This version of the ESI replaces the one published on 13.09.2022 to correct the error in the geometry provided.

Electronic Supplementary Information

First-Principles Modeling of the Highly Dynamical Surface Structure of a MoS_2 Catalyst with S-Vacancies

Po-Yuan Wang,[a] Bo-An Chen,[b] Yu-Chi Lee,[b] Cheng-chau Chiu*[a,c]

[a.] Department of Chemistry, National Sun Yat-sen University, Kaohsiung 80424, Taiwan.

[b.] Institute of Atomic and Molecular Sciences, Academia Sinica, Taipei 10617, Taiwan

[c.] Center for Theoretical and Computational Physics, National Sun Yat-sen University, Kaohsiung 80424, Taiwan

*Corresponding author's email: ccchiu@mail.nsysu.edu.tw

S1. Computational details

S1.1. DFT calculations

All spin-polarized, periodic DFT calculations were performed using the Vienna Ab initio Simulation Package (VASP), version 6.1.2.[S1–5] The structures considered in this work are all described with a 5×5 periodically repeated hexagonal supercell of 2H-MoS_2 with the cell constants $a = b = 15.910$ Å. This corresponds to optimized cell constants of 3.182 Å for the 1×1 unit cell of 2H-MoS_2. In the direction perpendicular to the MoS_2 basal plane, the simulation cell is chosen to have a height of 20 Å so the periodically repeated monolayers are separated from each other.

The electronic structures of the considered systems are evaluated with the generalized gradient approximation (GGA) exchange-correlation functional proposed by *Perdew*, *Burke*, and *Ernzerhof* (PBE).[S6,7] The core electrons in the modeled structures are treated with the projector augmented wave (PAW) method,[S8,9] while the valence electrons are explicitly described using a plane-wave basis with a cutoff energy of 520 eV. The Brillouin zone of the modeled structures is sampled with a 3×3×1, Γ-centered *k*-point grid. In addition, Gaussian smearing with a smearing width of 0.05 eV is used. The convergence criteria for the self-consistent field (SCF) calculation of the electronic energies requires the change in the energy between two SCF iteration steps to be less than 10^{-6} eV. Note that corrections to account for dispersive interactions have been omitted here, as they are expected to have a negligible effect on the calculated energetics.

During the geometry optimization, all atoms have been relaxed. However, the stress tensor on the unit cell has been neglected. The force tolerance of any optimized geometry is set to 0.02 eV/Å for any atom. For optimizing the transition state (TS) geometries, we have used two methods: the climbing-image nudged elastic band method (CI-NEB)[S10,11] and the Dimer method.[S12,13] For a TS optimization from scratch, we perform a CI-NEB calculation with eight images for pre-optimization and then continue to optimize with the Dimer methods. However, for several cases, we have been able to skip the CI-NEB step, as we could generate a sufficiently good initial guess geometry for the Dimer calculation from an optimized TS geometry of a diffusion process with a similar environment.

Numerical normal mode analyses within the harmonic approximation for all optimized geometries have been performed. To reduce the computational costs, only the contributions of the diffusing S-atom and the three nearest Mo-atoms to the Hessian have been considered for the normal mode analysis. In addition, it has been checked whether minimum geometries and transition states have exactly zero and one imaginary frequency, respectively. This is, with a few exceptions discussed below, always the case.

S1.2. Energetics

As mentioned in the main text, there are 256 distinct occupation patterns at the eight nearest-neighbor sites. For each of the 256 discussed diffusion processes, we obtained the energies for the the initial state (IS), in which the diffusing S-atom is on the "I" site, and the final state (FS), with the diffusing atom is on the "F" site, see Figure 1 of the main text. In most cases, there is only one TS along the diffusion pathway, see Figure S1(a). However, some cases, particularly when site 5 is unoccupied, feature two TSs along the diffusion pathway. To keep the kMC model simple, we have assumed that these processes also only feature one diffusion barrier, with the energetically higher-lying TS being the only TS of that diffusion process (see Figure S1(b)). With that, every diffusion process from I to F features one diffusion barrier E_a, defined as the energy difference between the TS and the IS, and a reaction energy E_r, calculated as the energy difference between FS and IS.

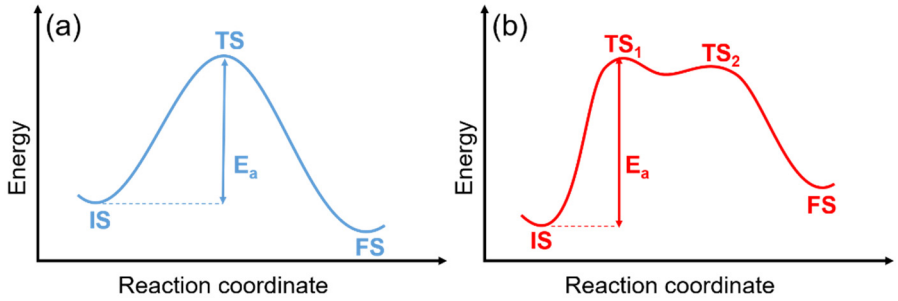

Figure S1. Schematic representation of the potential surface for a diffusion process with (a) one transition state and (b) two transition states. IS, $TS_{(i)}$, and FS indicate the diffusion processes' initial state, transition state(s), and final state. The barriers considered for the diffusion steps are marked with E_a.

In a few diffusion processes involving two TSs, the TSs can be very close to the intermediate state along the diffusion path. In such situations, we are sometimes unable to identify the imaginary vibrational frequency associated with the TS. Instead, the normal mode describing that reaction coordinate features a small real frequency below 100 cm^{-1}. In these cases, we simply treat that small, real frequency as an imaginary frequency and drop it during the calculation of the rate constant, as described in the next section. In Table S1, diffusion processes affected by this issue are underlined.

S1.3. Calculation of rate constants

To model the distribution and migration of the S-vacancies, we have first determined the rate constant k within the framework of transition state theory[S14] for each diffusion process. This has been done with the Eyring equation:

$$k = \frac{k_B T}{h}\left(\frac{q_{TS}}{q_{IS}}\right)\exp\left(\frac{-E_a}{k_B T}\right) \quad \text{(Eq. S1)}$$

Here, k_B stands for the Boltzmann constant, h for the Planck constant, and T for the absolute temperature. The partition functions of the TS and IS are abbreviated as q_{TS} and q_{IS}, respectively. E_a is the calculated diffusion barrier without zero-point energy corrections. The total partition function q of any chemical system can be described as the product of the contributions from the vibrational (q_{vib}), rotational (q_{rot}), translational (q_{trans}), and electronic degrees of freedom (q_{elec}):

$$q = q_{vib} q_{rot} q_{trans} q_{elec} \quad \text{(Eq. S2)}$$

Assuming that TS and IS feature sufficiently similar electronic structures, we can assume that $q_{elec,TS}$ and $q_{elec,IS}$ are equal. Similarly, $q_{trans,TS}$ and $q_{trans,IS}$ are also the same since the mass of the considered systems does not change upon diffusion. Lastly, if we further assume that all rotational degrees of freedom become vibrational degrees of freedom on a crystal surface, we can simplify the partition function term in Eq S1 to:

$$\frac{q_{TS}}{q_{IS}} = \frac{q_{vib,TS}}{q_{vib,IS}} \quad \text{(Eq. S3)}$$

with the vibrational contribution to the partition function being

$$q_{vib} = \prod_i^{real} \frac{\exp(-h\nu_i/2k_B T)}{1-\exp(-h\nu_i/k_B T)} \quad \text{(Eq. S4)}$$

In the equation above, ν_i stands for the real vibrational frequencies of the system under consideration.

S1.4. Optimized molecular structures and energies

We should first introduce a shorthand notation to describe the occupation on the nearest-neighbor sites and the corresponding diffusion processes. We will use a simple string with "x" and "o" to represent whether a site is occupied by an S-atom or empty. For example, a structure with sites 1, 4, 5, 6, and 8 featuring S-atoms and sites 2, 3, and 7 having a vacancy, as shown in Figure S2(a), would be abbreviated as "1x2o3o4x5x6x7o8x".

All optimized IS and TS structures reported in this work are collected in an attached zip file. Due to symmetry, each FS is symmetrically equivalent to (at least) one IS structure. Thus, the FS structures have not been explicitly optimized. For example, the IS of the diffusion process with sites 1, 2, and 3 being empty (1o2o3o4x5x6x7x8x), as shown in Figure S2(b), is equivalent to the FS of the diffusion process with sites 6, 7, and 8 being empty (1x2x3x4x5x6o7o8o) as illustrated in Figure S2(c). The same relation also applies to the FS of 1o2o3o4x5x6x7x8x and the IS of 1x2x3x4x5x6o7o8o. In other words, the diffusion processes 1o2o3o4x5x6x7x8x and 1x2x3x4x5x6o7o8o are the reverse processes of each other. As a reaction and its reverse reaction proceed *via* the same TS, one only needs to optimize the TS for the diffusion process 1o2o3o4x5x6x7x8x to obtain all relevant information for the TS of the diffusion 1x2x3x4x5x6o7o8o. All calculated reaction energies and diffusion barriers (in the forward direction) are listed in Table S1.

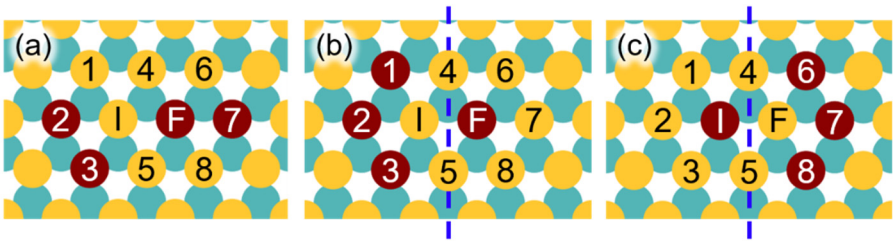

Figure S2. (a) Schematic representation of the (IS of the) diffusion process with sites 2, 3, and 7 being empty, labeled as 1x2o3o4x5x6x7o8x, (b) IS of the diffusion process 1o2o3o4x5x6x7x8x, (c) FS of diffusion process 1x2x3x4x5x6o7o8o. Reflecting the structure in (b) at the blue dashed line yields the structure shown in (c) and *vice versa*.

Table S1. The reaction energies (E_r) and diffusion barriers (E_a) for the 256 considered unique diffusion processes in the forward direction. All values are given in eV. Underlined processes are those processes for which we have not been able to identify the imaginary vibrational frequency for the TS.

Diffusion Process	E_r	E_a	Diffusion Process	E_r	E_a
1x2x3x4x5x6x7x8x	0.00	2.31	1x2x3x4x5o6x7x8o	-0.26	0.81
1o2x3x4x5x6x7x8x	-0.04	2.27	1x2x3x4x5x6o7o8x	-0.17	2.36
1x2o3x4x5x6x7x8x	-0.06	2.32	1x2x3x4x5x6o7x8o	0.06	2.37
1x2x3o4x5x6x7x8x	-0.03	2.29	1x2x3x4x5x6x7o8o	-0.11	2.34
1x2x3x4o5x6x7x8x	0.00	2.02	1o2o3o4x5x6x7x8x	0.38	2.70
1x2x3x4x5o6x7x8x	0.00	0.81	1o2o3x4o5x6x7x8x	0.39	2.21
1x2x3x4x5x6o7x8x	0.04	2.31	1o2o3x4x5o6x7x8x	-0.02	0.89
1x2x3x4x5x6x7o8x	0.06	2.39	1o2o3x4x5x6o7x8x	0.19	2.59
1x2x3x4x5x6x7x8o	0.03	2.33	1o2o3x4x5x6x7o8x	0.24	2.60
1o2o3x4x5x6x7x8x	0.17	2.53	1o2o3x4x5x6x7x8o	0.19	2.52
1o2x3o4x5x6x7x8x	-0.06	2.31	1o2x3o4o5x6x7x8x	0.06	2.04
1o2x3x4o5x6x7x8x	0.20	2.00	1o2x3o4x5o6x7x8x	0.09	0.88
1o2x3x4x5o6x7x8x	-0.16	0.73	1o2x3o4x5x6o7x8x	-0.03	2.36
1o2x3x4x5x6o7x8x	0.00	2.36	1o2x3o4x5x6x7o8x	0.03	2.40
1o2x3x4x5x6x7o8x	0.04	2.34	1o2x3o4x5x6x7x8o	-0.02	2.34
1o2x3x4x5x6x7x8o	0.00	2.27	1o2x3x4o5o6x7x8x	0.04	1.12
1x2o3o4x5x6x7x8x	0.11	2.45	1o2x3x4o5x6o7x8x	0.00	1.95
1x2o3x4o5x6x7x8x	-0.09	2.00	1o2x3x4o5x6x7o8x	0.29	2.10
1x2o3x4x5o6x7x8x	-0.09	0.74	1o2x3x4o5x6x7x8o	0.36	2.10
1x2o3x4x5x6o7x8x	-0.04	2.30	<u>1o2x3x4x5o6o7x8x</u>	0.00	0.72
1x2o3x4x5x6x7o8x	0.00	2.40	1o2x3x4x5o6x7o8x	-0.06	0.71
1x2o3x4x5x6x7x8o	-0.04	2.34	1o2x3x4x5o6x7x8o	-0.39	0.76
1x2x3o4o5x6x7x8x	-0.16	1.99	1o2x3x4x5x6o7o8x	-0.19	2.40
1x2x3o4x5o6x7x8x	0.26	1.07	1o2x3x4x5x6o7x8o	0.03	2.38
1x2x3o4x5x6o7x8x	0.00	2.26	1o2x3x4x5x6x7o8o	-0.10	2.27
1x2x3o4x5x6x7o8x	0.04	2.38	1x2o3o4o5x6x7x8x	-0.06	2.09
1x2x3o4x5x6x7x8o	0.00	2.39	1x2o3o4x5o6x7x8x	0.36	1.11
<u>1x2x3x4o5o6x7x8x</u>	0.00	0.97	1x2o3o4x5x6o7x8x	0.10	2.37
1x2x3x4o5x6o7x8x	-0.20	1.79	1x2o3o4x5x6x7o8x	0.19	2.50
1x2x3x4o5x6x7o8x	0.09	2.09	1x2o3o4x5x6x7x8o	0.12	2.52
1x2x3x4o5x6x7x8o	0.16	2.15	1x2o3x4o5o6x7x8x	-0.11	0.88
1x2x3x4x5o6o7x8x	0.16	0.89	1x2o3x4o5x6o7x8x	-0.29	1.81
1x2x3x4x5o6x7o8x	0.09	0.82	1x2o3x4o5x6x7o8x	0.00	2.08

Diffusion Process	E_r	E_a
1x2o3x4o5x6x7x8o	0.06	2.12
1x2o3x4x5o6o7x8x	0.06	0.77
1x2o3x4x5o6x7o8x	0.00	0.71
1x2o3x4x5o6x7x8o	-0.35	0.76
1x2o3x4x5x6o7o8x	-0.24	2.35
1x2o3x4x5x6o7x8o	-0.03	2.37
1x2o3x4x5x6x7o8o	-0.19	2.31
1x2x3o4o5o6x7x8x	0.11	1.08
1x2x3o4o5x6o7x8x	-0.36	1.75
1x2x3o4o5x6x7o8x	-0.06	2.06
1x2x3o4o5x6x7x8o	0.00	2.17
1x2x3o4x5o6o7x8x	0.39	1.15
1x2x3o4x5o6x7o8x	0.35	1.11
1x2x3o4x5o6x7x8o	0.00	0.97
1x2x3o4x5x6o7o8x	-0.19	2.33
1x2x3o4x5x6o7x8o	0.02	2.36
1x2x3o4x5x6x7o8o	-0.12	2.40
1x2x3x4o5o6o7x8x	-0.04	1.08
1x2x3x4o5o6x7o8x	0.11	0.98
1x2x3x4o5o6x7x8o	-0.11	0.97
1x2x3x4o5x6o7o8x	-0.39	1.82
1x2x3x4o5x6o7x8o	-0.06	1.98
1x2x3x4o5x6x7o8o	0.06	2.15
1x2x3x4x5o6o7o8x	0.02	0.91
1x2x3x4x5o6o7x8o	-0.09	0.78
1x2x3x4x5o6x7o8o	-0.36	0.76
1x2x3x4x5x6o7o8o	-0.38	2.32
1o2o3o4o5x6x7x8x	0.41	2.34
1o2o3o4x5o6x7x8x	0.43	1.16
1o2o3o4x5x6o7x8x	0.36	2.69
1o2o3o4x5x6x7o8x	0.46	2.74
1o2o3o4x5x6x7x8o	0.38	2.75
1o2o3x4o5o6x7x8x	0.18	1.16
1o2o3x4o5x6o7x8x	0.18	2.18
1o2o3x4o5x6x7o8x	0.48	2.32
1o2o3x4o5x6x7x8o	0.52	2.32
1o2o3x4x5o6o7x8x	0.13	0.86

Diffusion Process	E_r	E_a
1o2o3x4x5o6x7o8x	0.07	0.87
1o2o3x4x5o6x7x8o	-0.26	0.93
1o2o3x4x5x6o7o8x	0.00	2.56
1o2o3x4x5x6o7x8o	0.19	2.61
1o2o3x4x5x6x7o8o	0.08	2.50
1o2x3o4o5o6x7x8x	0.12	0.99
1o2x3o4o5x6o7x8x	-0.14	1.98
1o2x3o4o5x6x7o8x	0.15	2.15
1o2x3o4o5x6x7x8o	0.20	2.16
1o2x3o4x5o6o7x8x	0.23	0.96
1o2x3o4x5o6x7o8x	0.20	0.89
1o2x3o4x5o6x7x8o	-0.14	0.93
1o2x3o4x5x6o7o8x	-0.19	2.42
1o2x3o4x5x6o7x8o	0.00	2.49
1o2x3o4x5x6x7o8o	-0.10	2.37
1o2x3x4o5o6o7x8x	0.00	0.92
1o2x3x4o5o6x7o8x	0.15	1.00
1o2x3x4o5o6x7x8o	-0.06	1.00
1o2x3x4o5x6o7o8x	-0.18	2.00
1o2x3x4o5x6o7x8o	0.14	2.12
1o2x3x4o5x6x7o8o	0.28	2.16
1o2x3x4x5o6o7o8x	-0.13	0.73
1o2x3x4x5o6o7x8o	-0.23	0.72
1o2x3x4x5o6x7o8o	-0.47	0.72
1o2x3x4x5x6o7o8o	-0.36	2.33
1x2o3o4o5o6x7x8x	0.19	1.10
1x2o3o4o5x6o7x8x	-0.28	1.88
1x2o3o4o5x6x7o8x	0.04	2.11
1x2o3o4o5x6x7x8o	0.07	2.15
1x2o3o4x5o6o7x8x	0.47	1.19
1x2o3o4x5o6x7o8x	0.45	1.17
1x2o3o4x5o6x7x8o	0.09	1.00
1x2o3o4x5x6o7o8x	-0.08	2.42
1x2o3o4x5x6o7x8o	0.10	2.47
1x2o3o4x5x6x7o8o	0.00	2.46
1x2o3x4o5o6o7x8x	-0.15	0.85
1x2o3x4o5o6x7o8x	0.00	0.87

Diffusion Process	E_r	E_a	Diffusion Process	E_r	E_a
1x2o3x4o5o6x7x8o	-0.23	0.93	1o2o3x4o5x6o7x8o	0.31	2.34
1x2o3x4o5x6o7o8x	-0.48	1.84	1o2o3x4o5x6x7o8o	0.46	2.36
1x2o3x4o5x6o7x8o	-0.15	2.00	1o2o3x4x5o6o7o8x	0.00	0.85
1x2o3x4o5x6x7o8o	-0.04	2.07	1o2o3x4x5o6o7x8o	-0.12	0.88
1x2o3x4x5o6o7o8x	-0.07	0.79	1o2o3x4x5o6x7o8o	-0.34	0.88
1x2o3x4x5o6o7x8o	-0.20	0.69	1o2o3x4x5x6o7o8o	-0.19	2.53
1x2o3x4x5o6x7o8o	-0.45	0.71	1o2x3o4o5o6o7x8x	0.08	0.94
1x2o3x4x5x6o7o8o	-0.46	2.28	1o2x3o4o5o6x7o8x	0.24	1.07
1x2x3o4o5o6o7x8x	0.06	1.06	1o2x3o4o5o6x7x8o	0.02	1.02
1x2x3o4o5o6x7o8x	0.23	1.15	1o2x3o4o5x6o7o8x	-0.31	2.04
1x2x3o4o5o6x7x8o	0.00	0.93	1o2x3o4o5x6o7x8o	0.00	2.11
1x2x3o4o5x6o7o8x	-0.52	1.79	1o2x3o4o5x6x7o8o	0.15	2.16
1x2x3o4o5x6o7x8o	-0.20	1.96	1o2x3o4x5o6o7o8x	0.12	1.00
1x2x3o4o5x6x7o8o	-0.07	2.08	1o2x3o4x5o6o7x8o	0.00	0.85
1x2x3o4x5o6o7o8x	0.26	1.19	1o2x3o4x5o6x7o8o	-0.20	0.89
1x2x3o4x5o6o7x8o	0.14	1.07	1o2x3o4x5x6o7o8o	-0.35	2.41
1x2x3o4x5o6x7o8o	-0.09	0.91	1o2x3x4o5o6o7o8x	-0.14	0.92
1x2x3o4x5x6o7o8o	-0.38	2.37	1o2x3x4o5o6o7x8o	-0.08	0.87
1x2x3x4o5o6o7o8x	-0.18	0.97	1o2x3x4o5o6x7o8o	-0.12	0.98
1x2x3x4o5o6o7x8o	-0.12	0.86	1o2x3x4o5x6o7o8o	-0.20	2.07
1x2x3x4o5o6x7o8o	-0.19	0.91	1o2x3x4x5o6o7o8o	-0.24	0.98
1x2x3x4o5x6o7o8o	-0.41	1.93	1x2o3o4o5o6o7x8x	0.12	1.10
1x2x3x4x5o6o7o8o	-0.43	0.73	1x2o3o4o5o6x7o8x	0.30	1.16
1o2o3o4o5o6x7x8x	0.43	1.25	1x2o3o4o5o6x7x8o	0.06	1.01
1o2o3o4o5x6o7x8x	0.20	2.27	1x2o3o4o5x6o7o8x	-0.46	1.90
1o2o3o4o5x6x7o8x	0.51	2.44	1x2o3o4o5x6o7x8o	-0.15	2.01
1o2o3o4o5x6x7x8o	0.53	2.49	1x2o3o4o5x6x7o8o	0.00	2.06
1o2o3o4x5o6o7x8x	0.24	1.22	1x2o3o4x5o6o7o8x	0.34	1.22
1o2o3o4x5o6x7o8x	0.53	1.18	1x2o3o4x5o6o7x8o	0.20	1.09
1o2o3o4x5o6x7x8o	0.18	1.16	1x2o3o4x5o6x7o8o	0.00	0.93
1o2o3o4x5x6o7o8x	0.19	2.72	1x2o3o4x5x6o7o8o	-0.30	2.38
1o2o3o4x5x6o7x8o	0.35	2.76	1x2o3x4o5o6o7o8x	-0.28	0.87
1o2o3o4x5x6x7o8o	0.30	2.68	1x2o3x4o5o6o7x8o	-0.24	0.83
1o2o3x4o5o6o7x8x	0.14	1.06	1x2o3x4o5o6x7o8o	-0.30	0.86
1o2o3x4o5o6x7o8x	0.28	1.15	1x2o3x4o5x6o7o8o	-0.51	1.93
1o2o3x4o5o6x7x8o	0.08	1.16	1x2o3x4x5o6o7o8o	-0.53	0.65
1o2o3x4o5x6o7o8x	0.00	2.23	1x2x3o4o5o6o7o8x	-0.08	1.08

Diffusion Process	E_r	E_a
1x2x3o4o5o6o7x8o	-0.02	0.99
1x2x3o4o5o6x7o8o	-0.06	0.95
1x2x3o4o5x6o7o8o	-0.53	1.95
1x2x3o4x5o6o7o8o	-0.18	0.98
1x2x3x4o5o6o7o8o	-0.43	0.82
1o2o3o4o5o6o7x8x	0.37	1.21
1o2o3o4o5o6x7o8x	0.54	1.32
1o2o3o4o5o6x7x8o	0.31	1.22
1o2o3o4o5x6o7o8x	0.03	2.34
1o2o3o4o5x6o7x8o	0.31	2.37
1o2o3o4o5x6x7o8o	0.48	2.38
1o2o3o4x5o6o7o8x	0.42	1.26
1o2o3o4x5o6o7x8o	0.28	1.15
1o2o3o4x5o6x7o8o	0.12	1.16
1o2o3o4x5x6o7o8o	0.00	2.67
1o2o3x4o5o6o7o8x	0.00	1.06
1o2o3x4o5o6o7x8o	0.05	1.05
1o2o3x4o5o6x7o8o	0.02	1.13
1o2o3x4o5x6o7o8o	-0.03	2.31
1o2o3x4x5o6o7o8o	-0.42	0.84
1o2x3o4o5o6o7o8x	-0.05	1.00
1o2x3o4o5o6o7x8o	0.00	0.92
1o2x3o4o5o6x7o8o	-0.02	0.98
1o2x3o4o5x6o7o8o	-0.31	2.06
1o2x3o4x5o6o7o8o	-0.28	0.86
1o2x3x4o5o6o7o8o	-0.37	0.84
1x2o3o4o5o6o7o8x	-0.02	1.12
1x2o3o4o5o6o7x8o	0.02	1.00
1x2o3o4o5o6x7o8o	0.00	0.94
1x2o3o4o5x6o7o8o	-0.48	1.89
1x2o3o4x5o6o7o8o	-0.12	1.04
1x2o3x4o5o6o7o8o	-0.54	0.78
1x2x3o4o5o6o7o8o	-0.31	0.91
1o2o3o4o5o6o7o8x	0.24	1.27
1o2o3o4o5o6o7x8o	0.27	1.15
1o2o3o4o5o6x7o8o	0.27	1.17
1o2o3o4o5x6o7o8o	0.00	2.28

Diffusion Process	E_r	E_a
1o2o3o4x5o6o7o8o	0.00	1.05
1o2o3x4o5o6o7o8o	-0.24	1.03
1o2x3o4o5o6o7o8o	-0.27	0.87
1x2o3o4o5o6o7o8o	-0.27	0.90
1o2o3o4o5o6o7o8o	0.00	1.07

S2. Further results and illustrations

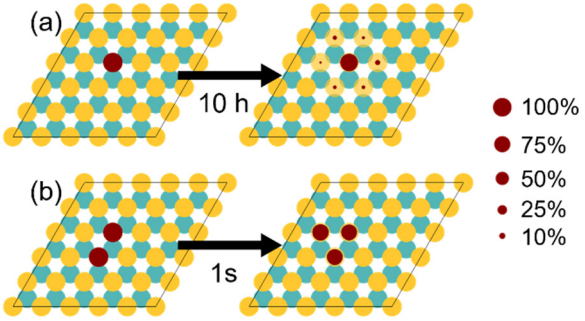

Figure S3. The surface models on the right show the probability of finding an S-vacancy on each site after (a) an isolated S-vacancy and (b) a 2S-vacancy on MoS_2, as shown on the left side, has been exposed to a temperature of 600 K for the indicated time. The probabilities of finding a vacancy on each site are calculated based on the results of 60 simulation runs.

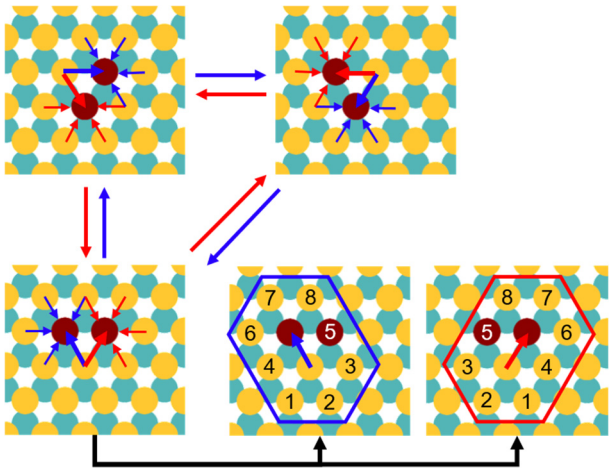

Figure S4. The diffusion processes of 2S-vacancies. The arrows in red/blue in the MoS_2 models mark all possible diffusion processes, with the arrow's direction indicating the S-atom's movement. The wider arrows are the dominating, low barrier diffusion processes. As exemplarily shown for the case at the bottom, these diffusion processes correspond to a situation where the not moving vacancy is on site 5, which leads to a low diffusion barrier. The diffusion processes marked by the broad arrows yield the surface structures as indicated by the reaction arrows of the same color. It can be seen that the 2S-vacancy only moves within a confined space consisting of three neighboring sites.

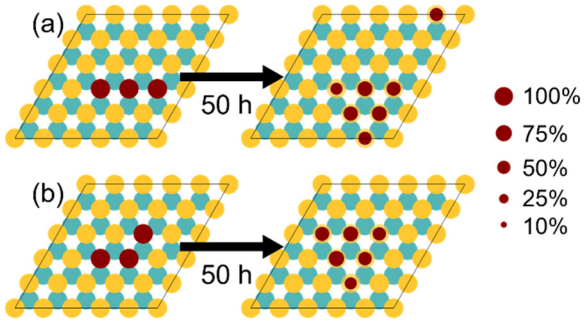

Figure S5. The surface models on the right show the probability of finding an S-vacancy on each site after the 3S-vacancies on the left have been exposed to 300 K for 50 h. The probabilities of finding a vacancy on each site are calculated based on the results of 60 simulation runs. It could be shown that both initial structures lead to the same distribution of vacancy sites after 50 h at 300 K.

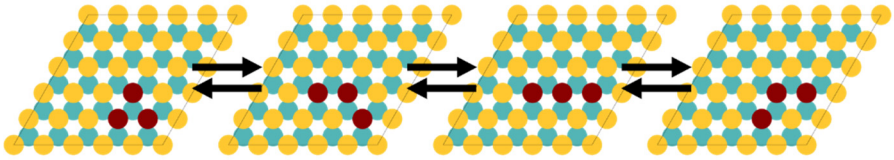

Figure S6. The conversion of the different types of 3S-vacancies into each other that play a role in the equilibrium between hcp-type (left), v-shaped (second from left and right), and linear 3S-vacancies (second from right). The hcp-type 3S-vacancies can rapidly convert into v-shaped 3S-vacancies, which in return can quickly change into a linear 3S-vacancy, and *vice versa*. Note that the fcc-type 3S-vacancy is not a part of this equilibrium at room temperature as any reaction from or to it is hindered by a high diffusion barrier.

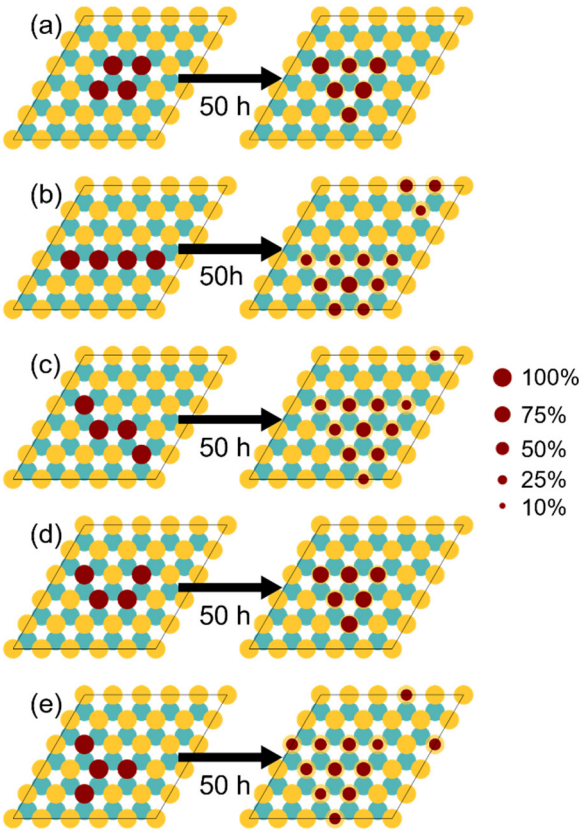

Figure S7. The surface models on the right show the probability of finding an S-vacancy on each site after the various initial arrangements of 4S-vacancies on the left side have been exposed to a temperature of 300 K for 50 h. The probabilities of finding a vacancy on each site are calculated based on the results of 60 simulation runs.

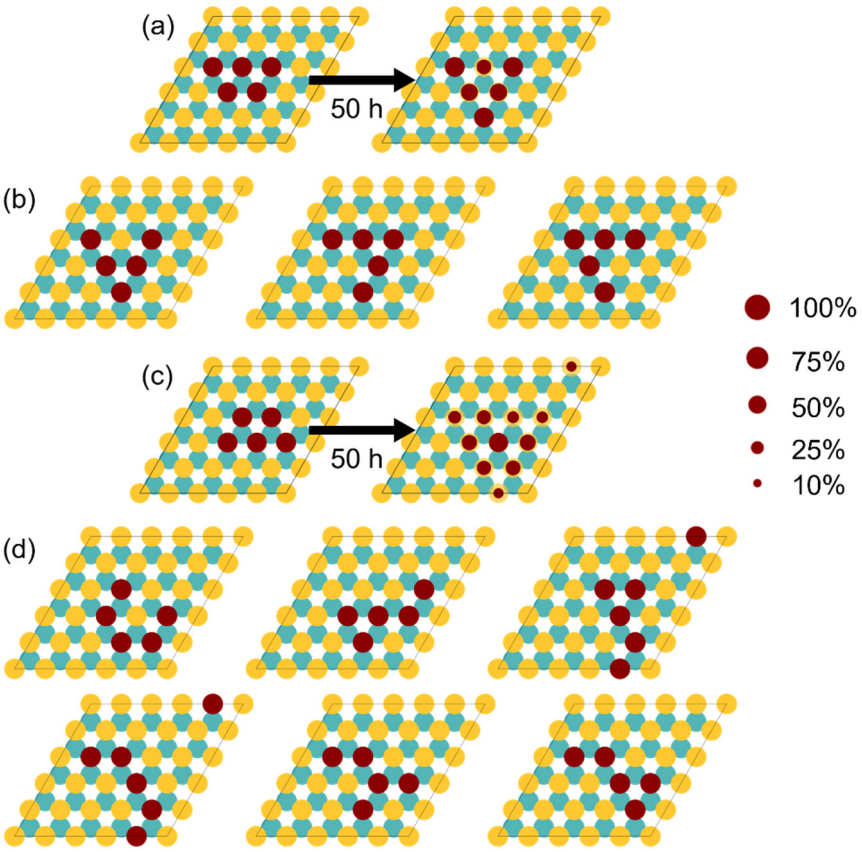

Figure S8. The left side of (a) shows a 5S-vacancy that can be understood as two fcc-type and one hcp-type 3S-vacancies. The right side illustrates the probability of finding an S-vacancy on each site after exposing the structure on the left to 300 K for 50 hours. (b) shows the different arrangements of the 5S-vacancy after 50 hours, which lead to the probability distribution shown in (a). (c) and (d) show the corresponding information for a 5S-vacancy that can be understood as one fcc-type and two hcp-type 3S-vacancies. The probabilities of finding a vacancy on each site are calculated based on the results of 60 simulation runs.

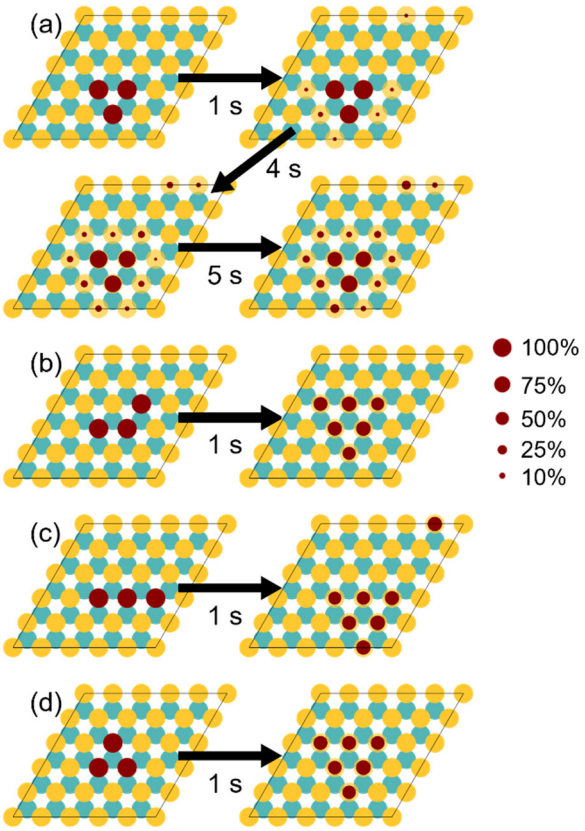

Figure S9. The surface models on the right show the probability of finding an S-vacancy on each site after the various initial arrangements of 3S-vacancies on MoS_2 on the left side have been exposed to a temperature of 600 K for 1 s. The probabilities of finding a vacancy on each site are calculated based on the results of 60 simulation runs. The probability distributions for (b)-(d) are similar, while the result for fcc-type 3S-vacancy shown in (a) differs visibly. Thus, the simulation time for (a) has been extended to 10 s to see whether it will resemble the other systems after a longer equilibration time. Note that the values over the reaction arrows denote the time interval between the shown probability distributions.

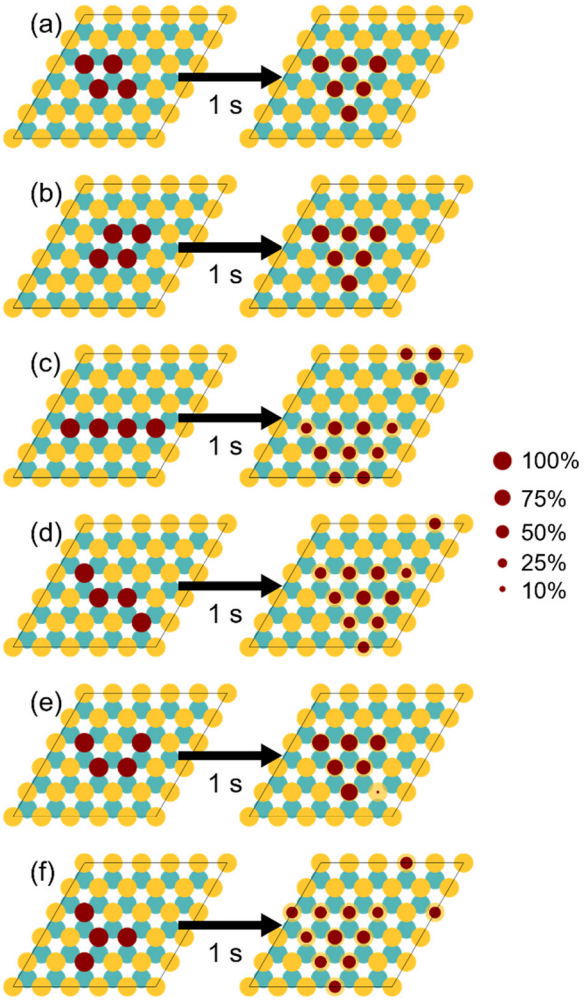

Figure S10. The surface models on the right show the probability of finding an S-vacancy on each site after the various initial arrangements of 4S-vacancies on the left side have been exposed to a temperature of 600 K for 1 s. The probabilities of finding a vacancy on each site are calculated based on the results of 60 simulation runs.

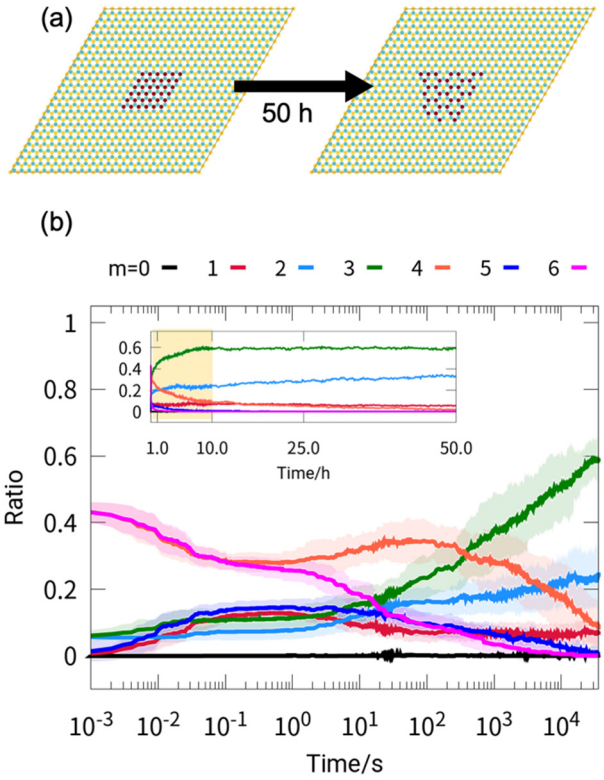

Figure S11. (a) MoS_2 surface with a rhombic nS-vacancy (6% vacancy ratio) and the simulated structure after exposing it to a temperature of 300 K for 50 hours. (b) shows the ratio of S-vacancies with m neighboring S-vacancy sites as a function of time, averaged over 32 simulations under identical conditions. The backgrounds of the graphs in a lighter color mark the standard deviation.

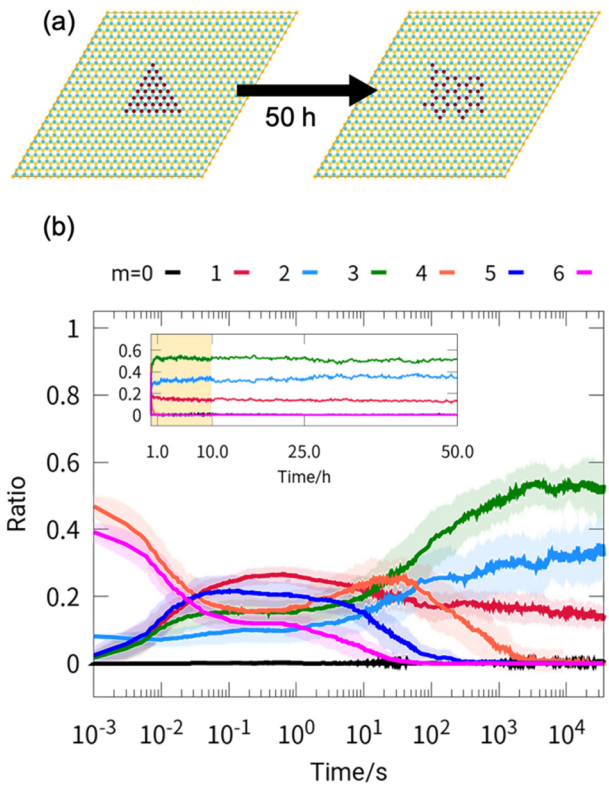

Figure S12. (a) MoS_2 surface with a triangular nS-vacancy (6% vacancy ratio) and the simulated structure after exposing it to a temperature of 300 K for 50 hours. (b) shows the ratio of S-vacancies with m neighboring S-vacancy sites as a function of time, averaged over 32 simulations under identical conditions. The backgrounds of the graphs in a lighter color mark the standard deviation.

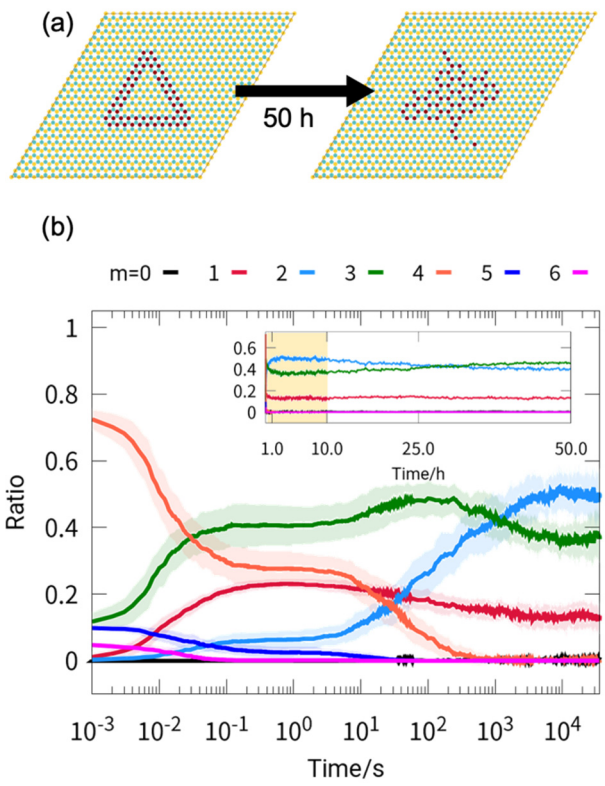

Figure S13. (a) MoS_2 surface with a "hollow" triangular nS-vacancy (10% vacancy ratio) and the simulated structure after exposing it to a temperature of 300 K for 50 hours. (b) shows the ratio of S-vacancies with m neighboring S-vacancy sites as a function of time, averaged over 32 simulations under identical conditions. The backgrounds of the graphs in a lighter color mark the standard deviation.

References

(S1) Kresse, G.; Hafner, J. Ab. Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47*, 558–561.

(S2) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metalamorphous- Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49*, 14251–14269.

(S3) Kresse, G.; Furthmüller, J. 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) Kresse, G.; Furthmü, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169–11186.

(S5) Kresse, G.; Hafner, J. Norm-Conserving and Ultrasoft Pseudopotentials for First-Row and Transition Elements. *J. Phys. Condens. Matter* **1994**, *6*, 8245–8257.

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

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

(S8) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953–17979.

(S9) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59*, 1758–1775.

(S10) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *J. Chem. Phys.* **2000**, *113*, 9901–9904.

(S11) Henkelman, G.; Jónsson, H. Improved Tangent Estimate in the Nudged Elastic Band Method for Finding Minimum Energy Paths and Saddle Points. *J. Chem. Phys.* **2000**, *113*, 9978–9985.

(S12) Henkelman, G.; Jónsson, H. A Dimer Method for Finding Saddle Points on High Dimensional Potential Surfaces Using Only First Derivatives. *J. Chem. Phys.* **1999**, *111*, 7010–7022.

(S13) Heyden, A.; Bell, A. T.; Keil, F. J. 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.

(S14) Eyring, H. The Activated Complex in Chemical Reactions. *J. Chem. Phys.* **1935**, *3*, 107–115.