

Supporting Information - Approach to Calculate the Free Energy Changes of Surface Reactions Using a Free Energy Decomposition upon *Ab Initio* Brute- force Molecular Dynamics Trajectories

Jiayan Xu, Hao Huang, P. Hu*

*School of Chemistry and Chemical Engineering, Queen's University Belfast, Belfast BT9 5AG,
UK*

Corresponding Author

*E-mail address: p.hu@qub.ac.uk

S1. Statistical Mechanics Approximation for Gibbs Free Energy

To obtain Gibbs free energy G at the given temperature T , statistical mechanics approximations merely considering the vibrational contribution are directly applied on structures optimized by density functional theory (DFT) calculations at 0 K, which is in the form of¹

$$G(T) = E_{pot} + E_{ZPE} + \int_0^T C_{V,vib} dT - TS_{vib} \quad (1)$$

where E_{pot} is the potential energy, E_{ZPE} is the zero point energy (ZPE), and $C_{V,vib}$ and S_{vib} are the vibrational constant-volume heat capacity and the vibrational entropy respectively. E_{pot} is the ground state energy from the DFT. E_{ZPE} can be calculated from the DFT based frequency calculation. The framework of the AIMD used in this study did not take the nuclear quantum effect into consideration. Thus, E_{ZPE} was excluded from the thermodynamic correction.

The potential energies and thermodynamically corrected free energies are listed in **Table S1**.

Table S1. Potential Energy and Free Energy along the NEB Pathway

Image	Potential Energy (eV)	Free Energy at 300 K (eV)	Notes
1	0.0	0.0	Initial State
2	0.0128	0.0129	
3	0.0137	0.0037	
4	0.4586	0.4686	
5	0.9161	0.9461	Transition State
6	0.3885	0.4086	
7	0.1864	0.3764	
8	-0.0210	-0.0110	
9	-0.0530	-0.0530	Final State

S2. Enhanced Sampling Methods

By the statistical mechanics, the free energy is defined as the multi-integral below

$$A = -kT \ln \int e^{-E(\mathbf{q}^N)/kT} d\mathbf{q}^N \quad (2)$$

where k is the Boltzmann constant and $E(\mathbf{q}^N)$ is the potential energy function of N atoms' general coordinates. There are plentiful enhanced sampling methods to calculate the potential of mean force (PMF). The constrained molecular dynamics (CMD)² and the umbrella sampling (US)³ are two most frequently used methods both in surface and biochemical reactions.

For the CMD, one will evaluate the free energy gradient (FEG) of the selected collective variable (CV) ξ at given values. The most used mathematical formula is⁴

$$\frac{\partial A}{\partial \xi} = \frac{\langle Z_\xi^{-1/2} (\lambda_\xi + \frac{k_B T}{2Z_\xi} \nabla_\xi \cdot \nabla Z_\xi) \rangle}{\langle Z_\xi^{-1/2} \rangle} \quad (3)$$

where Z_ξ is the metric tensor, λ_ξ is the Lagrange multiplier and the rest in the numerator is the thermo correction. The above formula can be easily extended to the system with few constraints. Generally, the SHAKE algorithm is used to control the selected CV during the simulation. In our CMD-related simulations, we took 0.5 ps to equilibrate the system and about 3.0 ps to average the FEG later. Most simulations produced the FEG with a standard error less than 0.04 eV/Å (see data in **S3** and **S7**), which is estimated by the block average algorithm.

When it comes to the US, a restraining potential will be added to the selected CV. Though the harmonic potential is used the most, only can the gaussian function be used in the implementation of VASP. The gaussian potential G at given value of CV ξ_0 is in the form of

$$G = h \exp\left(-\frac{(\xi - \xi_0)^2}{2w^2}\right) \quad (4)$$

where h is the height and w is the width. To construct the PMF along the C-O bond distance R_{C-O} , we used a -5 eV height and 0.5 Å width gaussian function which was added to R_{C-O} from 1.2 Å to 3.3 Å with an interval of 0.1 Å. The collected samples in this series of simulations generated a distribution of R_{C-O} , which lasted about 3.0 ps after a 0.5 ps relaxation. Therefore,

weighted histogram analysis method (WHAM)⁵ converted the probability into the free energy by an iterative scheme. The US along the path-CV is discussed separately in S5.

S3. Error Analysis of CMD and US along C-O Bond Distance

The PMF along R_{C-o} of CO oxidation on Pt (111) surface at 300 K was calculated by CMD and US listed in Table S2 and Table S3 respectively.

Table S2. FEG and Standard Error by CMD

R_{C-o} (Å)	FEG (eV/Å)	Standard Error (eV/Å)	Notes
3.300	-0.0323	0.0180	Initial State
3.200	-0.0825	0.0078	
3.100	-0.1809	0.0116	
3.000	-0.2806	0.0113	
2.900	-0.3927	0.0133	
2.800	-0.5248	0.0159	
2.700	-0.6649	0.0078	
2.600	-0.8213	0.0255	
2.500	-1.0447	0.0102	
2.400	-1.2366	0.0271	
2.300	-1.4446	0.0424	
2.200	-1.6375	0.0228	
2.100	-0.3318	0.0261	
2.001	0.0517	0.0134	Transition State
1.900	0.2065	0.0129	
1.800	0.5626	0.0104	
1.700	1.2118	0.0240	
1.600	1.4155	0.0217	

R_{C-o} (Å)	FEG (eV/Å)	Standard Error (eV/Å)	Notes
1.500	3.0181	0.0234	
1.394	2.3951	0.0252	
1.304	0.7338	0.0396	
1.280	0.0146	0.0493	Final State
1.218	2.9766	0.1737	

Table S3. Free Energy and Standard Error by US

R_{C-o} (Å)	Free Energy (eV)	Standard Error (eV)	Notes
3.300	0.0000	0.0000	Initial State
3.200	0.0021	0.0004	
3.100	0.0404	0.0010	
3.000	0.0633	0.0018	
2.900	0.0778	0.0026	
2.800	0.1331	0.0028	
2.700	0.1935	0.0033	
2.600	0.2579	0.0037	
2.500	0.3226	0.0039	
2.400	0.4179	0.0041	
2.300	0.5243	0.0042	
2.200	0.6584	0.0042	
2.100	0.7996	0.0044	
2.000	0.8332	0.0046	Transition State
1.900	0.8296	0.0048	
1.800	0.7684	0.0048	
1.700	0.6947	0.0050	

R_{C-O} (Å)	Free Energy (eV)	Standard Error (eV)	Notes
1.600	0.4982	0.0054	
1.500	0.2307	0.0055	
1.4	-0.0456	0.0054	
1.3	-0.2953	0.0054	Gas State
1.2	-0.4175	0.0054	

S4. On-the-fly Free Energy Gradient Analysis of C-Pt Bond Distance

Selecting reference bonds for the path-CV can be a tough task. To realize a fast screen of bonds responsible for the free energy change, we took the on-the-fly free energy gradient analysis. For few brute-force trajectories beginning from the TS, we could calculate the instant FEG of a selected CV by the vector form of FEG proposed by Darve et al.⁶ Since a plausible estimation of the FEG must be established on its average, we split the CV value into few small bins and averaged the instant FEGs. The analysis on $R_{C=O}$ is shown in **Figure S1**. The bond distance changes around its equilibrated distance. Thus, we concluded that there is no net free energy contribution from C-O triple bond during the reaction.

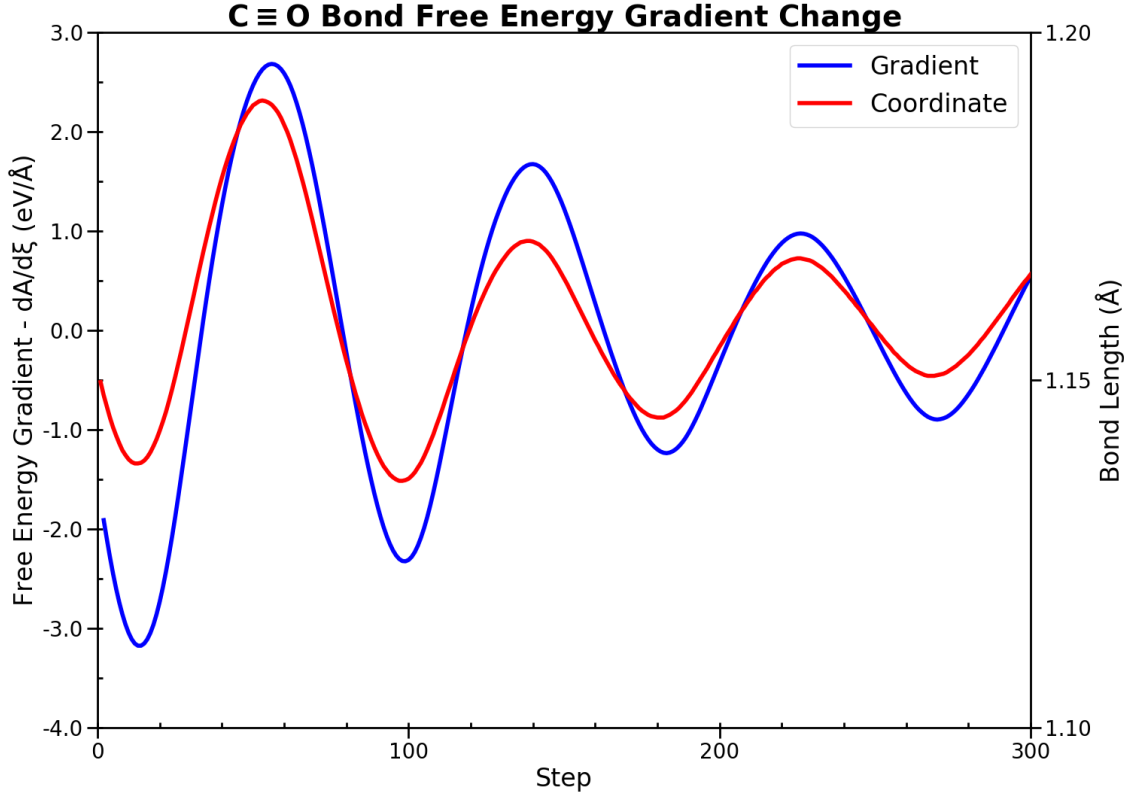


Figure S1. $R_{C\equiv O}$ and its FEG during a brute-force trajectory. The bond distance and its FEG fluctuate at a certain period.

S5. Path Collective Variable: Implementation and Sampling

The path-CV⁷ taking bonds as the reference can be calculated as⁸

$$s(\Xi) = \frac{1}{P-1} \frac{\sum_{i=1}^P (i-1) e^{-\lambda(\Xi-\Xi_i)^2}}{\sum_{i=1}^P e^{-\lambda(\Xi-\Xi_i)^2}} \quad (5)$$

$$z(\Xi) = -\frac{1}{\lambda} \ln\left(\sum_{i=1}^P e^{-\lambda(\Xi-\Xi_i)^2}\right) \quad (6)$$

where $\Xi = \{\xi_1, \dots, \xi_N\}$ is path vector whose entries are bond distances, P is the number of reference structures and λ is the coefficient which smooths the path. In our simulations, we took five bond distances, R_{C-O} , R_{C-pt_D} , R_{O-pt_A} , R_{O-pt_B} , and R_{O-pt_C} , to form the path-CV. The value of λ is recommended as the reciprocal of the distance between two neighbor reference structures. Since we took 30 reference structures, λ was set to 100 \AA^{-2} .

To calculate the PMF along the path-CV, we utilized the US. Here, we took the harmonic potential as the restraint. The potential can be written as

$$H = \frac{1}{2}k(\xi - \xi_0)^2 \quad (7)$$

where k is the spring constant and ξ can be the path progress s or the path deviation z . We adopted the spring constant 500 eV for s and 80 eV for z . To add this restraint to the system, we have to calculate the Jacobian matrix which transforms the cartesian coordinates to the path-CV. The mathematical formulas are

$$\frac{\partial s}{\partial \mathbf{R}} = \frac{1}{P-1} \frac{\frac{\partial g}{\partial \mathbf{R}} f - g \frac{\partial f}{\partial \mathbf{R}}}{f^2}, \quad \frac{\partial z}{\partial \mathbf{R}} = -\frac{1}{\lambda f} \frac{\partial f}{\partial \mathbf{R}} \quad (8)$$

$$f(\Delta \Xi_i) = \sum_{i=1}^P e^{-\lambda(\Delta \Xi_i)^2}; \quad g(\Delta \Xi_i) = \sum_{i=1}^P (i-1) e^{-\lambda(\Delta \Xi_i)^2} \quad (9)$$

$$\frac{\partial f}{\partial \mathbf{R}} = \frac{\partial \sum_{i=1}^P e^{-\lambda(\Xi - \Xi_i)^2}}{\partial \mathbf{R}} = \sum_{i=1}^P -2\lambda e^{-\lambda(\Xi - \Xi_i)^2} (\Xi - \Xi_i) \frac{\partial \Xi^T}{\partial \mathbf{R}} \quad (10)$$

$$\frac{\partial g}{\partial \mathbf{R}} = \frac{\partial \sum_{i=1}^P (i-1) e^{-\lambda(\Xi - \Xi_i)^2}}{\partial \mathbf{R}} = \sum_{i=1}^P -2\lambda (i-1) e^{-\lambda(\Xi - \Xi_i)^2} (\Xi - \Xi_i) \frac{\partial \Xi^T}{\partial \mathbf{R}} \quad (11)$$

where $\mathbf{R} = \{x_1, \dots, x_{3N}\}$ is the cartesian coordinate and $\Delta \Xi_i = \Xi - \Xi_i$ is the difference between the current structure and the i -th reference structure. This series of simulations for path-CV was carried out from $s = 0.0$ to $s = 1.0$ at an interval of 0.02 while z was controlled around 0.05. The system at each sample point first equilibrated for 0.5 ps and then produced for 3.0 ps. The final 2-D PMF spanned by s and z was constructed by the 2-D WHAM.

S6. Selected Brute-force Molecular Dynamics Trajectory

The free energy decomposition analysis was performed on the selected brute-force MD trajectory starting from the TS. Those trajectories were carried out at a timestep of 0.2 fs. Since the integration timestep is really small, most of the output trajectories were very similar. Thus, we selected two trajectories, one from TS to FS and the other from TS to IS, as shown in **Figure S2**.

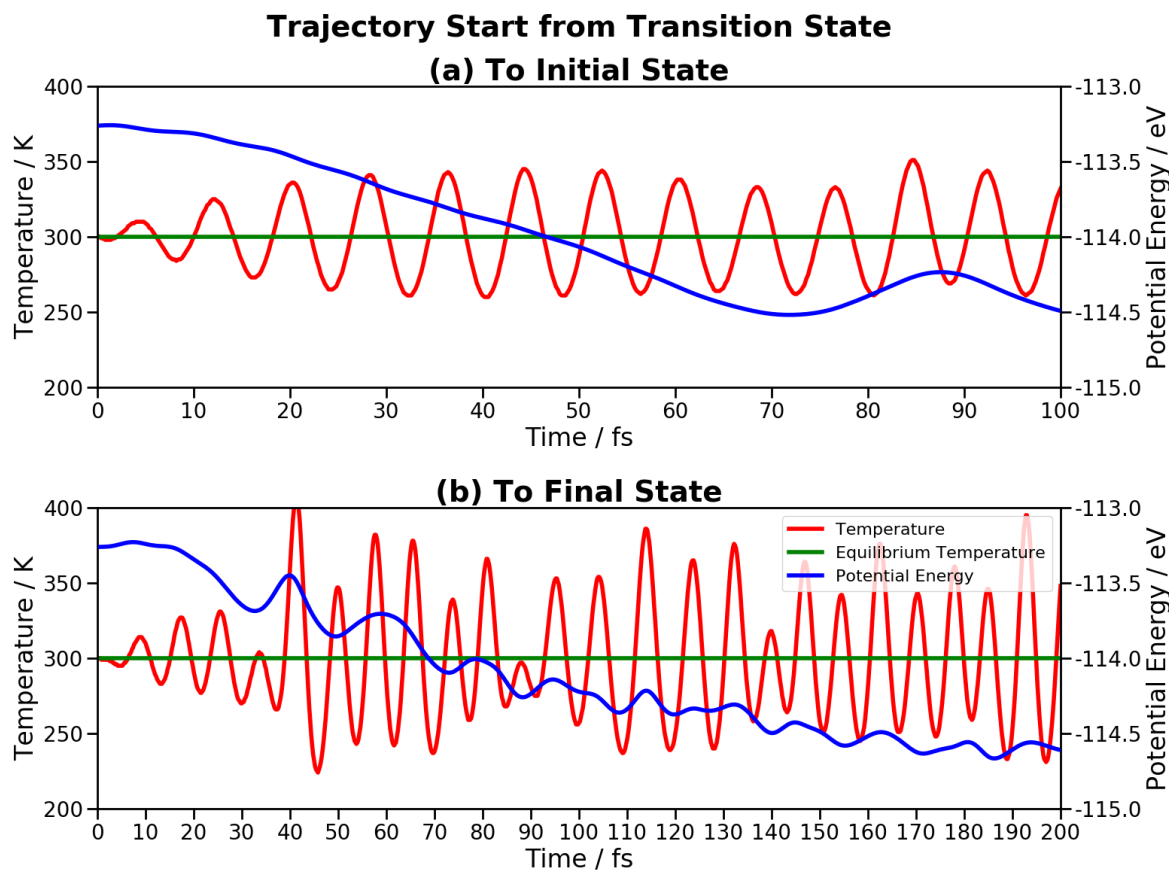


Figure S2. Two trajectories used to construct the reaction path. The trajectory ends with the initial state is shown in (a) while the other stops at the final state is shown in (b). For each trajectory, the potential energy and temperature changes are plotted.

S7. Error Analysis of CMD along Linear Combination CVs

For the further verification of insights from the free energy decomposition analysis, we chose $S(R_{C-O}, R_{O-Pt_A})$ for the reaction before the TS and R_{C-O} with $S(R_{O-Pt_B}, R_{O-Pt_C})$ after the TS. The related information about these two processes are shown in **Table S4** and **Table S5**.

Table S4. Free Energy Gradient and Standard Error by CMD before TS

$S(R_{C-O}, R_{O-Pt_A})$	FEG (eV/Å)	Standard Error (eV/Å)	Notes
1.300	-0.0125	0.0181	Initial State
1.200	-0.0650	0.1157	

1.100	-0.1052	0.0133	
1.000	-0.1822	0.0163	
0.900	-0.2500	0.0137	
0.800	-0.3294	0.0051	
0.700	-0.4123	0.0085	
0.600	-0.4915	0.0048	
0.500	-0.5785	0.0074	
0.400	-0.6460	0.0066	
0.300	-0.6772	0.0099	
0.200	-0.6967	0.0120	
0.100	-0.6501	0.0124	
0.000	-0.5959	0.0128	
-0.100	-0.5283	0.0138	
-0.200	-0.4669	0.0140	
-0.300	-0.4212	0.0193	
-0.400	-0.4057	0.0104	
-0.500	-0.3660	0.0079	
-0.600	-0.3623	0.0081	
-0.700	-0.3215	0.0090	
-0.800	-0.2668	0.0143	
-0.900	-0.0599	0.0205	
-1.000	0.2875	0.0631	Transition State

Table S5. Free Energy Gradient and Standard Error by CMD after TS

<i>S</i>	<i>R_{C-O}</i>	FEG/Error in <i>S</i> (eV/Å)	FEG/Error in <i>R</i> (eV/Å)	Notes
0.000	2.000	0.0010 / 0.0052	-0.0517 / 0.0121	Transition State

0.000	1.900	-0.0006 / 0.0084	0.1552 / 0.0152	
0.000	1.800	0.0026 / 0.0100	0.5836 / 0.0109	
0.000	1.700	0.0006 / 0.0035	0.9881 / 0.0264	
0.000	1.600	0.0016 / 0.0056	1.3813 / 0.0205	
0.000	1.500	0.0028 / 0.0066	1.5277 / 0.0326	
0.000	1.400	0.0060 / 0.0093	1.4170 / 0.0705	
0.100	1.400	-0.0164 / 0.0135	1.2353 / 0.0725	
0.200	1.400	-0.0979 / 0.0173	1.4393 / 0.0203	
0.300	1.400	-0.1203 / 0.0190	1.7492 / 0.0585	
0.400	1.400	-0.2673 / 0.0162	1.8273 / 0.0392	
0.500	1.400	-0.3346 / 0.0244	2.0948 / 0.0402	
0.600	1.400	-0.4181 / 0.0222	2.3947 / 0.0415	
0.700	1.400	-0.4513 / 0.0245	2.4030 / 0.0359	
0.800	1.400	-0.3981 / 0.0184	2.4927 / 0.0262	
0.900	1.400	-0.4008 / 0.0145	2.5652 / 0.0370	
1.000	1.400	-0.3070 / 0.0179	2.5080 / 0.0305	
1.100	1.400	-0.1865 / 0.0182	2.5720 / 0.0396	
1.200	1.400	-0.1272 / 0.0302	2.5886 / 0.0319	
1.200	1.280	-0.0086 / 0.0197	-0.0837 / 0.0355	Final State

References

- (1) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models, 2nd Edition*. John Wiley & Sons Ltd, 2004.
- (2) Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R. Constrained Reaction Coordinate Dynamics for the Simulation of Rare Events. *Chem. Phys. Lett.* **1989**, *156* (5), 472-477.

- (3) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23* (2), 187-199.
- (4) Fleurat-Lessard, P.; Ziegler, T. Tracing the Minimum-energy Path on the Free-energy Surface. *J. Chem. Phys.* **2005**, *123* (8), 084101.
- (5) Roux, B. The Calculation of the Potential of Mean Force using Computer Simulations. *Comput. Phys. Commun.* **1995**, *91*, 275-282.
- (6) Darve, E.; Rodriguez-Gomez, D.; Pohorille, A. Adaptive Biasing Force Method for Scalar and Vector Free Energy Calculations. *J. Chem. Phys.* **2008**, *128* (14), 144120.
- (7) Branduardi, D.; Gervasio, F. L.; Parrinello, M. From A to B in Free Energy Space. *J. Chem. Phys.* **2007**, *126* (5), 054103.
- (8) Zinovjev, K.; Martí, S.; Tuñón, I. A Collective Coordinate to Obtain Free Energy Profiles for Complex Reactions in Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8* (5), 1795-1801.