Supporting Information

Mondal and Warshel 10.1073/pnas.1800054115

S1. Empirical Valence Bond Simulations

The empirical valence bond (EVB) is a well-known and successfully used method in the field of enzyme catalysis (36). It represents the reacting system in a simpler but in a realistic way as a superposition of different resonance forms (diabatic states). Therefore, it can give quantitative comparison of the effect of different environments on the reaction potential surfaces. The resonance forms correspond to the classical valence bond descriptions of the reactant, intermediate, and product states. The potential energy of a diabatic state can be represented as a combination of two force fields. In EVB, the general molecular mechanics (MMs) force field Enzymix (37) is used to simulated part of the system that is not involved in the reaction, whereas a quantum empirical force field is used to represents the reaction center. The potential energy function can be represented as follows:

$$\in_i = \alpha^i_{\text{gas}} + U_{\text{intra}}(R,Q) + U_{Ss}(R,Q,r,s) + U_{ss}(r,q). \qquad \textbf{[S1]}$$

R and Q represent the position and the charges (calculated quantum mechanically) of atoms in the reaction center (solute) in the ith diabatic state, whereas r and q represent the position and charges of atoms of the surrounding protein and solvent. α^i_{gas} denotes the gas phase energy in the ith diabatic state, where all of the fragments of the solute are infinitely apart. U_{intra} is the intramolecular potential of the solute. U_{Ss} represents the intermolecular interaction potential between the solute (S) and the surrounding (s) atoms. U_{ss} represents interaction between surrounding atoms [surrounding–surrounding (ss)].

The potential energy function for the diabatic states can be used to build the EVB Hamiltonian H_{EVB} (Eq. **S2**) where the off-diagonal (coupling) term H_{ij} can be approximated as a function of the reacting bonds and assumed to be same in the gas phase, solution, and in the protein:

$$H_{\text{EVB}} = \begin{vmatrix} \in_{ii} - E^g & H_{ij} \\ H_{ij} & \in_{jj} - E^g \end{vmatrix}, \qquad \textbf{[S2]}$$

$$H_{\text{EVB}} C_g = E_g C_g. \qquad \textbf{[S3]}$$

The adiabatic ground-state energy E_g, and the corresponding eigen vector C_g can be obtained by solving the secular equation (Eq. **S3**).

To calculate the complete reaction profile, we run molecular dynamics (MD) simulation on a mapping potential (ϵ_m) (Eq. **S4**), which is a linear function of the diabatic potentials of the initial (state 1) and the final state (state 2) (for example, in the case of a single step reaction, the reactant and product are the initial and the final state, respectively):

$$\in_m = \lambda_m \epsilon_1 + (1-\lambda_m)\epsilon_2, \quad \text{where} \quad 0 \le \lambda_m \le 1. \qquad \textbf{[S4]}$$

The parameter λ_m varies between 0 and 1 in $N+1$ windows as the initial state changes to the final state.

The free energy associated with the change in potential from ϵ_1 to ϵ_2 can be calculated using free energy perturbation/umbrella sampling (FEP/US) method. The free-energy functional $\Delta G(\lambda)$ can be represented as follows:

$$\Delta G(\lambda_n) = \Delta G(\lambda_0 \to \lambda_n) = \sum_{i=0}^{N-1} \delta G(\lambda_i \to \lambda_{i+1}), \qquad \textbf{[S5]}$$

where

$$\delta G(\lambda_i \to \lambda_{i+1}) = -\left(\frac{1}{\beta}\right) \ln\left[\left\langle e^{(-(\epsilon_{i+1}-\epsilon_i))\beta} \right\rangle_i\right]. \qquad \textbf{[S6]}$$

In Eq. **S6**, $\beta = 1/k_B T$, where k_B denotes the Boltzmann constant and T is the temperature of the simulations. The angular bracket ($< >_i$) indicates an average of the quantity within the bracket, calculated on the potential surface ϵ_i.

The activation free energy ($\Delta g^{\neq}$), which requires the calculation of the probability of reaching the transition state for the trajectory that moves on E_g, can be calculated using the following FEP/US equation:

$$\exp[-\Delta g(X^n)\beta] = \exp[-\Delta G(\lambda_m)\beta] \left\langle \exp\left[-\left(E_g(X^n) - \epsilon_m(X^n)\right)\beta\right] \right\rangle_m, \qquad \textbf{[S7]}$$

where X^n denotes the reaction coordinate in terms of energy gap $\epsilon_2 - \epsilon_1$.

In our current EVB simulations, the following protein crystal structures [EF-G′, PDB ID code 4V90 (9); EF-G, PDB ID code 2BV3 (6); EF-Tu′, PDB ID codes 2XQD and 2XQE (7); and EF-Tu, PDB ID code 1EFT (4)] were used. All water molecules [except the nucleophilic water molecule, and the water close to the γ-phosphate (for EF-G′)] and all ions [except the Mg^{2+} coordinated to GTP and another Mg^{2+}, which is shown to be coordinated with Asp22 (for EF-G′); see figure 5 of ref. 9] were discarded from the PDB files before starting any simulation. All water molecules were added using the software Molaris-XG (38). For the active GTPases, the whole ribosome was not taken in the simulation. As a simulation system, we took all of the residues of EF-Tu and EF-G in its ribosome-bound form and only the critical nucleotides from 23S rRNA (A2654 to G2664 for EF-G and C2652 to A2666 for EF-Tu). Although eight more nucleotides [compared with our previous study on EF-Tu (15)] were taken in this current study, the overall conclusion remains the same as in our previous study. Thus, we assumed that the addition of any extra nucleotide in the simulation system would be redundant. To study the mutational effects, the PDB of the wild-type proteins (after the modification mentioned above) were mutated using Chimera 1.10.2 (32). The rotamer [from Dunbrack Rotamer library (39)] conformation with highest probability was selected for the mutation.

For all simulations, we used a spherical system ($r = 20$ Å) where the center of the catalytic site (EVB region) was taken as the center of the simulation system. The spherical boundary waters were subject to polarization and radial restraint according to the surface constraint all-atom solvent (SCAAS) model (40). These surface constraints are introduced as boundary conditions in the SCAAS model to make the finite system behave as if it is part of an infinite system. The local reaction field (LRF) approach (41) was used to treat the long-range effects.

The initial system was heated slowly starting from 10 to 300 K for 200 ps with 1-fs time step (fixed EVB region). The final structure obtained from the heating was used to relax further by decreasing the constraint on the EVB region for another 200 ps. The structure obtained from the last step was used to relax the system further for another 100 ps (with 0.3 kcal/mol position constraint on the EVB region and 0.03 kcal/mol on the protein

atoms outside the EVB region). We finally generated three structures from the final relaxation step to use in the EVB calculations. The same EVB parameters were used in the relaxation as well as in the EVB calculations. The force field parameters for the rest of the protein and water were taken from the polarizable force field ENZYMIX (38). The free-energy profiles were calculated by the FEP/US approach, where the total FEP simulation was divided into 31 mapping frames, with 20 ps each. Three independent runs were performed to take the average for the EVB calculations. All parameters used in EVB calculations are given below.

S2. Water Flooding Simulations

The water flooding (WF) approach (25) can be used to generate water configurations that are likely to be present at the site of interest. In this method, the energy of insertion of a water molecule is calculated using the linear response approximation (LRA) and the linear interaction energy (LIE) approach, and finally a postprocessing Monte Carlo (MC) is run to sort the energies.

The MC criterion applied to accept addition and deletion of a water can be represented by Eqs. **S8** and **S9**, respectively:

$$P_{m+1} = \min\left\{1, \left(\frac{1}{N+1}\right)\exp\left[-\left((\Delta G_{m+1} - \Delta G_m)\beta - B'_{m+1}\right)\right]\right\}, \quad \textbf{[S8]}$$

$$P_m = \min\left[1, N\exp\left((\Delta G_{m+1} - \Delta G_m)\beta - B'_m\right)\right], \quad \textbf{[S9]}$$

where the LHS of the equations represent the probability of accepting a move; ΔG_i is the approximate free energy of configuration i of the internal water molecules; and parameter B' in Eqs. **S8** and **S9** can be represented as $B'_i = \Delta G_{\text{bulk}}\beta + \ln N_i$. The term ΔG_{bulk} in the expression of B' defines the chemical potential of the system, and N_i denotes the average number of waters in the system. More detailed description about WF can be found in ref. 25.

The WF simulations were performed as follows: At the beginning, all water molecules (except the nucleophilic water) and all ions (except the above mentioned Mg^{2+}s) were discarded from the PDB files, then we minimized the potential energy of the system for 3,000 steps using steepest descent method. The structure obtained from the last energy minimization step was used as a starting structure for the WF calculation. In the WF calculation, the protein was immersed in a solvent sphere of radius 18 Å using the SCAAS, and LRF model was used to treat the long-range electrostatics. A 6-Å radius from the γ-phosphate of GTP was defined to introduce excess number of water by reducing the van der Waals cutoff distance (r_{cut}) from 2.8 to 2.1 Å. After 6 Å, the default r_{cut} was used to generate the rest of the water molecules. A spherical hard wall was introduced (r = 6 Å from the γ-phosphate of GTP) to prevent the waters from crossing the hard wall. Thus, the inside waters could not escape from the hard wall sphere and the outside waters could not enter inside the hard wall sphere. After generating water configuration using 10-ps MD simulation, a postprocessing MC (B' = −12 kcal/mol) simulation was performed to get minimum energy water configuration around the active site.

S3. Electrostatic Group Contribution Simulations

Electrostatic group contribution calculations have been extensively used in our group to qualitatively explain the effect of various residues in different processes like drug resistance, and enzyme catalysis. Study of individual mutational effect on free-energy changes could also be an option, if we want to estimate this contribution in a more quantitative way. However, the PDLD/S-LRA method could be a reasonable approach for qualitative estimations. The PDLD/S-LRA approach is used to calculate the solvation free energy for RS and TS state. The change in free energy upon the moving the solute (i) from water (w) to protein (p) can be represented as [based on the semimicroscopic version of the protein dipole Langevin dipole method (PDLD/S)]:

$$\Delta U^{w\to p}_{sol,i} = \left[-\Delta G^{w}_{sol,i}(q=q_i) + \Delta G^{p}_{sol,i}(q=q_i) - \Delta G^{p}_{sol,i}(q=0)\right]\left(\frac{1}{\epsilon_p} - \frac{1}{\epsilon_w}\right) + \Delta U_{q\mu}\frac{1}{\epsilon_p}. \quad \textbf{[S10]}$$

In the original PDLD/S approach, only one configuration is considered, so the PDLD/S free energy is defined as an effective potential U. $\Delta G^{s}_{sol,i}(q=q_i)$ represents contribution of the solvent (s) in the solvation free energy when i is in charge state q_i. ϵ_w represents the dielectric constant of water, whereas ϵ_p represents the scaling factor that is introduced to take care of the interactions that are not treated explicitly. The term $\Delta U_{q\mu}$ represents the interaction between the charge q and the surrounding protein dipoles μ.

To account for the effect of protein reorganization associated with the moving of charge from water to protein, we used the LRA approximation (Eq. **S11**):

$$\Delta G^{w\to p}_{sol,i} = \frac{1}{2}\left[\left\langle \Delta U^{w\to p}_{sol,i}\right\rangle_{q=q_i} + \left\langle \Delta U^{w\to p}_{sol,i}\right\rangle_{q=0}\right], \quad \textbf{[S11]}$$

where $\langle\rangle_q$ represents the molecular dynamic average obtained by using a potential that corresponds to the charge state q.

To calculate the individual contribution of the residue j on the solvation of i, the following equation can be used:

$$\Delta\Delta G^{ij}_{sol} = \Delta G\left(q_i = 0 \to q_i = \overline{q_i}\right)^{w\to p}_{q_j=\overline{q_j}} - \Delta G\left(q_i = 0 \to q_i = \overline{q_i}\right)^{w\to p}_{q_j=0}, \quad \textbf{[S12]}$$

where $\Delta G(q_i{=}0 \to q_i{=}\overline{q_i})^{w\to p}_{q_j=\overline{q_j}}$ denotes the change in the free energy of moving i with charge $\overline{q_i}$ from water to protein when residue j has total charge $\overline{q_j}$. Eq. **S12** can be approximated, where $\Delta U^{j}_{q\mu}1/\epsilon_p$ can be considered as the major contributing term in $\Delta\Delta G^{ij}_{sol}$. This approximation is valid for polar residues only, since the free energy of charging i from 0 to $\overline{q_i}$ is small for polar residues. In the case of ionic residues, the contribution from the other free-energy terms are not negligible compared with $\Delta U^{j}_{q\mu}$, and the significant compensation of $\Delta U^{j}_{q\mu}$ can be considered by using higher scaling factor ϵ_p for ionic residues. For further clarification, see ref. 31.

The PDLD/S-LRA calculations used the EVB structures of the RS and the TS for the active as well as inactive GTPase. The obtained structures were used to calculate the generalized solvation free energy (the electrostatic interaction with the surrounding) for the nucleophilic water and the phosphate groups separately in the reactant and transition state of the active and inactive form, using the PDLD/S-LRA (31). The LRA group contributions, obtained from the solvation free-energy calculation, were used to calculate the relative contribution of different protein residues on going from reactant state to the transition state both for the active and the inactive form. The relative contributions obtained for the active form was then subtracted from that of the inactive form for each residue to get the allosteric activation contribution of different residues.

S4. Estimating the Conformational and Chemical Landscape

To analyze the relation between conformational free-energy change with catalytic free-energy change, we need intermediate structures between the active and inactive form of the enzyme. Targeted molecular dynamics (TMD) simulation was used to generate those intermediate structures between active EF-G (EF-G′) and EF-G, which were subsequently used in FEP/EVB calculations to get the relevant activation free energies. (See Fig. 6 in main text.)

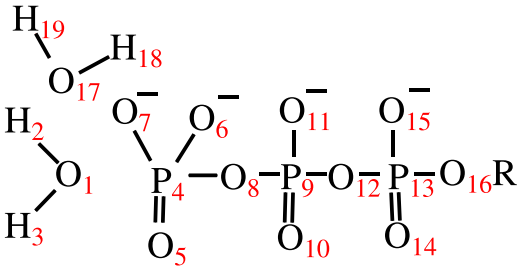

Fig. S1. EVB atoms used to define different resonance states.

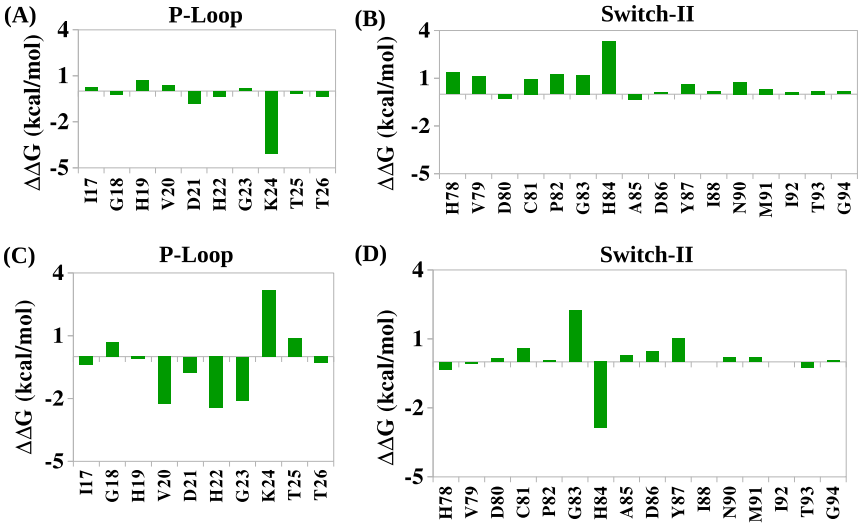

Fig. S2. Allosteric contributions (the change in the calculated group contributions upon moving from EF-Tu to EF-Tu′ for the transition from the RS to the TS) of different residues of the P-loop and switch II of EF-Tu in the 1W mechanism. *A* and *B*, and *C* and *D* represent the allosteric contribution for changing the charges of the phosphate groups of GTP and the nucleophilic water, respectively [effective dielectric constant ε_{eff} for ionizable residues (except Lys) =10; Lys = 20].

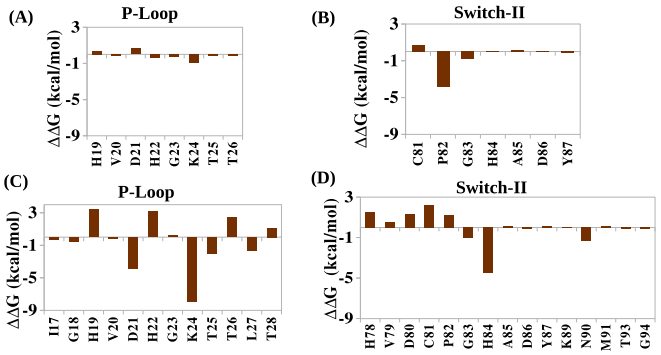

Fig. S3. Allosteric contributions (the change in the calculated group contributions upon moving from EF-Tu to EF-Tu′ for the transition from the RS to the TS) of different residues of the P-loop and switch II of EF-Tu in the 2W mechanism, where the PGH loop was constrained. *A* and *B*, and *C* and *D* represent the allosteric contribution for changing the charges of the phosphate groups of GTP and the nucleophilic water, respectively [effective dielectric constant ε_{eff} for ionizable residues (except Lys) = 10; Lys = 20].

Table S1. Partial charges used in the 2W mechanism

	Reactant		Intermediate	
Atom number (Fig. S1)	Atom type	Charge	Atom type	Charge
1	O_0	−0.899	O_q	−0.760
2	H_0	0.425	H_0	0.468
3	H_0	0.444	H_0	0.456
4	P_0	1.520	P_+	1.498
5	O^-	−1.010	O^-	−0.933
6	O^-	−0.996	O^-	−0.827
7	O^-	−1.058	O^-	−0.927
8	O_p	−0.712	O_D	−0.924
9	P_0	1.600	P_0	1.458
10	O^-	−1.012	O^-	−1.053
11	O^-	−0.960	O^-	−1.018
12	O_p	−0.530	O_p	−0.550
13	P_0	1.250	P_0	1.250
14	O^-	−0.846	O^-	−0.870
15	O^-	−0.846	O^-	−0.870
16	O_p	−0.470	O_p	−0.470
17	O_0	−0.800	O_0	−0.800
18	H_0	0.450	H_0	0.450
19	H_0	0.450	H_0	0.450

The charges were calculated using B3LYP functional with 6-31+G(d,p) basis set.

Table S2. Partial charges used in the 1W mechanism

	Reactant		Intermediate		Product	
Atom number	Atom type	Charge	Atom type	Charge	Atom type	Charge
1	O_0	−0.899	O_N	−1.200	O_M	−0.765
2	H_0	0.425	H_0	0.450	H_0	0.464
3	H_0	0.444	H_0	0.300	H_0	0.464
4	P_0	1.520	P_0	1.430	P_0	1.378
5	O^-	−1.015	O^-	−0.914	O^-	−0.891
6	O^-	−0.971	O^-	−0.866	O^-	−0.885
7	O^-	−0.985	O_H	−0.755	O_H	−0.835
8	O_p	−0.705	O_p	−0.650	O_D	−0.923
9	P_0	1.600	P_0	1.600	P_0	1.527
10	O^-	−1.012	O^-	−1.012	O^-	−1.056
11	O^-	−0.960	O^-	−0.960	O^-	−1.016
12	O_p	−0.530	O_p	−0.530	O_p	−0.550
13	P_0	1.250	P_0	1.250	P_0	1.250
14	O^-	−0.846	O^-	−0.846	O^-	−0.846
15	O^-	−0.846	O^-	−0.846	O^-	−0.846
16	O_p	−0.470	O_p	−0.470	O_p	−0.470

The charges were calculated using B3LYP functional with 6-31+G(d,p) basis set.

Table S3. EVB parameters: Part 1

Bonding interaction: $V_b = D_0[1 - e^{-\mu(b-b_0)^2}] + k_b(b-b_0)^2$

Bond type	D_0, kcal·mol^{-1}	b_0, Å	μ, Å^{-1}	k_b, kcal·mol^{-1}·Å^{-2}
P_0 – O^-	120.0	1.50	2.0	0.0
P_+ – O^-	100.0	1.52	2.0	0.0
P_0 – O_p	85.0	1.70	2.0	0.0
P_+ – O_D	0.0	2.70	0.0	10.0
P_0 – O_D	70.1	1.60	2.0	0.0
P_0 – O_0	0.0	3.10	2.0	10.0
P_+ – O_q	75.0	2.20	2.0	0.0
P_0 – O_N	0.0	3.10	2.0	10.0
P_0 – O_M	85.0	1.70	2.0	10.0
O_H – H_0	102.0	0.96	1.0	0.0
O_N – H_0	102.0	0.96	1.0	0.0
O_0 – H_0	102.0	0.96	1.0	0.0
O_M – H_0	102.0	0.96	1.0	0.0
O_q – H_0	50.0	1.10	1.0	0.0

Table S4. EVB parameters: Part 2

Angle interaction: $V_\theta = D_\theta[1 - e^{-(\theta-\theta_0)^2/\sigma}] + k_\theta(\theta-\theta_0)^2$

Angle type	k_θ, kcal·mol^{-1}·rad^{-2}	θ, degree	σ, rad^2	D_θ, kcal·mol^{-1}
O^- – P_0 – O^-	140.0	119.0	1.0	0.0
O^- – P_+ – O^-	100.0	120.0	1.0	0.0
O^- – P_0 – O_p	70.0	109.0	1.0	0.0
O^- – P_0 – O_0	10.0	80.0	1.0	0.0
O^- – P_+ – O_D	10.0	90.0	1.0	0.0
H_0 – O_0 – P_0	50.0	120.0	1.0	0.0
O^- – P_+ – O_q	100.0	90.0	1.0	0.0
O^- – P_0 – O_D	100.0	109.0	1.0	0.0
O_p – P_0 – O_p	70.0	100.0	1.0	1.0
P_0 – O_p – P_0	100.0	120.0	1.0	0.0
O_M – P_0 – O_H	70.0	109.0	1.0	1.0
O_M – P_0 – O^-	70.0	109.0	1.0	1.0
P_0 – O_M – H_0	80.0	106.0	1.0	0.0
H_0 – O_0 – H_0	80.0	105.0	1.0	0.0
H_0 – O_q – H_0	80.0	105.0	1.0	0.0
P_+ – O_q – H_0	80.0	120.0	1.0	0.0
P_0 – O_H – H_0	80.0	106.0	1.0	0.0

Table S5. EVB parameters: Part 3

Torsional angle interaction: $V_\varphi = k_\varphi[1 + \cos(n\varphi - \varphi_0)]$

Angle type	k_φ, kcal·mol^{-1}·rad^{-2}	n	φ, degree
O^- – P_0 – O_p – P_0	2.0	3.0	0.0
O_p – P_0 – O_H – H_0	2.0	3.0	0.0
O^- – P_+ – O_D – P_0	2.0	3.0	0.0
P_0 – P_0 – O_H – H_0	2.0	3.0	0.0
O^- – P_+ – O_q – H_0	2.0	3.0	0.0
O^- – P_0 – O_0 – H_0	2.0	3.0	0.0
O^- – P_0 – O_M – H_0	2.0	3.0	0.0
O_H – P_0 – O_0 – H_0	2.0	3.0	0.0

Table S6. EVB parameters: Part 4

Improper torsional angle interaction: $V_{\varphi} = k_{\varphi}[1 + \cos(n\varphi - \varphi_0)]$

Angle type	k_{φ}	n	φ, degree
$O_q - H_0 - H_0 - P_0$	30.0	2.0	180.0

Table S7. EVB parameters: Part 5

Nonbonded interaction: $V_{nb} = Ce^{-\alpha r}$(atoms bonded in one of the EVB states)

Atom type	C, kcal·mol^{-1}	α
$P_0 - O^-$	17,000	3.0
$P_0 - O_P$	17,000	3.0
$P_+ - O^-$	17,000	3.0
$P_+ - O_D$	17,000	3.0
$P_+ - O_q$	17,000	3.0
$P_0 - O_M$	17,000	3.0
$P_0 - O_H$	17,000	3.0
$H_0 - O_0$	4,000	4.0
$H_0 - O_q$	4,000	4.0
$H_0 - O_H$	4,000	4.0
$H_0 - O_N$	4,000	4.0
$H_0 - O_M$	4,000	4.0

Table S8. EVB parameters: Part 6

Lennard Jones potential: $V_{LJ}^{ij} = A_i A_j / r_{ij}^{12} - B_i B_j / r_{ij}^{6}$ (atoms never bonded)

Atom type	A, kcal·mol^{-1}·Å^{12}	B, kcal·mol^{-1}·Å^{12}
P_0	2,454.0	47.0
P_+	2,454.0	47.0
O_0	636.0	23.0
O_P	636.0	23.0
O_D	1,379.0	35.0
O_q	636.0	23.0
O_H	636.0	23.0
O^-	1,379.0	35.0
O_M	1,379.0	23.0
O_N	1,379.0	35.0
H_0	5.0	0.0

Table S9. EVB parameters: Part 7

	Off-diagonal elements: $H_{ij} = Ae^{-\mu r}$		Gas phase shift: α	
Reactions	A	μ	α_1	α_2
2W mechanism: RS→INT	90.1	0.0	0.0	−30.0
1W mechanism: PT step	9.1	0.6	0.0	−27.6
1W mechanism: NA step	142.0	0.0	0.0	−95.0