Electronic Supplementary Information

A Comparative Study on the Linear Scaling Relations for the Diffusion of S-Vacancies on MoS_2 and WS_2

Po-Yuan Wang,[1‡] Chun-Chi Yeh,[1,2‡] Ming-Jia Chiu,[1] Cheng-chau Chiu*[1,3,4]

[1] Department of Chemistry, National Sun Yat-sen University, Kaohsiung 80424 Taiwan
[2] National Chaochou Senior High School, Pingtung 92047, Taiwan
[3] Green Hydrogen Research Center, National Sun Yat-sen University, Kaohsiung 80424 Taiwan
[4] Center for Theoretical and Computational Physics, National Sun Yat-sen University, Kaohsiung 80424, Taiwan

*Corresponding author's email: ccchiu@mail.nsysu.edu.tw,
phone: +886-7-5252000#3935

‡These two authors contributed equally

S1. Further detail on the computational models and methods

S1.1. Density Functional Theory calculations

We have used the Vienna Ab initio Simulation Package (VASP, version 6.1.2)[S1–5] to conduct spin-polarized, 3-dimensional periodic Density Functional Theory (DFT) calculations using the generalized gradient approximation (GGA) exchange-correlation functional PBE named after *Perdew*, *Burke*, and *Ernzerhof*.[S6,7] As mentioned in the main text, the DFT data for MoS_2 have been taken from our earlier study.[S8,9] However, for the sake of a complete documentation, the corresponding parameters for the calculations are also shown here. All DFT data for MoS_2 and WS_2 mentioned here have been obtained from calculations using a periodically repeated, hexagonal 5×5 supercell of the H-phase monolayer structure of the respective compound. For MoS_2, the cell parameters $a = b$ of the 1×1 unit cell are optimized to 3.182 Å, while the corresponding value for WS_2 is 3.186 Å. To avoid unphysical interactions between a MoS_2/WS_2 monolayer and its period image, the height of the simulation cells is set to 20 Å if not stated otherwise. The valence electrons in the considered systems are described with a plane-wave basis with a kinetic energy cutoff of 520 eV. At variance, the projector augmented wave (PAW) method[S10,11] has been used to treat the core electrons. To evaluate the band energies, a Γcentered, 3×3×1 *k*-point grid is used to sample the Brillouin zone. To make the self-consistent field (SCF) calculation of the electronic structures smoother, Gaussian smearing with a smearing width of 0.05 eV is used. A SCF calculation is considered converged if the difference between the energies of two successive SCF steps is smaller than 10^{-6} eV.

For the geometry optimization, all atoms have been relaxed, while the stress tensor on the unit cell has not been considered. The convergence criteria for the geometry optimization requires the force on each atom to be smaller than 0.02 eV/Å for any atom. The optimization of the transition states (TSs) is carried out with a combination of the climbing-image nudged elastic band method (CI-NEB)[S12,13] and the Dimer method.[S14,15]

The vibrational frequencies are calculated numerically within the harmonic approximation. To reduce the computational costs, this work has only a partial Hessian matrix with contributions from the diffusing S-atom, and the three nearest metal atoms.

S1.2. Kinetic Monte Carlo simulations

The present work uses a self-implemented program for the kinetic Monte Carlo (kMC) simulations. In the following, the working principle of kMC simulations is briefly outlined using the system under study as an example. The defective transition metal disulfide structure, for instance, WS_2, as shown in **Figure S1 (a)**, is represented by a hexagonal grid. At a given point in time t, the transition metal disulfide model is in a configuration α. First, the configuration α is analyzed, and all possible S-vacancy diffusion processes i identified. Based on the occupation of the nearest-neighbor S-sites, each diffusion process will be assigned a rate constant k_i, calculated for a given temperature T according to Section 2 of the main text. Note that the same type of diffusion occurring at different sites, e.g., the diffusion processes 1 to 6 in **Figure S1 (a)**, count as separate processes. Each of the diffusion processes is a reaction that converts the configuration α to a not further specified, new configuration that is not α ("non-α"). Thus, k_{sum}, the sum over the rate constants k_i of all possible diffusion processes as visualized in **Figure S1 (b)**, can be interpreted as the rate constant of the unimolecular "decay" of α, i.e., the reaction of α to non-α. To estimate the time t+Δt' at which that reaction occurs, one uses the following equation:

$$p = exp\left(- k_{sum}\Delta t\right) \qquad \text{Eq S1}$$

which describes the probability p that ***no*** reaction has occurred after a time span Δt, see **Figure S1 (c)**. Once this probability is known, one could simply determine the period Δt' after which the occurrence of a reaction is "likely." However, it is unclear to which values of p this situation corresponds. Thus, in the context of kMC, one selects a random number r between 0 and 1. For $p \leq r$, the occurrence of a reaction is considered as probable. By replacing p in Eq S1 with r, one can determine the time t+Δt' at which the reaction from α to non-α becomes likely, where

$$\Delta t' = - \ln r / k_{sum} \qquad \text{Eq S2}$$

This time can be considered as the time when the reaction occurs.

The next question that needs to be clarified is which product configuration is being formed. Each of the possible diffusion processes shown in **Figure S1 (a)** leads to the formation of a different product configuration. As visualized in **Figure S1 (b)**, one uses

the rate constants k_i of the diffusion processes as weighting factors and randomly selects a process to execute, which leads to the formation of a new configuration. By continuously repeating the calculation of the reaction time and the random selection of a possible diffusion process, one can model how the local distribution of the S-vacancies on a MoS_2/WS_2 monolayer changes with time. Note that grid-based kMC, as used for this work, typically only allows the considered species, i.e., S-atoms and vacancies, to occupy specific positions described by the grid points. Potential distortion of the local environments resulting from the presence of a S-vacancy is not accounted for.

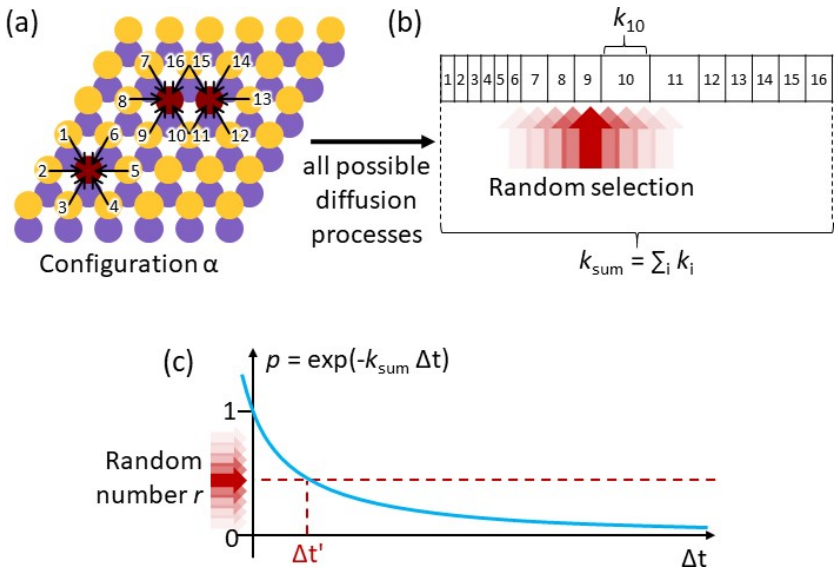

Figure S1 A brief outline of the kMC method: (a) shows the periodic model for a transition metal dichalcogenide monolayer, e.g., WS_2, here represented by a 5×5 hexagonal grid. Yellow: S-atom; violett: metal atom; red: S-vacancy. The black arrows and numbers indicate the possible diffusion processes, marking the moving direction of the S atoms. (b) schematically illustrates the calculation of k_{sum} and the random selection of a diffusion process to execute. For the random selection, the rate constant k_i for each reaction i is used as a weighting factor for the probability of being selected. (c) shows the probability p that no reaction has occurred as a function of the time span Δt as well as how this function and a random number r between 0 and 1 are used to estimate the time required for the reaction of configuration α.

S2. Calculated reaction and activation energies for S-vacancy diffusion on WS_2

Table S1. The reaction energies (E_r) and diffusion barriers (E_a) for the 256 considered unique diffusion processes in the forward direction (migration of an S-atom from site I to site F). The occupancy at the nearest-neighbor sites 1-8 is indicated with "x" (S-atom) and "o" (vacancy). All values are given in eV. The underlined entries are those diffusion steps for which we have not been able to identify the imaginary vibrational frequency for the TS. Consistent with our earlier work,[S8] we have removed the smallest real frequency, which is typically below 100 cm^{-1}, when calculating the rate constants.

Diffusion Process	E_r	E_a	Diffusion Process	E_r	E_a
1x2x3x4x5x6x7x8x	0.00	2.79	1x2x3x4x5x6x7o8o	-0.15	2.82
1o2x3x4x5x6x7x8x	-0.01	2.74	1o2o3o4x5x6x7x8x	0.42	3.28
1x2o3x4x5x6x7x8x	-0.04	2.81	1o2o3x4o5x6x7x8x	0.41	2.64
1x2x3o4x5x6x7x8x	-0.01	2.84	1o2o3x4x5o6x7x8x	0.03	1.30
1x2x3x4o5x6x7x8x	0.00	2.41	1o2o3x4x5x6o7x8x	0.21	3.08
<u>1x2x3x4x5o6x7x8x</u>	0.00	1.20	1o2o3x4x5x6x7o8x	0.28	3.08
1x2x3x4x5x6o7x8x	0.01	2.75	1o2o3x4x5x6x7x8o	0.21	3.00
1x2x3x4x5x6x7o8x	0.04	2.85	1o2x3o4o5x6x7x8x	0.08	2.50
1x2x3x4x5x6x7x8o	0.01	2.85	1o2x3o4x5o6x7x8x	0.15	1.19
1o2o3x4x5x6x7x8x	0.23	3.04	1o2x3o4x5x6o7x8x	-0.01	2.91
1o2x3o4x5x6x7x8x	-0.01	2.87	1o2x3o4x5x6x7o8x	0.06	2.94
1o2x3x4o5x6x7x8x	0.22	2.43	1o2x3o4x5x6x7x8o	0.00	2.92
1o2x3x4x5o6x7x8x	-0.14	1.12	1o2x3x4o5o6x7x8x	0.05	1.47
1o2x3x4x5x6o7x8x	0.00	2.81	1o2x3x4o5x6o7x8x	0.00	2.37
1o2x3x4x5x6x7o8x	0.05	2.79	1o2x3x4o5x6x7o8x	0.30	2.50
1o2x3x4x5x6x7x8o	0.00	2.73	1o2x3x4o5x6x7x8o	0.35	2.47
1x2o3o4x5x6x7x8x	0.15	2.97	<u>1o2x3x4x5o6o7x8x</u>	0.00	1.11
1x2o3x4o5x6x7x8x	-0.07	2.36	1o2x3x4x5o6x7o8x	-0.06	1.09
1x2o3x4x5o6x7x8x	-0.07	1.12	1o2x3x4x5o6x7x8o	-0.42	1.06
1x2o3x4x5x6o7x8x	-0.05	2.75	1o2x3x4x5x6o7o8x	-0.21	2.88
1x2o3x4x5x6x7o8x	0.00	2.86	1o2x3x4x5x6o7x8o	0.01	2.86
1x2o3x4x5x6x7x8o	-0.05	2.85	1o2x3x4x5x6x7o8o	-0.11	2.70
1x2x3o4o5x6x7x8x	-0.13	2.38	1x2o3o4o5x6x7x8x	-0.05	2.47
1x2x3o4x5o6x7x8x	0.30	1.41	1x2o3o4x5o6x7x8x	0.39	1.47
1x2x3o4x5x6o7x8x	0.00	2.73	1x2o3o4x5x6o7x8x	0.11	2.81
1x2x3o4x5x6x7o8x	0.05	2.90	1x2o3o4x5x6x7o8x	0.20	3.02
1x2x3o4x5x6x7x8o	0.00	2.98	1x2o3o4x5x6x7x8o	0.12	3.08
1x2x3x4o5o6x7x8x	0.00	1.26	1x2o3x4o5o6x7x8x	-0.10	1.19
1x2x3x4o5x6o7x8x	-0.22	2.21	1x2o3x4o5x6o7x8x	-0.30	2.20
1x2x3x4o5x6x7o8x	0.07	2.43	1x2o3x4o5x6x7o8x	0.00	2.39
1x2x3x4o5x6x7x8o	0.13	2.51	1x2o3x4o5x6x7x8o	0.05	2.44
1x2x3x4x5o6o7x8x	0.14	1.26	1x2o3x4x5o6o7x8x	0.06	1.14
1x2x3x4x5o6x7o8x	0.07	1.19	1x2o3x4x5o6x7o8x	0.00	1.09
1x2x3x4x5o6x7x8o	-0.30	1.12	1x2o3x4x5o6x7x8o	-0.38	1.06
1x2x3x4x5x6o7o8x	-0.23	2.81	1x2o3x4x5x6o7o8x	-0.28	2.80
1x2x3x4x5x6o7x8o	0.01	2.89	1x2o3x4x5x6o7x8o	-0.06	2.88

Diffusion Process	E_r	E_a
1x2o3x4x5x6x7o8o	-0.20	2.82
1x2x3o4o5o6x7x8x	0.15	1.39
1x2x3o4o5x6o7x8x	-0.35	2.12
1x2x3o4o5x6x7o8x	-0.05	2.40
1x2x3o4o5x6x7x8o	0.00	2.60
1x2x3o4x5o6o7x8x	0.42	1.48
1x2x3o4x5o6x7o8x	0.38	1.44
1x2x3o4x5o6x7x8o	0.00	1.23
1x2x3o4x5x6o7o8x	-0.21	2.79
1x2x3o4x5x6o7x8o	0.00	2.93
1x2x3o4x5x6x7o8o	-0.12	2.96
1x2x3x4o5o6o7x8x	-0.05	1.41
1x2x3x4o5o6x7o8x	0.10	1.29
1x2x3x4o5o6x7x8o	-0.15	1.23
1x2x3x4o5x6o7o8x	-0.41	2.22
1x2x3x4o5x6o7x8o	-0.08	2.43
1x2x3x4o5x6x7o8o	0.05	2.52
1x2x3x4x5o6o7o8x	-0.03	1.27
1x2x3x4x5o6o7x8o	-0.15	1.04
1x2x3x4x5o6x7o8o	-0.39	1.08
1x2x3x4x5x6o7o8o	-0.42	2.86
1o2o3o4o5x6x7x8x	0.41	2.76
1o2o3o4x5o6x7x8x	0.47	1.47
1o2o3o4x5x6o7x8x	0.37	3.20
1o2o3o4x5x6x7o8x	0.48	3.33
1o2o3o4x5x6x7x8o	0.39	3.27
1o2o3x4o5o6x7x8x	0.20	1.47
1o2o3x4o5x6o7x8x	0.18	2.62
1o2o3x4o5x6x7o8x	0.49	2.70
1o2o3x4o5x6x7x8o	0.52	2.68
1o2o3x4x5o6o7x8x	0.15	1.25
1o2o3x4x5o6x7o8x	0.10	1.26
1o2o3x4x5o6x7x8o	-0.27	1.25
1o2o3x4x5x6o7o8x	0.00	3.14
1o2o3x4x5x6o7x8o	0.19	3.12
1o2o3x4x5x6x7o8o	0.10	2.97
1o2x3o4o5o6x7x8x	0.16	1.30
1o2x3o4o5x6o7x8x	-0.13	2.43
1o2x3o4o5x6x7o8x	0.17	2.59
1o2x3o4o5x6x7x8o	0.20	2.58
1o2x3o4x5o6o7x8x	0.27	1.27
1o2x3o4x5o6x7o8x	0.25	1.22
1o2x3o4x5o6x7x8o	-0.13	1.19
1o2x3o4x5x6o7o8x	-0.19	2.93
1o2x3o4x5x6o7x8o	0.00	3.03

Diffusion Process	E_r	E_a
1o2x3o4x5x6x7o8o	-0.09	2.92
1o2x3x4o5o6o7x8x	0.00	1.15
1o2x3x4o5o6x7o8x	0.16	1.28
1o2x3x4o5o6x7x8o	-0.08	1.29
1o2x3x4o5x6o7o8x	-0.18	2.45
1o2x3x4o5x6o7x8o	0.13	2.57
1o2x3x4o5x6x7o8o	0.28	2.47
1o2x3x4x5o6o7o8x	-0.15	1.10
1o2x3x4x5o6o7x8o	-0.27	1.00
1o2x3x4x5o6x7o8o	-0.48	1.03
1o2x3x4x5x6o7o8o	-0.37	2.83
1x2o3o4o5o6x7x8x	0.20	1.38
1x2o3o4o5x6o7x8x	-0.28	2.19
1x2o3o4o5x6x7o8x	0.04	2.46
1x2o3o4o5x6x7x8o	0.06	2.56
1x2o3o4x5o6o7x8x	0.48	1.51
1x2o3o4x5o6x7o8x	0.47	1.49
1x2o3o4x5o6x7x8o	0.08	1.29
1x2o3o4x5x6o7o8x	-0.10	2.87
1x2o3o4x5x6o7x8o	0.09	3.00
1x2o3o4x5x6x7o8o	0.00	2.98
1x2o3x4o5o6o7x8x	-0.16	1.12
1x2o3x4o5o6x7o8x	0.00	1.17
1x2o3x4o5o6x7x8o	-0.26	1.19
1x2o3x4o5x6o7o8x	-0.49	2.22
1x2o3x4o5x6o7x8o	-0.17	2.42
1x2o3x4o5x6x7o8o	-0.04	2.42
1x2o3x4x5o6o7o8x	-0.10	1.16
1x2o3x4x5o6o7x8o	-0.25	0.97
1x2o3x4x5o6x7o8o	-0.47	1.02
1x2o3x4x5x6o7o8o	-0.48	2.85
1x2x3o4o5o6o7x8x	0.08	1.37
1x2x3o4o5o6x7o8x	0.26	1.45
1x2x3o4o5o6x7x8o	0.00	1.23
1x2x3o4o5x6o7o8x	-0.52	2.16
1x2x3o4o5x6o7x8o	-0.20	2.37
1x2x3o4o5x6x7o8o	-0.06	2.50
1x2x3o4x5o6o7o8x	0.27	1.52
1x2x3o4x5o6o7x8o	0.13	1.32
1x2x3o4x5o6x7o8o	-0.08	1.21
1x2x3o4x5x6o7o8o	-0.39	2.89
1x2x3x4o5o6o7o8x	-0.20	1.28
1x2x3x4o5o6o7x8o	-0.16	1.13
1x2x3x4o5o6x7o8o	-0.20	1.18
1x2x3x4o5x6o7o8o	-0.41	2.36

Diffusion Process	E_r	E_a
1x2x3x4x5o6o7o8o	-0.47	1.00
1o2o3o4o5o6x7x8x	0.41	1.51
1o2o3o4o5x6o7x8x	0.17	2.62
1o2o3o4o5x6x7o8x	0.49	2.85
1o2o3o4o5x6x7x8o	0.50	2.80
1o2o3o4x5o6o7x8x	0.55	1.52
1o2o3o4x5o6x7o8x	0.55	1.50
1o2o3o4x5o6x7x8o	0.17	1.47
1o2o3o4x5x6o7o8x	0.19	3.26
1o2o3o4x5x6o7x8o	0.34	3.34
1o2o3o4x5x6x7o8o	0.30	3.24
1o2o3x4o5o6o7x8x	0.13	1.33
1o2o3x4o5o6x7o8x	0.29	1.45
1o2o3x4o5o6x7x8o	0.05	1.47
1o2o3x4o5x6o7o8x	0.00	2.65
1o2o3x4o5x6o7x8o	0.29	2.80
1o2o3x4o5x6x7o8o	0.46	2.67
1o2o3x4x5o6o7o8x	0.00	1.22
1o2o3x4x5o6o7x8o	-0.14	1.17
1o2o3x4x5o6x7o8o	-0.33	1.22
1o2o3x4x5x6o7o8o	-0.19	3.07
1o2x3o4o5o6o7x8x	0.10	1.28
1o2x3o4o5o6x7o8x	0.28	1.37
1o2x3o4o5o6x7x8o	0.02	1.28
1o2x3o4o5x6o7o8x	-0.29	2.51
1o2x3o4o5x6o7x8o	0.00	2.54
1o2x3o4o5x6x7o8o	0.17	2.57
1o2x3o4x5o6o7o8x	0.14	1.32
1o2x3o4x5o6o7x8o	0.00	1.14
1o2x3o4x5o6x7o8o	-0.18	1.17
1o2x3o4x5x6o7o8o	-0.34	3.01
1o2x3x4o5o6o7o8x	-0.13	1.19
1o2x3x4o5o6o7x8o	-0.10	1.18
1o2x3x4o5o6x7o8o	-0.11	1.23
1o2x3x4o5x6o7o8o	-0.17	2.46
1o2x3x4x5o6o7o8o	-0.55	0.97
1x2o3o4o5o6o7x8x	0.11	1.34
1x2o3o4o5o6x7o8x	0.30	1.43
1x2o3o4o5o6x7x8o	0.04	1.23
1x2o3o4o5x6o7o8x	-0.46	2.21
1x2o3o4o5x6o7x8o	-0.17	2.40
1x2o3o4o5x6x7o8o	0.00	2.40
1x2o3o4x5o6o7o8x	0.33	1.55
1x2o3o4x5o6o7x8o	0.18	1.35
1x2o3o4x5o6x7o8o	0.00	1.25

Diffusion Process	E_r	E_a
1x2o3o4x5x6o7o8o	-0.30	2.94
1x2o3x4o5o6o7o8x	-0.29	1.16
1x2o3x4o5o6o7x8o	-0.28	1.10
1x2o3x4o5o6x7o8o	-0.30	1.14
1x2o3x4o5x6o7o8o	-0.49	2.36
1x2o3x4x5o6o7o8o	-0.55	0.95
1x2x3o4o5o6o7o8x	-0.05	1.42
1x2x3o4o5o6o7x8o	-0.02	1.26
1x2x3o4o5o6x7o8o	-0.04	1.20
1x2x3o4o5x6o7o8o	-0.50	2.31
1x2x3o4x5o6o7o8o	-0.17	1.30
1x2x3x4o5o6o7o8o	-0.41	1.10
1o2o3o4o5o6o7x8x	0.32	1.46
1o2o3o4o5o6x7o8x	0.51	1.55
1o2o3o4o5o6x7x8o	0.26	1.52
1o2o3o4o5x6o7o8x	0.01	2.74
1o2o3o4o5x6o7x8o	0.27	2.76
1o2o3o4o5x6x7o8o	0.46	2.76
1o2o3o4x5o6o7o8x	0.41	1.57
1o2o3o4x5o6o7x8o	0.26	1.38
1o2o3o4x5o6x7o8o	0.12	1.44
1o2o3o4x5x6o7o8o	0.00	3.28
1o2o3x4o5o6o7o8x	0.00	1.36
1o2o3x4o5o6o7x8o	0.02	1.37
1o2o3x4o5o6x7o8o	0.02	1.40
1o2o3x4o5x6o7o8o	-0.01	2.74
1o2o3x4x5o6o7o8o	-0.41	1.16
1o2x3o4o5o6o7o8x	-0.02	1.35
1o2x3o4o5o6o7x8o	0.00	1.23
1o2x3o4o5o6x7o8o	0.01	1.24
1o2x3o4o5x6o7o8o	-0.27	2.49
1o2x3o4x5o6o7o8o	-0.26	1.12
1o2x3x4o5o6o7o8o	-0.32	1.14
1x2o3o4o5o6o7o8x	-0.02	1.39
1x2o3o4o5o6o7x8o	-0.01	1.23
1x2o3o4o5o6x7o8o	0.00	1.23
1x2o3o4o5x6o7o8o	-0.46	2.30
1x2o3o4x5o6o7o8o	-0.12	1.32
1x2o3x4o5o6o7o8o	-0.51	1.04
1x2x3o4o5o6o7o8o	-0.28	1.27
1o2o3o4o5o6o7o8x	0.20	1.51
1o2o3o4o5o6o7x8o	0.21	1.44
1o2o3o4o5o6x7o8o	0.24	1.47
1o2o3o4o5x6o7o8o	0.00	2.73
1o2o3o4x5o6o7o8o	0.00	1.36

Diffusion Process	E_r	E_a
1o2o3x4o5o6o7o8o	-0.20	1.30
1o2x3o4o5o6o7o8o	-0.21	1.23
1x2o3o4o5o6o7o8o	-0.24	1.23

Diffusion Process	E_r	E_a
1o2o3o4o5o6o7o8o	0.00	1.38

S3. Approximation of the partition function term in the Eyring Equation

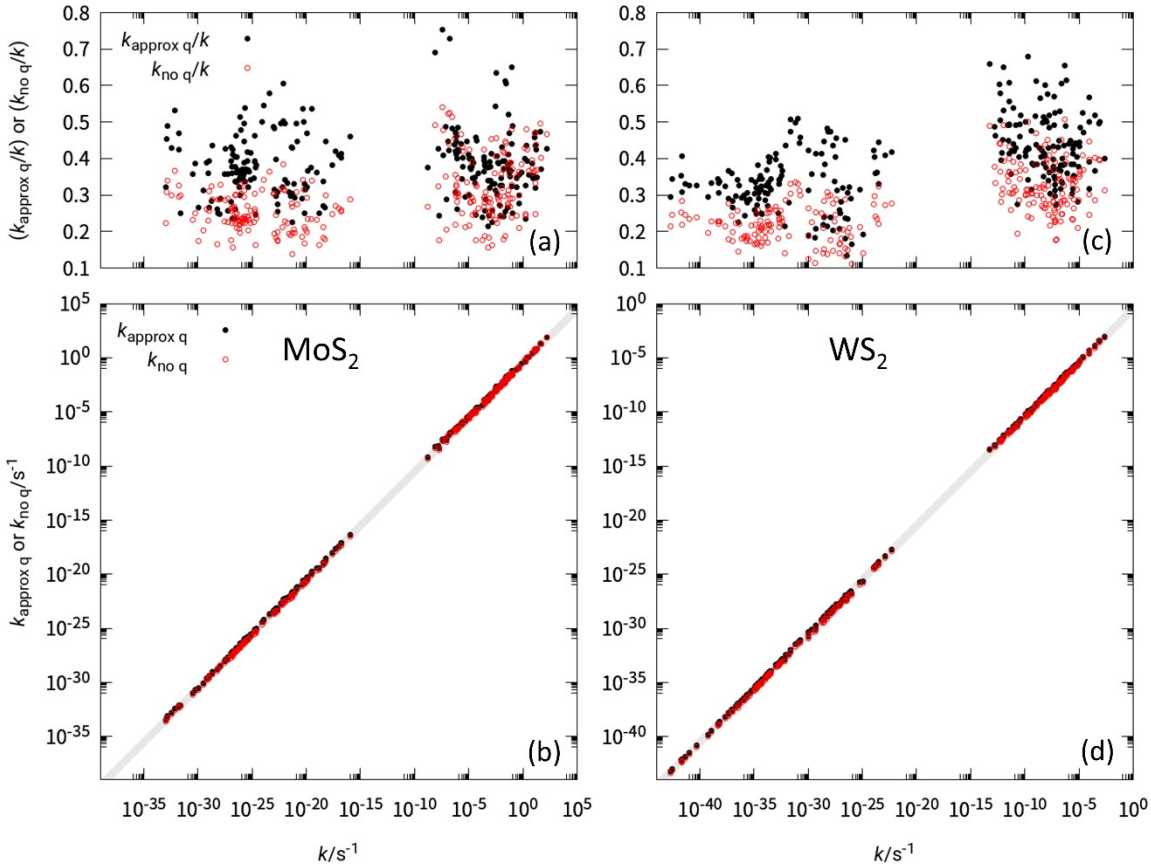

Figure S2. Approximated rate constants $k_{approx\ q}$ and $k_{no\ q}$ for the S-vacancy diffusion processes on MoS_2 and WS_2 at 300 K compared to the corresponding rate constants k obtained self-consistently at DFT level. (a) and (c) show the ratio between the approximated rate constants, $k_{approx\ q}$ or $k_{no\ q}$, and the actual value at DFT level, k, plotted against k. Independent of the actual value of k, both $k_{approx\ q}$ and $k_{no\ q}$ tend to underestimate it. In (b) and (d), the approximated rate constants are plotted against the actual DFT-level rate constant. Data points in the gray area indicate that an approximated rate constant is between k and $k/10$.

S4. Comparison of diffusion barriers obtained with different methods or approximations

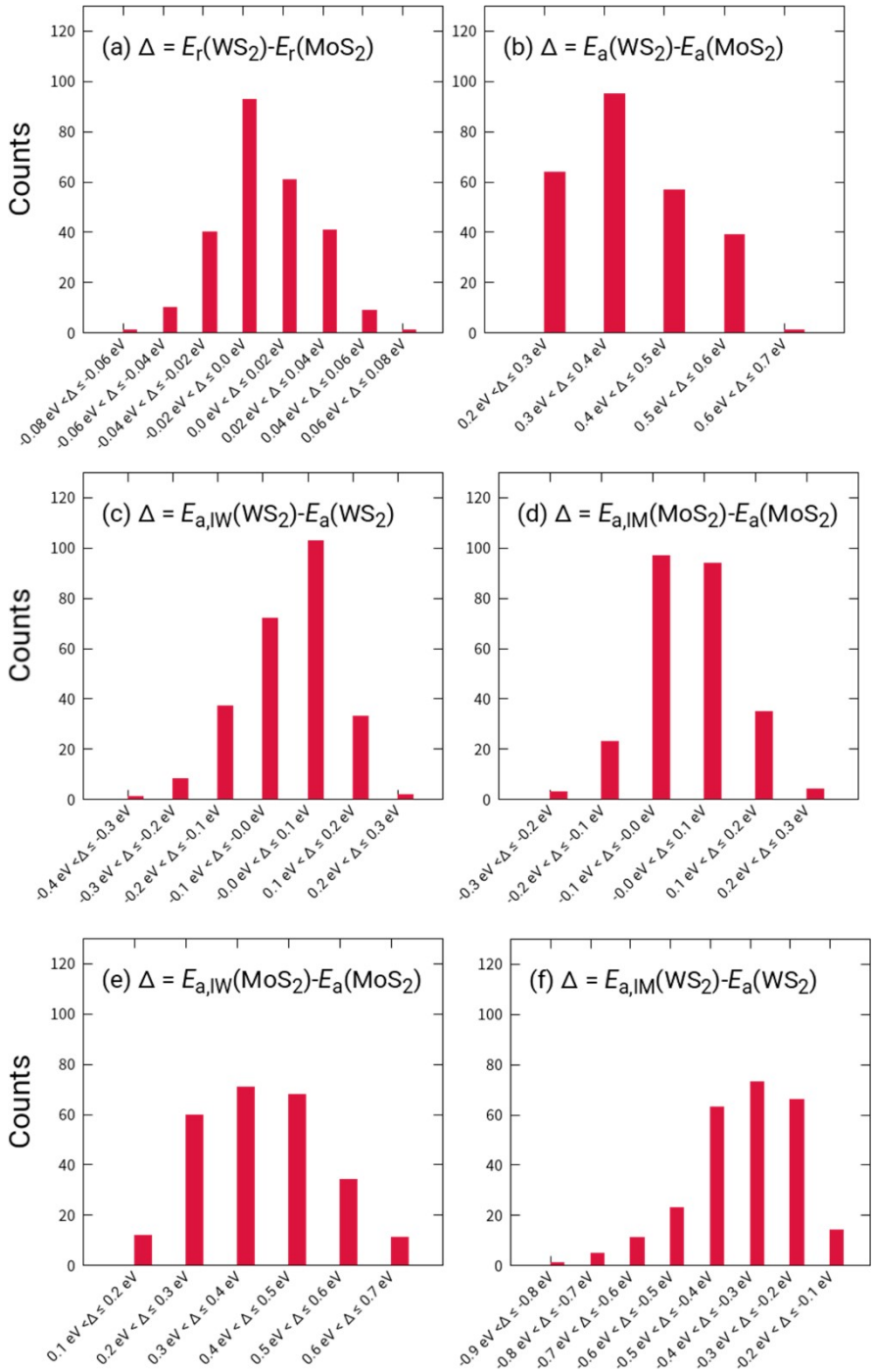

Figure S3. The first two histograms show the distribution of the difference in the (a) E_r and (b) E_a values calculated at DFT level for the different S-vacancy diffusion processes on WS_2 and MoS_2. The histograms (c) to (f) compare the barriers obtained from linear scaling $E_{a,lY}$ to those calculated at DFT level. Note that the subscript lY denotes how the applied linear scaling relations have been obtained, while the parenthesis indicates to which system the linear scaling relations have been applied.

S5. Linear scaling relations for S-vacancy diffusion on WS_2

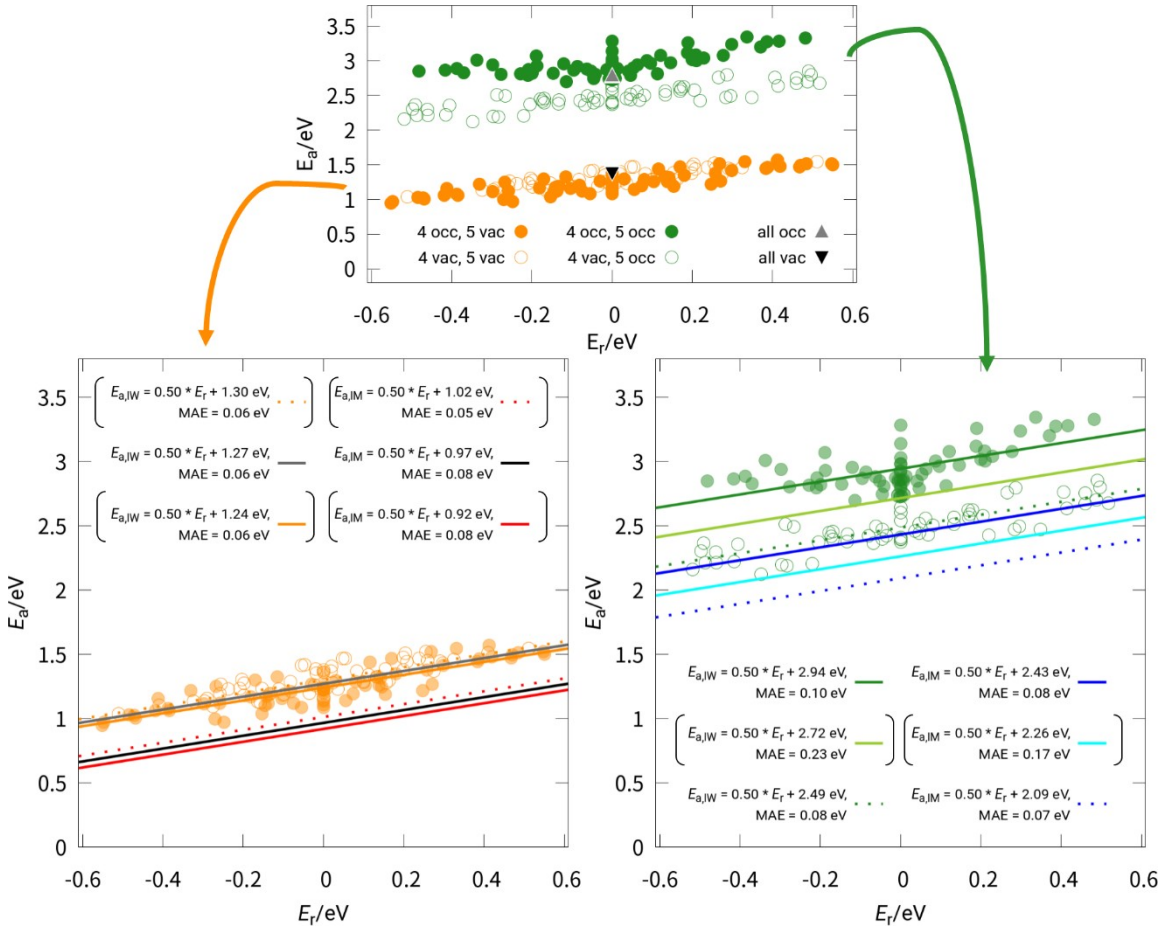

Figure S4. The top figure shows the E_r and E_a values for the S-vacancy diffusion on WS_2. The data points are marked according to the occupation of sites 4 and 5. The diffusion processes with sites 1-8 all occupied by an S atom and all vacant, respectively, are also highlighted. The figures at the bottom left and right focus on the data for the diffusions with site 5 and site 4, respectively, being vacant. In addition, the derived linear scaling relations ($E_{a,IW}$) and the associated mean absolute error (MAE) are shown. Corresponding relations obtained for MoS_2 ($E_{a,IM}$) reported in Refs S8 and S9 are also shown for comparison. In the figure bottom left, the dotted orange (red for MoS_2) line refers to the data points where sites 4 and 5 are both vacant (empty, orange circles), the orange (red), solid line to those where site 4 is occupied and site 5 vacant (filled, orange circles). The gray (black), solid line takes both sets of data points into account. In the bottom right figure, the dark green (dark blue for MoS_2), solid lines are fitted to the data points where sites 4 and 5 are both occupied (filled, dark green circles). The dark green (dark blue), dotted line is associated with the diffusion where site 5 is occupied and site 4 vacant (empty, dark green circles). The light green (cyan), solid line is fitted to all data points where site 5 is occupied. The linear scaling relations not used for the present work are in parentheses.

S6. Comment on the gradient of the linear scaling relations

Figure S4 has shown that all linear scaling relations mentioned in this work, whether for MoS_2 or WS_2, feature the general form of

$$E_{\mathrm{a,lY}} = 0.5 \times E_{\mathrm{r,lY}} + c \qquad \text{Eq S3}$$

In other words, the gradient is always 0.5. This is not a coincidence and is explained in the following. The diffusion processes considered in this work can be divided into two groups. In the first group of diffusion processes, the initial and product state structures are symmetrically equivalent, so their reaction energies E_r are 0. Due to symmetry, these diffusion steps are also their own reverse process. All these processes contribute to data points where E_r is 0 and E_a a not further specified value. The second group consists of all remaining diffusion processes with inequivalent initial and product states. Without loss of generality, one can assume that the "forward" diffusion is always an endothermic process with the reaction energy E_r' and barrier E_a'. Then, the reverse diffusion, which is also included in the considered diffusion processes, is exothermic and has the reaction energy $-E_\mathrm{r}'$ and a barrier of $E_\mathrm{a}'–E_\mathrm{r}'$. A line passing through these two data points will always have the gradient g, with

$$g = \frac{E_a^{'} - \left(E_a^{'} - E_r^{'}\right)}{E_r^{'} - \left(-E_r^{'}\right)} = \frac{E_r^{'}}{2E_r^{'}} = 0.5 \qquad \text{Eq S4}$$

As the data points for all pairs of forward and backward reactions can be connected by a line with a gradient of 0.5, it is easy to see why the derived linear scaling relations also have a gradient of 0.5. The diffusion processes from the first group all have E_r values of 0. In other words, their E_r values are exactly at the center of the E_r values of any pair of forward and backward diffusion with inequivalent initial and product states. In this case, the data points associated with the first group of diffusion reactions do not affect the gradient of the linear scaling relations, only the additive constant c in Eq S3.

S7. Comment on the differences between the diffusion barriers on MoS_2 and WS_2

The generally higher diffusion barriers for S-vacancies on WS_2 compared to those on MoS_2 are likely associated with the stronger metal-S bonds in the former. The stronger bonds can be seen from Figure S5. For this figure, we have analyzed the 5×5 super cells used to describe the defective MoS_2 and WS_2 monolayers in the DFT simulations, see Section S1.1, and calculated the average binding energy $E_{M\text{-}S}$ between a sulfur atom and the metal sites (M = Mo or W) of the monolayer *via* the equation

$$E_{M-S} = \left[E(M_{25}S_{50-n}) + \frac{n}{8} \times E(S_8) - E(M_{25}S_{50})\right] \Big/ n \qquad \text{Eq S5}$$

$E(M_{25}S_{50})$ is the electronic energy of the 5×5 supercell of pristine MS_2, while $E(M_{25}S_{50\text{-}n})$ is the electronic energy of the simulation cell with n S-vacancies. $E(S_8)$ is the energy of a S_8 molecule in gas phase, which is used as a more or less arbitrarily chosen reference for the energy of the S atoms. The figure shows that breaking the metal-S bonds to create a vacancy structure on WS_2 always requires more energy than the corresponding process on MoS_2.

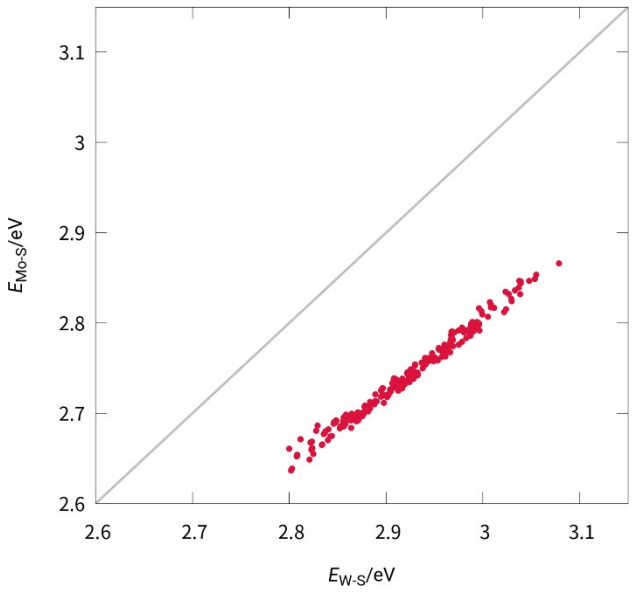

Figure S5. Binding energies $E_{M\text{-}S}$ (M = Mo or W) calculated according to Eq S5. $E_{M\text{-}S}$ values obtained from corresponding vacancy structures on MoS_2 and WS_2 are plotted against each other. The circumstance that all data points are located below the diagonal means that the W-S interaction is generally stronger than the Mo-S interaction.

The question that comes up next is why the W-S interaction is stronger than the Mo-S interaction. The stronger metal-S bond in WS_2 is also reflected in the electronic density of states (DOS), as shown in Figure S6. Before discussing the figure, we should address some technical details. The figure shows the DOS for MoS_2 and WS_2 evaluated with a 1×1 unit cell and a 15×15×1 *k*-point grid. The other simulation parameters are as described in Section S1.1. In addition, we also show the DOS for a hypothetical system in which the layers of metal atoms and S-atoms, respectively, are separated by a vacuum space of 16.67 Å width. For this calculation, the height of the unit cell has been increased to 50 Å while keeping the other settings unchanged. Unlike most DOS plots in literature, which are aligned at the Fermi level, we have aligned the DOS plots in Figure S6 at the energy, where the band arising from the sulfur *s*-orbitals has its maximum DOS value. If the maximum DOS values differ for different spin directions, the DOS is aligned at the band of the majority spin. This uncommon choice for the energy reference is motivated by the circumstance that it allows a clearer idea of how the band energies shift once the metal and the S-layers interact. The sulfur *s*-orbital is the energetically lowest-lying atomic orbital that is present in both WS_2 and MoS_2. The reason for aligning at a low-lying band is that its band energy is expected to be least affected by the interaction between the metal and the sulfur atoms.

Figure S6 (b) and (c) show that the DOS of the valence bands of MoS_2 and WS_2 are roughly located at the same energies. In contrast, the DOS for the separated metal and S-layers in Figure S6 (a) and (d) show larger differences, particularly the states around the Fermi level, which are mainly associated with the *d*-states of the metal atoms. To quantify the energy difference between the electrons in the *d*-states of the isolated Mo- and W-layers, we have estimated the center of the occupied *d*-states via the following equation:

$$d-band\ center = \frac{\int_{-\infty}^{E_{Fermi}} PDOS_d(E)\cdot E dE}{\int_{-\infty}^{E_{Fermi}} PDOS_d(E) dE} \approx \frac{\sum_{E=-\infty}^{E_{Fermi}} PDOS_d(E)\cdot E}{\sum_{E=-\infty}^{E_{Fermi}} PDOS_d(E)}$$

Eq S6

where $PDOS_d(E)$ is the DOS projected on the metal *d*-orbitals at a given energy *E*, for which the $PDOS_d(E)$ value is recorded. As visible from Figure S6 (a) and (d), the *d*-states of the W-layer are slightly higher in energy than the *d*-states of the Mo-layer. Given the similarity in the DOS for MoS_2 and WS_2, it can be stated that the formation of WS_2 gives rise to a stronger stabilization of the monolayer structure compared to the situation of MoS_2. As the diffusion of the S-vacancies comes along with breaking metal-S bonds to allow an S-atom to move to the vacancy site, this stronger metal-sulfur interaction may rationalize the higher diffusion barriers determined for WS_2.

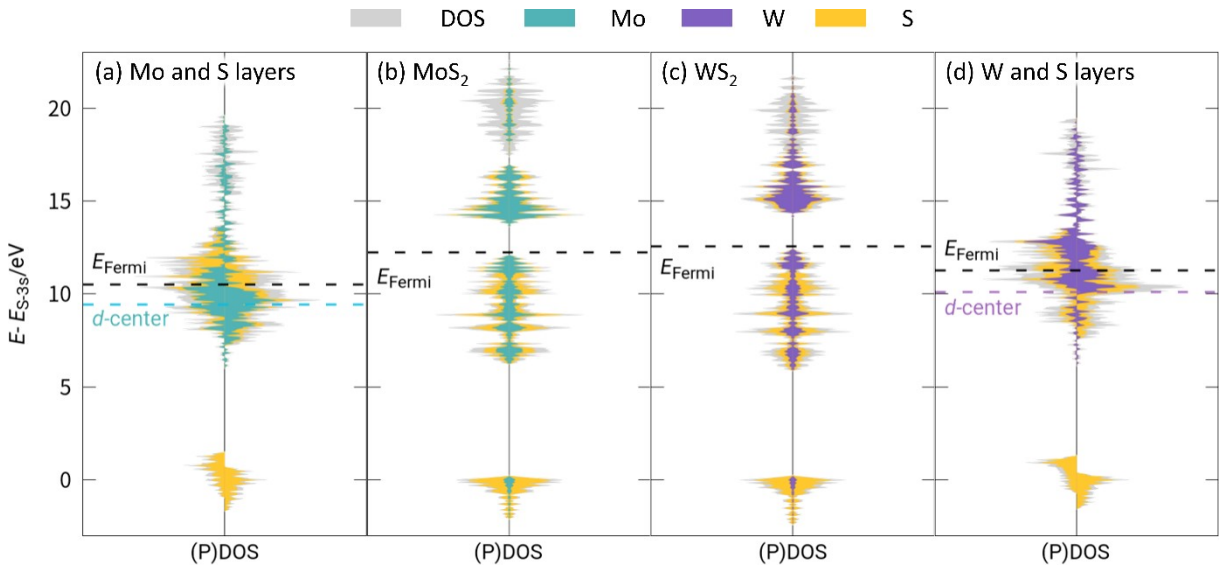

Figure S6. DOS and PDOS for pristine (b) MoS_2 and (c) WS_2 monolayers and, in (a) and (d), the corresponding systems, in which the metal and the sulfur layers are separated by vacuum space. A black dashed line marks the Fermi level. The dashed lines in the same color as the PDOS for the metal states mark the center of the occupied *d*-states. The band energies are aligned at the energy where the (majority spin) sulfur 3*s*-band shows its maximum in the DOS. Note that the value of the PDOS is indicated by the *visible* area of the respective color.

References

S1 G. Kresse and J. Hafner, Ab initio molecular dynamics for liquid metals, *Phys. Rev. B*, 1993, **47**, 558–561.

S2 G. Kresse and J. Hafner, Ab initio molecular-dynamics simulation of the liquid-metalamorphous- semiconductor transition in germanium, *Phys. Rev. B*, 1994, **49**, 14251–14269.

S3 G. Kresse and J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.*, 1996, **6**, 15–50.

S4 G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B - Condens. Matter Mater. Phys.*, 1996, **54**, 11169–11186.

S5 G. Kresse and J. Hafner, Norm-conserving and ultrasoft pseudopotentials for first-row and transition elements, *J. Phys. Condens. Matter*, 1994, **6**, 8245.

S6 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

S7 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized Gradient Approximation Made Simple[Phys. Rev. Lett. 77, 3865 (1996)], *Phys. Rev. Lett.*, 1997, **78**, 1396–1396.

S8 P.-Y. Wang, B.-A. Chen, Y.-C. Lee and C. Chiu, First-principles modeling of the highly dynamical surface structure of a MoS2 catalyst with S-vacancies, *Phys. Chem. Chem. Phys.*, 2022, **24**, 24166–24172.

S9 P.-Y. Wang, B.-A. Chen, Y.-C. Lee and C. Chiu, Correction: First-principles modeling of the highly dynamical surface structure of a MoS2 catalyst with S-vacancies, *Phys. Chem. Chem. Phys.*, 2023, **25**, 21109–21110.

S10 P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B*, 1994, **50**, 17953–17979.

S11 G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B*, 1999, **59**, 1758–1775.

S12 G. Henkelman, B. P. Uberuaga and H. Jónsson, Climbing image nudged elastic band method for finding saddle points and minimum energy paths, *J. Chem. Phys.*, 2000, **113**, 9901–9904.

S13 G. Henkelman and H. Jónsson, Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points, *J. Chem. Phys.*, 2000, **113**, 9978–9985.

S14 G. Henkelman and H. Jónsson, A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives, *J. Chem. Phys.*, 1999, **111**, 7010–7022.

S15 A. Heyden, A. T. Bell and F. J. Keil, Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method, *J. Chem. Phys.*, 2005, **123**, 1–14.