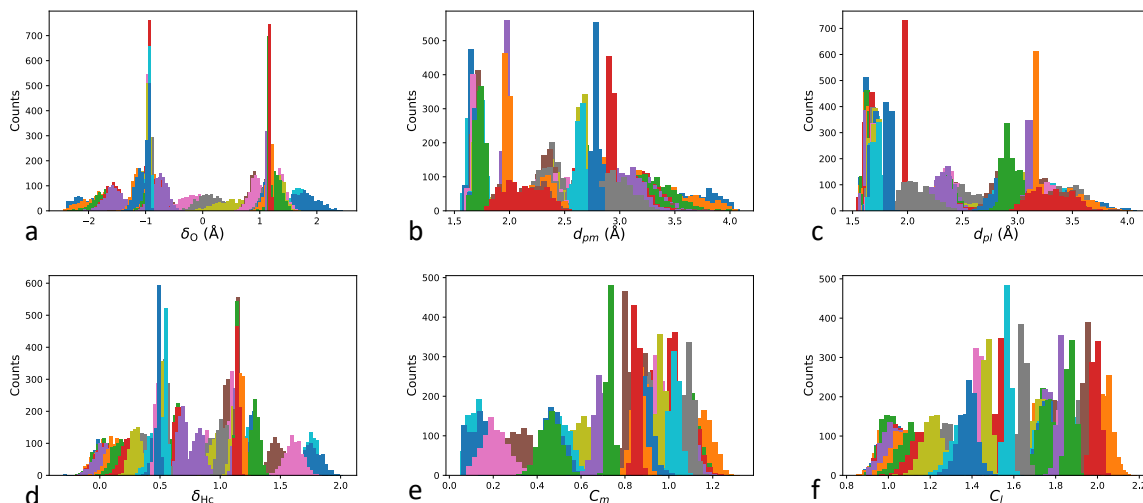


Figure S4: Overlap between umbrella sampling (US) windows, as represented along the different molecular distances as defined in the main text, for the phosphoester bond formation in the monoanion with possibility of solvent-assisted proton transfer. Each color corresponds to one US replica.

2. Phosphate dianion reactivity

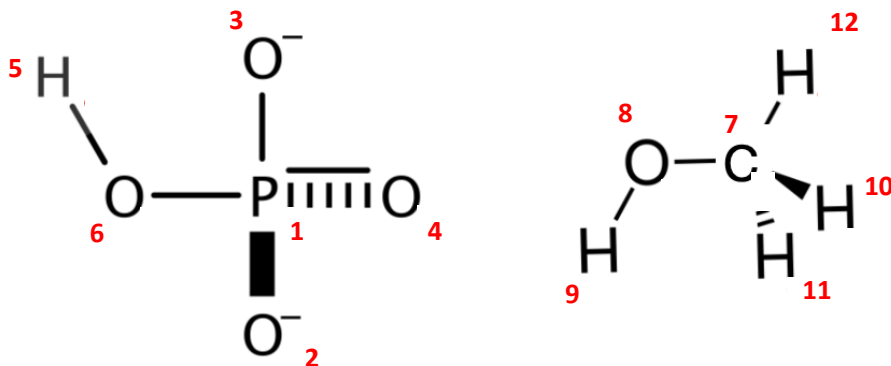


Figure S5: Identification of the atoms of phosphate dianion and methanol from 1 to 12.

Table S8: Definition of the molecular angles

Angles	α	β_1	γ_1	δ	β_2	γ_2	β_3	γ_3
Atomic index	$O_6P_1O_8$	$O_6P_1O_2$	$O_2P_1O_8$	$C_7O_8P_1$	$O_6P_1O_3$	$O_3P_1O_8$	$O_6P_1O_4$	$O_4P_1O_8$

Table S9: Constraints applied on angles or angle differences during the simulations.

Angles	Values a_i (rad)		Force constant k_i (kcal/mol/rad ²)
	Lower wall	Upper wall	
$\beta_1-\beta_2$	-0.5	0.5	400
$\beta_1-\beta_3$	-0.5	0.5	400
$\beta_3-\beta_2$	-0.5	0.5	400
$\gamma_1-\gamma_2$	-0.5	0.5	400
$\gamma_1-\gamma_3$	-0.5	0.5	400
$\gamma_3-\gamma_2$	-0.5	0.5	400
δ	1.745	/	500

Table S10: Constraints applied on distances or distance differences during the simulations.

Atomic index	Values a_i (Å)		Force constant k_i (kcal/mol/Å ²)
	Lower wall	Upper wall	
$d_{pl}: P_1O_6$	/	4	500
$d_{pm}: P_1O_8$	/	4	500
$d_{pl}-d_{pm}$	-2.4	2.4	400
O_2P_1	/	1.8	500
O_3P_1	/	1.8	500
O_4P_1	/	1.8	500
$H_{10}C_7$	/	1.25	500
$H_{11}C_7$	/	1.25	500
$H_{12}C_7$	/	1.25	500
O_8C_7	1.35	1.6	500
O_6H_5	/	1.2	500

Table S11: Constraints applied on the number of coordination during the simulations.

Group A	Group B	Values		Force constant k_i (kcal/mol)
		Lower wall	Upper wall	
O_6	$H_5, H_9,$ and the H of the solvent	0.8	2.5	500
O_8	$H_5, H_9,$ and the H of the solvent	/	1.5	500

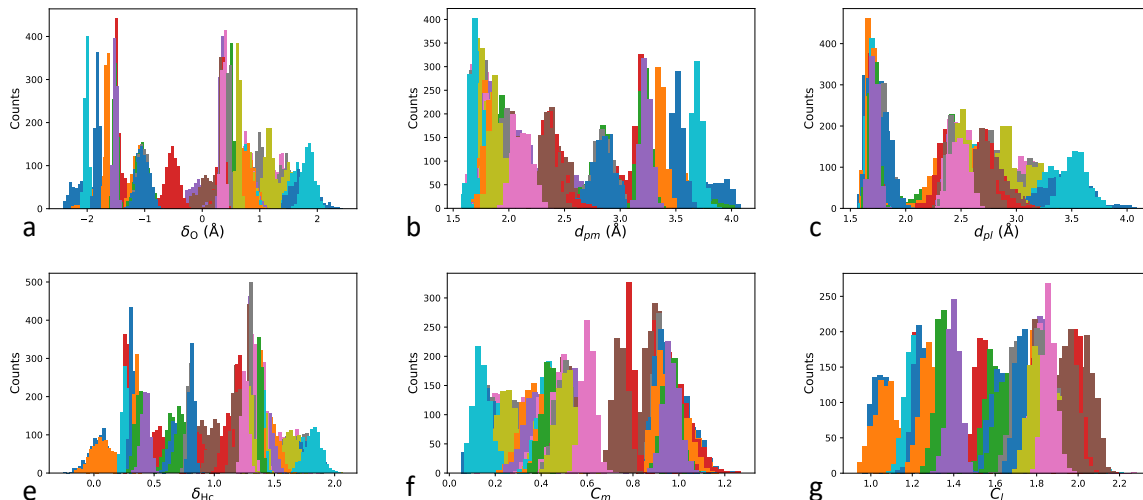


Figure S6: Overlap between umbrella sampling (US) windows, as represented along the different molecular distances as defined in the main text, for the phosphoester bond formation in the dianion with possibility of solvent-assisted proton transfer. Each color corresponds to one US replica.

3. DFT-based energetic corrections to the free-energy paths

For the PDP (phosphate diprotonated) species, we have calculated the single point energies for 10-40 structures randomly chosen around each point of the string along the free-energy path, both at the DFTB level, as well as at the BLYP DFT level, including dispersion interaction corrections (D3, similar to those tested for the description of the water bulk phase, see main text). We then compared the free-energy paths obtained from our aiMD DFTB simulations, to those for which the energetic contribution to the free-energy was corrected as follows: $G_{corr} = G_{DFTB} + E_{DFT} - E_{DFTB}$. This approach is simplistic and probably suffers from many approximations: the structures are those of DFTB, the calculations are performed in the gas phase in the absence of explicit solvent molecules (which could in particular impact the energies of species with large partial and/or net charges), and the entropic contributions are assumed to be similar at the two quantum levels. While a more systematic investigation would involve a much more elaborate and refined approach, these

calculations enable a crude and rough estimate of the robustness of our conclusions.

As expected, single point calculations on the DFTB structures show some differences between the BLYP-D3 and the SCC-DFTB methods; however, we note that the "corrected" free-energy profiles exhibit the same qualitative behaviors, and that the dissociative-associative mechanism remains much more favored than an associative-dissociative mechanism.

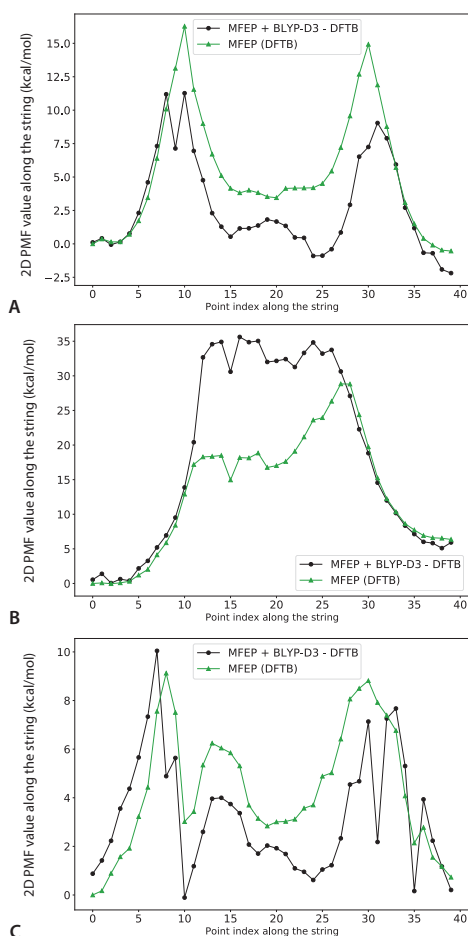


Figure S7: MFEP including (green) or not (black) corrections based on gas phase single point calculations at the BLYP-D3 level for (A) the dissociative-associative mechanism with distance-based proton transfer coordinates, (B) the associative mechanism with distance-based proton transfer coordinates, and (C) the dissociative-associative mechanism with coordination-based proton transfer coordinates.