PNAS

Supporting Information for

The role of dynamics in heterogeneous catalysis: surface diffusivity and N_2 decomposition on Fe(111)

L. Bonati, D. Polino, C. Pizzolitto, P. Biasi, R. Eckert, S. Reitmeier, R. Schlögl and M. Parrinello

E-mail: luigi.bonati@iit.it, michele.parrinello@iit.it

This PDF file includes:

Figs. S1 to S18
Table S1
Legends for Movies S1 to S3

Other supporting materials for this manuscript include the following:

Movies S1 to S3

Contents

1. Validation of the machine learning potential

In Table S1, we report the composition of the dataset used for the ML potential fitting. We first checked the entire distribution of errors on the various systems (Fig. S1). Furthermore, to assess the reliability of the potential in generating new configurations, we compared the radial distribution function of Fe predicted by the MLP with that obtained from the DFT simulations (Fig. S2). Then, we monitored the uncertainty of model predictions by calculating the standard deviation of an ensemble of MLPs on the configurations generated at different temperatures (Fig. S3).

Type	Formula	Training	Validation
AIMD	Fe45	4218	744
AIMD	Fe72	306	53
AIMD	Fe108	1465	258
AIMD	Fe45 N	1782	314
AIMD	Fe45 N2	2012	354
AIMD	Fe45 N4	1746	307
AIMD	Fe72 N	138	24
AIMD	Fe72 N2 H12	3496	616
AIMD	Fe72 N4 H20	1598	282
AIMD	Fe72 N6 H30	716	126
Active learning	Fe45	1128	198
Active learning	Fe72	336	59
Active learning	Fe108	1151	202
Active learning	Fe45 N	265	46
Active learning	Fe45 N2	5202	917
Active learning	Fe45 N4	272	48
Active learning	Fe72 N2	804	141
TOTAL		26635	4689

Table S1. Dataset for MLP construction. Composition of the data set used for the fitting of the machine learning potential. A wide variety of systems were simulated: pure iron, with one or more nitrogen molecules and with nitrogen atoms, and - though not used in this study - also configurations with hydrogen. In all cases 5, 8, or 12 layers were simulated. Configurations were extracted every 2 or 5 fs from AIMD depending on whether they were metadynamics or standard MD simulations. As for the active learning part candidates were extracted every 1 ps and then selected based on the uncertainty on the predictions of a model ensemble.

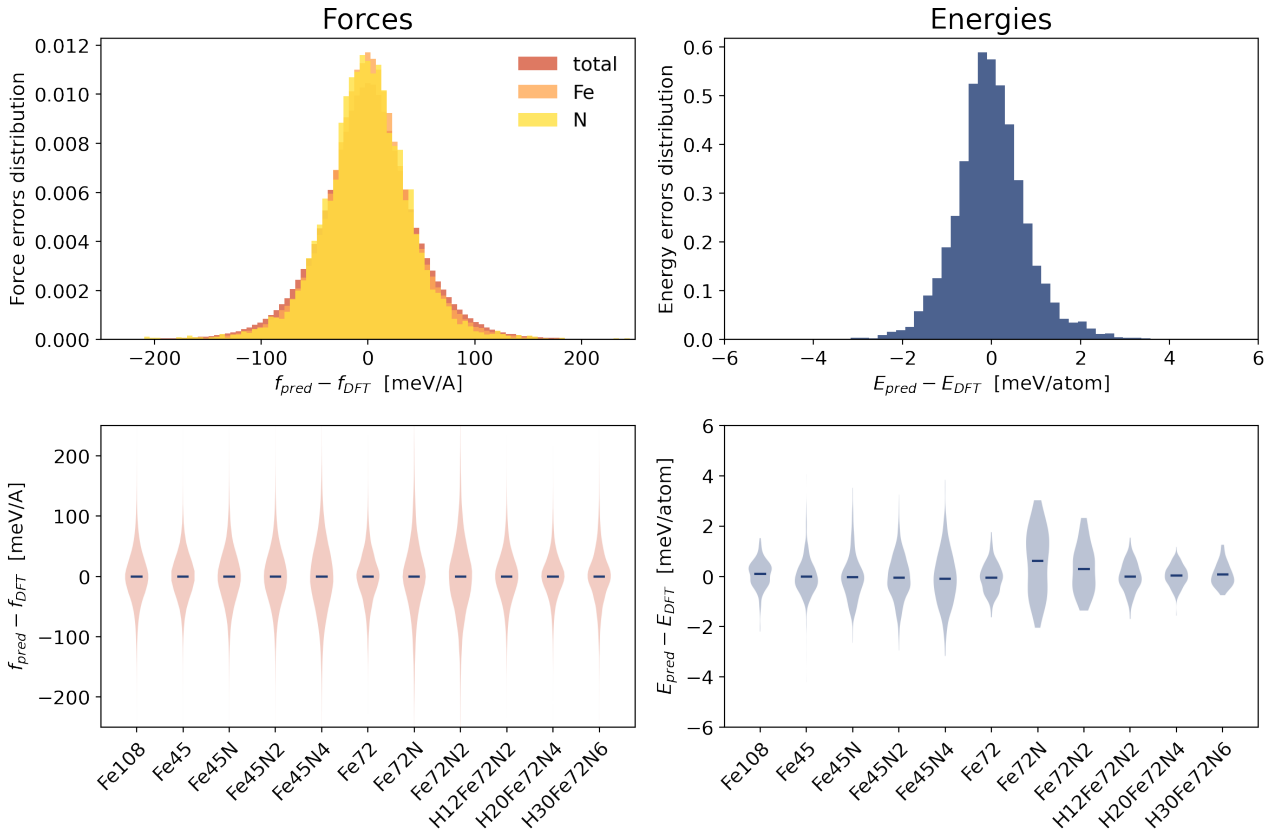

Fig. S1. Accuracy of energy and force predictions. The upper figures report the histogram of the errors in the forces (left) and energies (right) in the test set (corresponding to 5% of the final dataset used for training and validation of the ensemble of NN potentials). For the forces, the total distribution was reported along with the specific distributions of iron and nitrogen atoms, which are remarkably close. In the lower figures, we decomposed the error distributions for each system. The violin plot represents the error distributions on the forces (left) and energies (right) of the configurations grouped by chemical formula. The lines represent the mean value for each system.

Although the system has been optimized with data of the Fe(111) surface, we tested its ability to describe also other surfaces, namely the (110) and (211) surfaces. The RMSE for energies are 2 and 3.4 meV/atom and 56 and 78 meV/Åfor forces for the two surfaces respectively. These values are only slightly higher than those for (111), indicating a good description of the different surfaces.

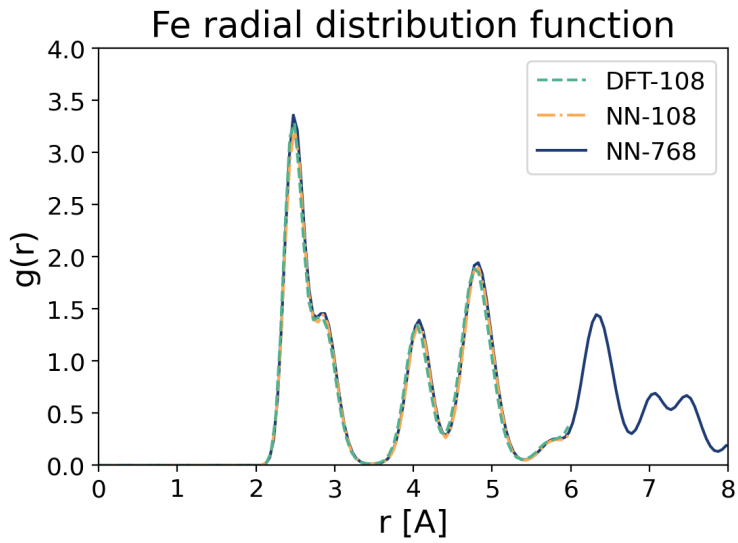

Fig. S2. Fe radial distribution function. Comparison of the $g(r)$ computed from DFT and from the machine learning potential. Two different simulations are performed, one with the same number of atoms as the DFT simulation (108 atoms) and one with a larger system (768 atoms). To make a fair comparison with the AIMD data, simulations of 10 ps were performed also with the MLP.

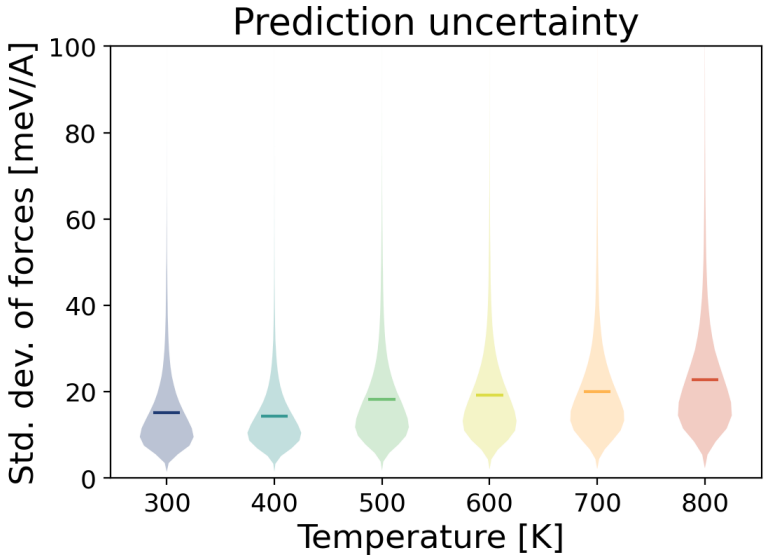

Fig. S3. Uncertainty of force predictions over MD simulations. Violin plot representing the distribution of the standard deviation of the force predictions for the configurations generated by the MLP in representative MD trajectories. The standard deviation is calculated from 4 MLPs trained on different permutations of training+validation dataset. This is then used as proxy for the uncertainty of the predictions. The lines denote the average values at each temperature.

2. Temperature dependence of Fe surface morphology

In Fig. S4 we further validated the quality of surface dynamics results, performing single-point calculations to make sure that the our potential matches the underlying DFT PES. Then we further characterize the mechanism of self-diffusion of Fe atoms. The radial distribution function (Fig. S5) shows that the BCC order is preserved also in the diffusive surface. The scatter plot of atomic positions (Fig. S6) shows the movement of surface atoms as a function of temperature across the whole temperature range, while the displacement distributions (Fig. S7) indicate that the diffusion happens via jumps between lattice sites. Finally, the distribution of the environment similarity (Fig. S8) is plotted for all the temperatures studied.

The following movies are also available:

Movie S1. Surface dynamics at T=300 K vs T=700 K.

Movie S2. Surface atoms diffusion at T=700 K.

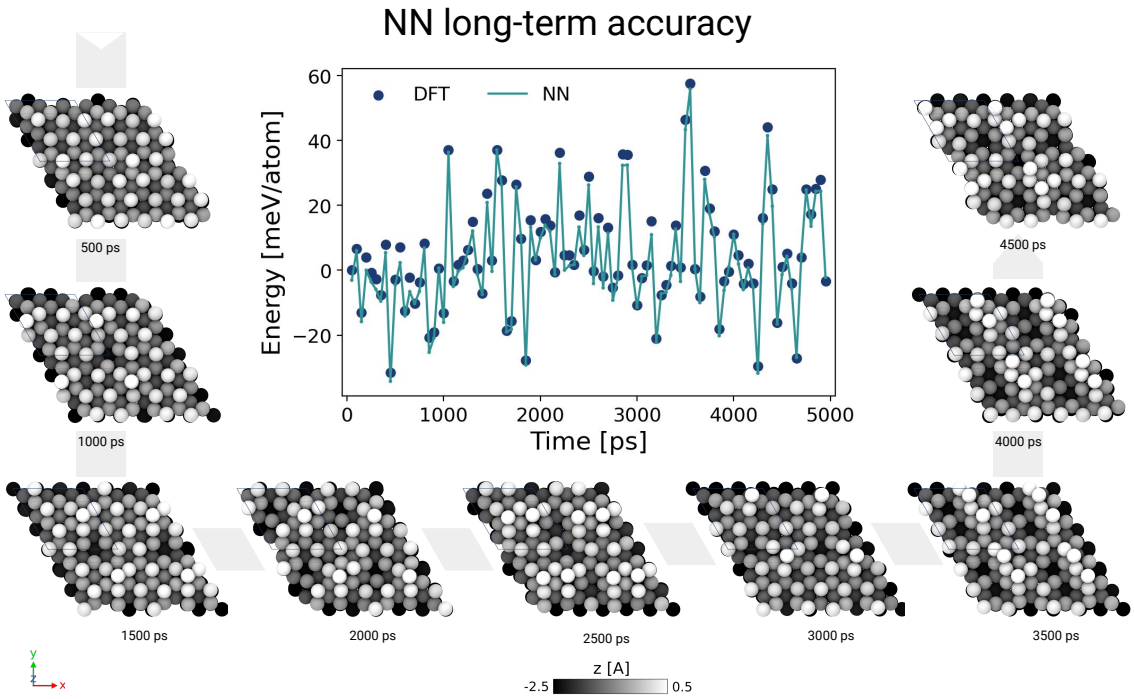

Fig. S4. Accuracy of the potential for surface dynamics. The accuracy over a long-term simulation of the surface is assessed by extracting configurations every 50 ps and comparing the NN predictions with single-point DFT calculations. Along all the trajectory, the NN potential describes well the energetics of the system regardless of the dynamical movement that surface atoms undergo, as illustrated in the snapshots around the figure. The temperature is set to $T = 700$ K, and the system simulated has 108 atoms and 12 layers, of which the bottom two are fixed.

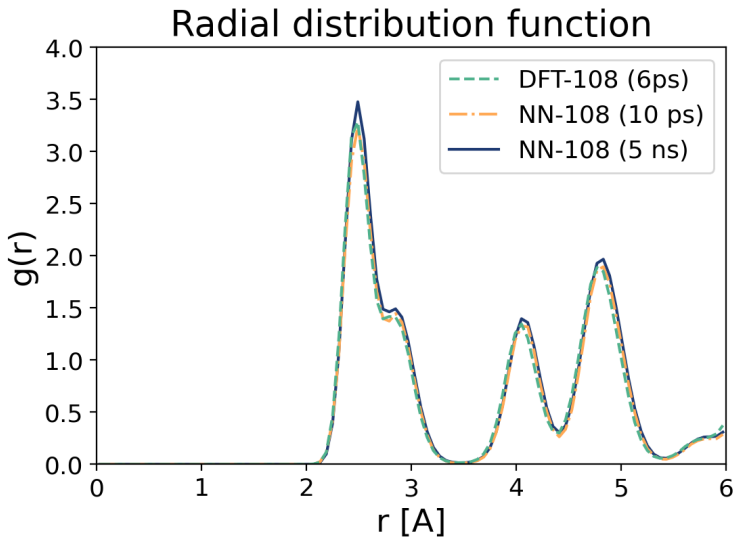

Fig. S5. Radial distribution function for the ideal and diffusive surface. We compare the radial distribution function at T=700K for the ideal surface, obtained from short simulations (DFT-108 6 ps and NN-108 10 ps) already plotted in Fig. S2, with the one extracted from the 5 ns simulation of Fig. S4. The high similarity between them, together with the fact that the second and third neighbor shells are well distinguished, at variance with a liquid-like $g(r)$, clearly show that the short-range order is preserved.

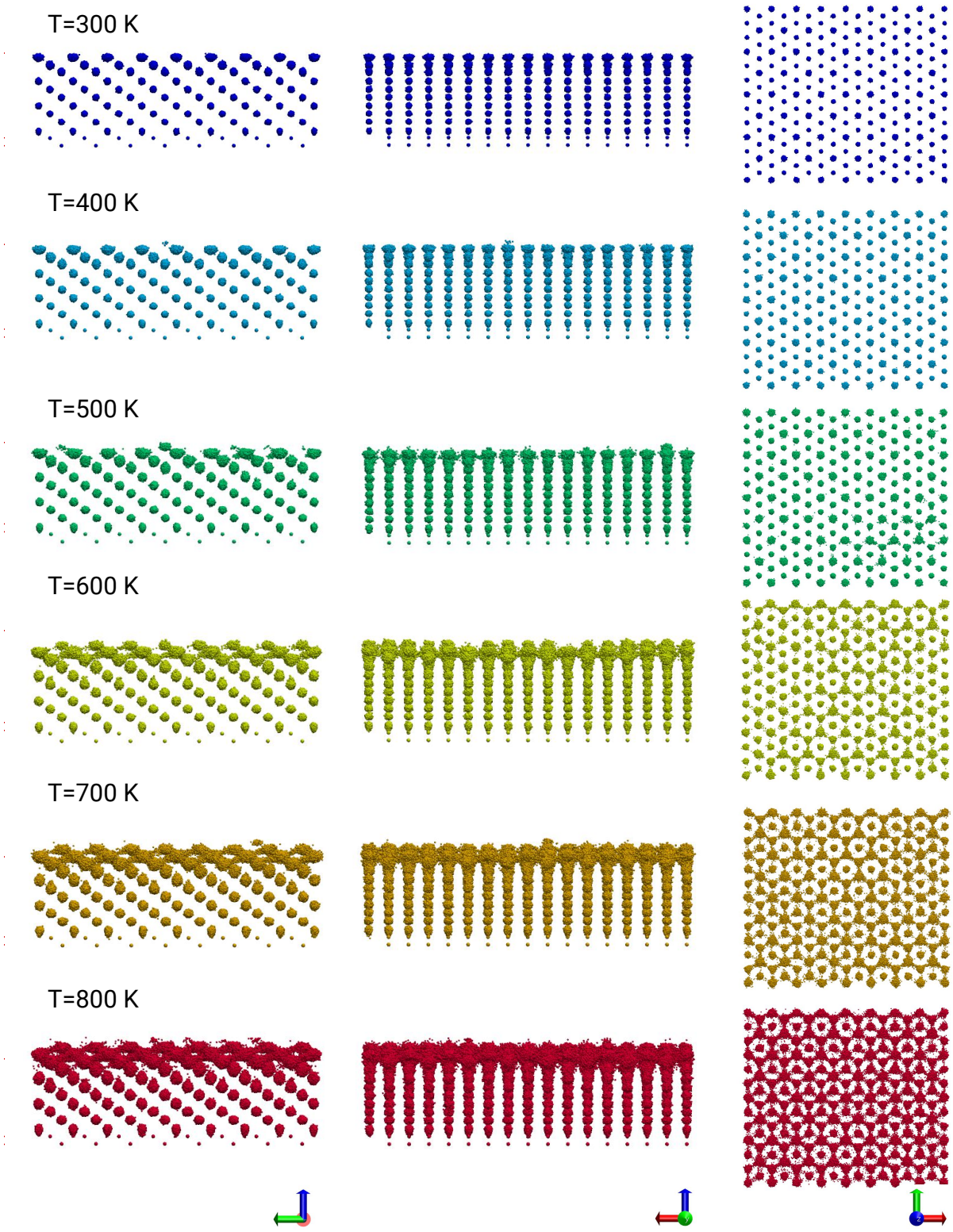

Fig. S6. Scatter plot of atomic position as a function of temperature. Each row represents a temperature, while each column reports the view from different directions according to the bottom of the figure. Same as Fig. 4 in the main text but for the whole temperature range.

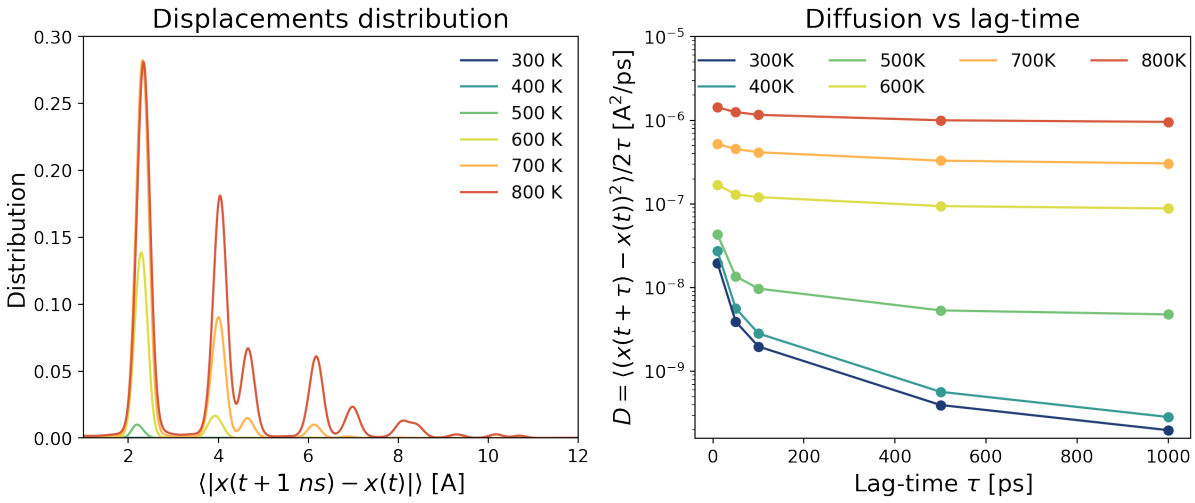

Fig. S7. Surface diffusion additional analysis. (left) Distribution of the time-lagged displacements of surface atoms with a lag time $\tau = 1$ ns. The peaks describe a jump diffusion between lattice sites rather than a liquid-like one. (right) Convergence of the diffusion coefficient computed from the average time-lagged displacements, as a function of the lag-time τ.

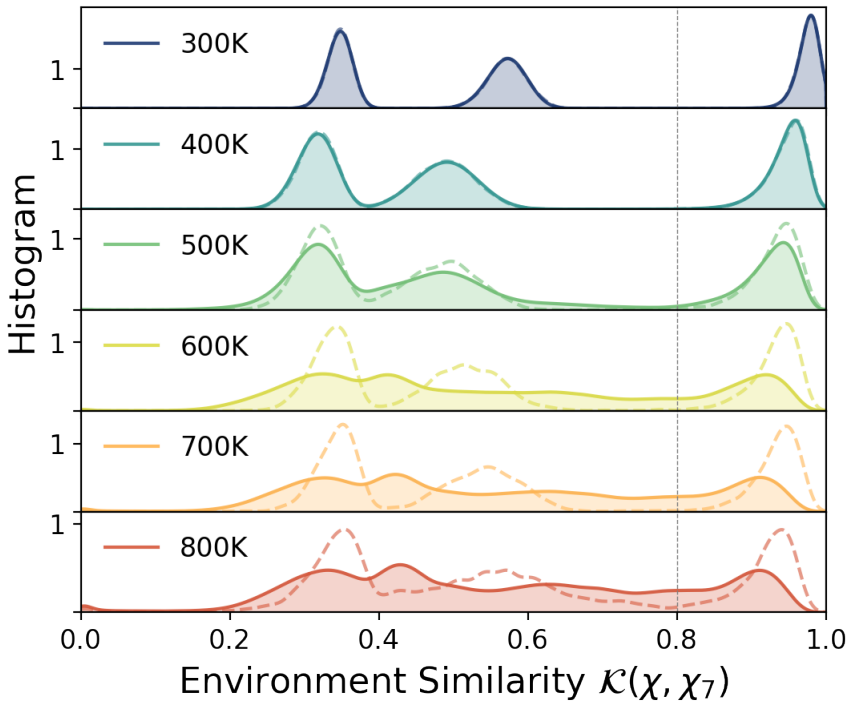

Fig. S8. Environment Similarity distribution. The histogram is calculated for the atoms in the surface for the equilibrated system (filled curves) and for the ideal surface accounting only for thermal fluctuations (calculated from the first 10 ps of the trajectory, checking that no diffusion of the surface atoms is observed). The threshold to decide whether or not a site is χ_7 is chosen to be equal to 0.8 (dotted vertical line), based on the minima of the ideal dashed curves.

3. Nitrogen adsorption and dissociation mechanism

Fig. S9 shows the free energy surfaces as a function of the charge transferred q and the distance of the N_2 molecule from the surface, to highlight the adsorption mechanism between the different sites. In Fig. S10 we report the free energy of dissociative chemisorption as a function of the N-N distance d and the charge q across the whole temperature range. Fig. S11 reports the same FES for two representative temperatures but where the surface is fixed, to highlight the effect of surface dynamics on the destabilization of the α' adsorption state. Fig.S12 shows an analysis of the collective variables used in the OPES simulations. Fig. S13 reports the free energy profiles computed along the distance and the charge transfer with a block average analysis. S14-S15 show the minimum free energy pathways in the $d-q$ space and the 1D projection in units of k_BT.

The following movie is also available: **Movie S3.** N_2 **decomposition at T=300 K and T=700 K.**

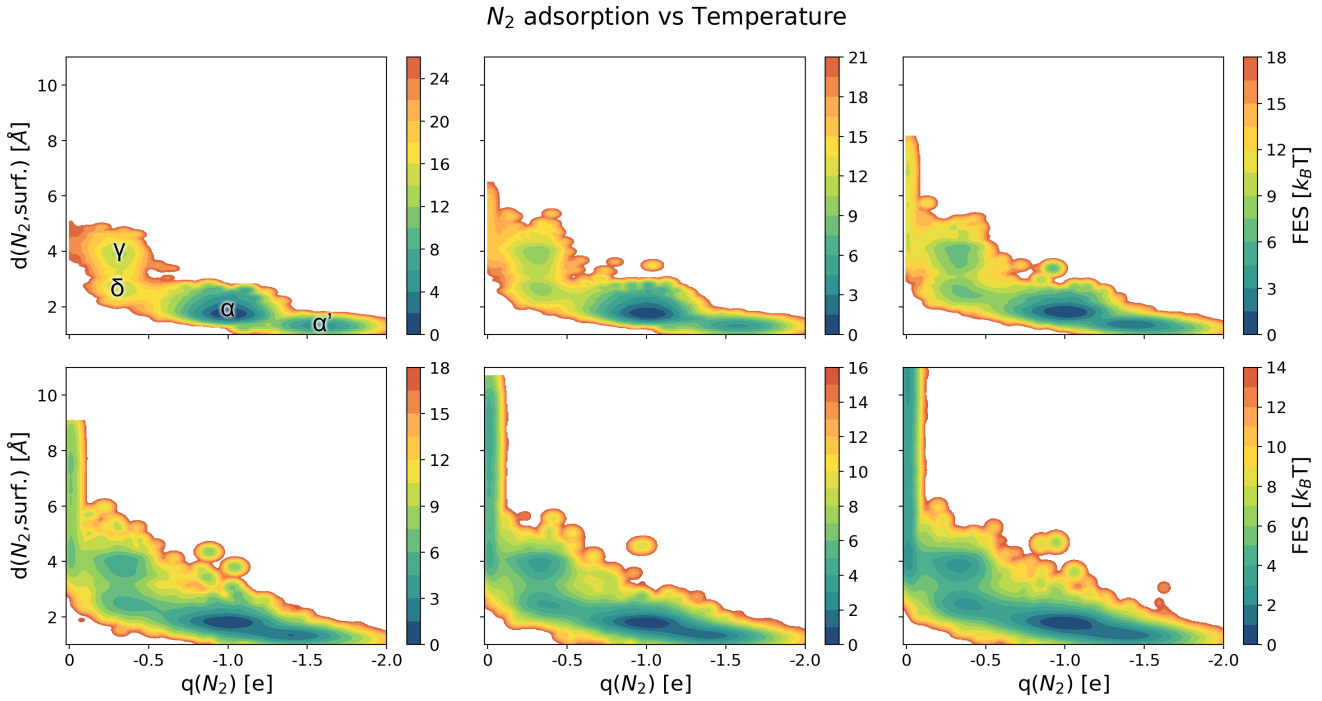

Fig. S9. Adsorption mechanism vs temperature. Free energy profile of N_2 adsorption as a function of N_2 partial charge and the distance between N_2 center of mass and surface. To remove thermal fluctuations and to account for the roughening of the surface a two-step procedure is used to compute the distance between the surface and the moleucule. First, the trajectory is smoothed with a running average of the atomic positions of the Fe atoms. Then, a surface mesh is constructed with the Alpha-Shape method, and the distance from the center of mass of N_2 to the surface mesh. The two states around q=0.3 with different heights from the surface correspond to the N_2 molecule being adsorbed vertically on top of the first or the second layer (γ and δ sites).

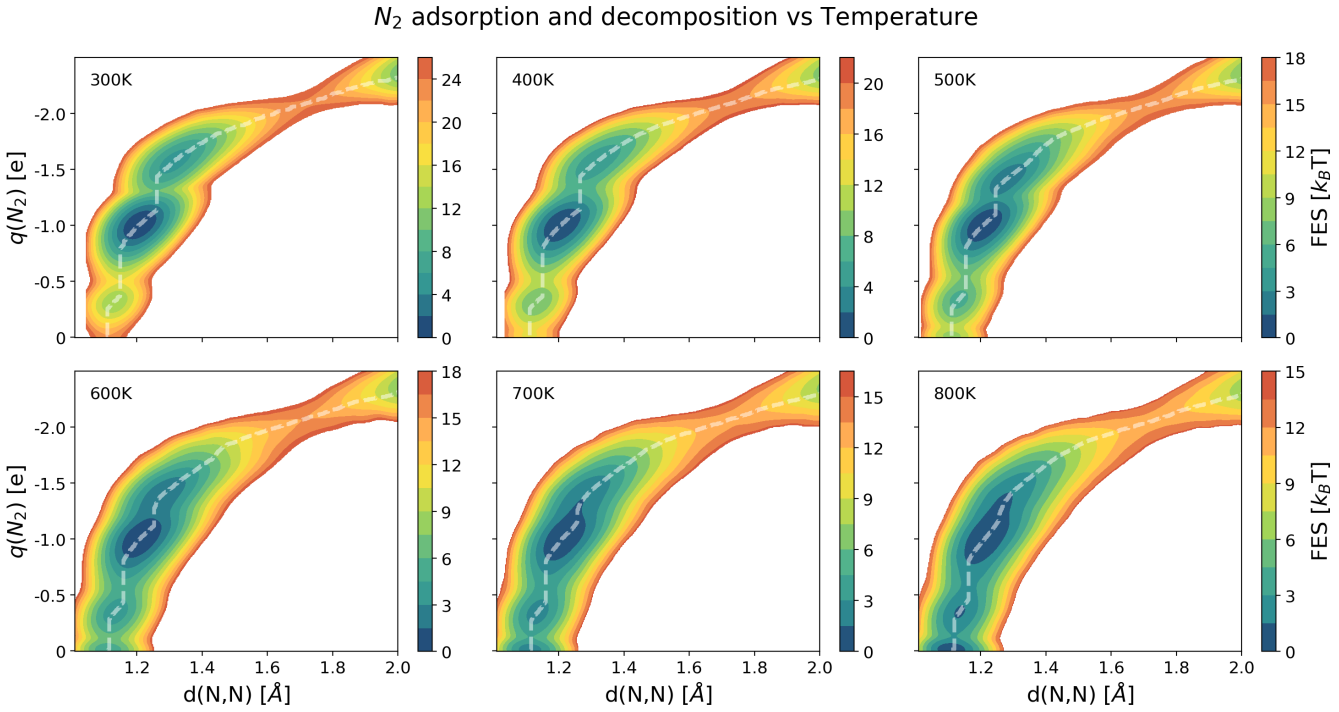

Fig. S10. N_2 **decomposition mechanism.** The free energy surface is computed as a function of N-N distance and N_2 charge for all the temperatures investigated. White dashed lines denote the minimum free energy pathway.

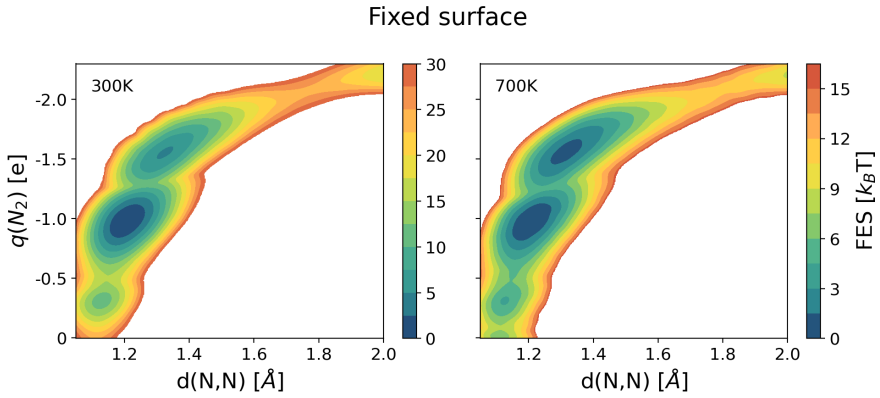

Fig. S11. N_2 **decomposition mechanism on a static surface.** Free energies as a function of N-N distance and N_2 at T=300 K and T=700 K from OPES simulations in which all the surface atoms are fixed. In this case no qualitative change between the low and high-temperature regimes are observed, except for a change of the relative stability of the different metastable states. Notably, at high temperature the alpha' precursor state is not destabilized but on the contrary is more stabilized. This is additional evidence that the observed results are caused by the major dynamic changes taking place at the surface, which therefore need to be fully accounted for.

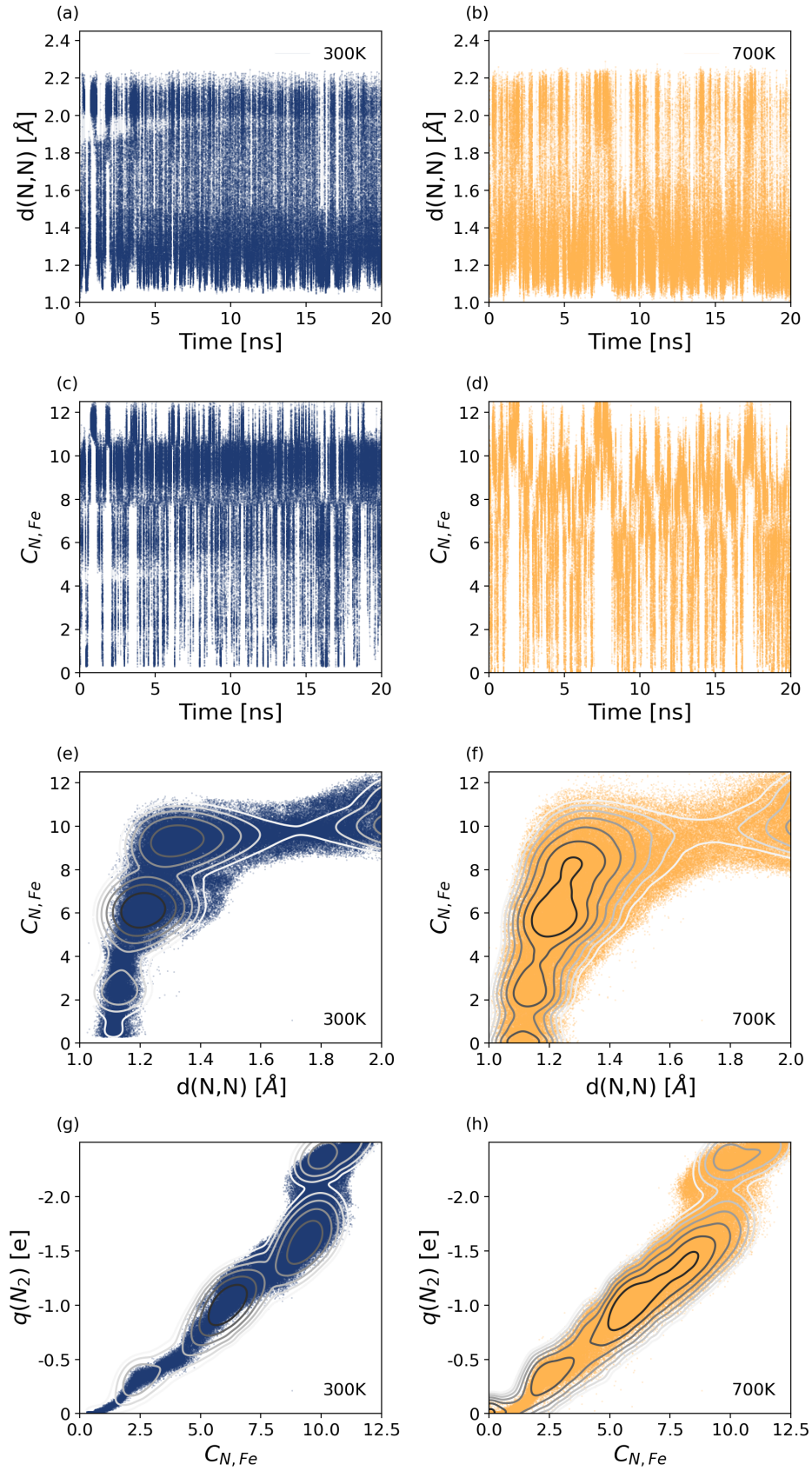

Fig. S12. Collective variables analysis. Time evolution of the CVs used and scatter plots of the explored space during the OPES simulations at the T=300 K (left panels) and T=700 K (right panels). (a-b) Time evolution of the N-N distance. (c-d) Time evolution of the N-Fe coordination number. (e-f) Scatter plot of the distance d and N-Fe coordination number. (g-h) Scatter plot of the N-Fe coordination and the charge transferred to the N_2 molecule. In panels e-h lines represent the isolines of the free energy surfaces, plotted every 4 k_BT (for T=300) and 2 k_BT (for T=700) with colors ranging from black to white.

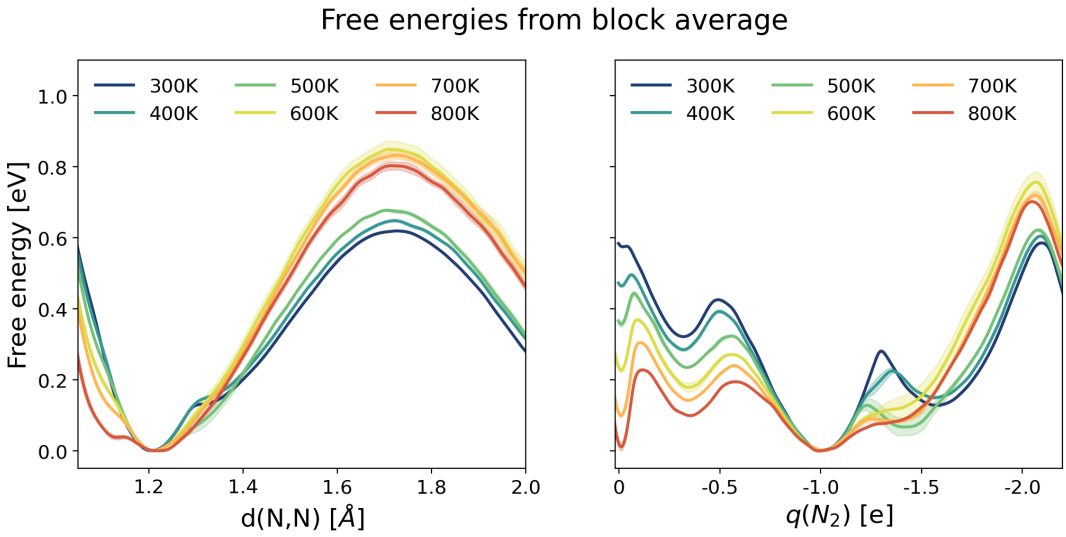

Fig. S13. Free energy profiles and uncertainties. FES as a function of distance N-N (left) and charge N_2 (right) are calculated using a block-weighted average citeInvernizzi2019 using three blocks of 5 ns each (the first 5 ns of the trajectory were discarded). Standard deviations are displayed with shaded curves. For 500 and 600 K temperatures, the maximum uncertainty on FES is 0.02 eV, while for all others it is always below 0.01 eV.

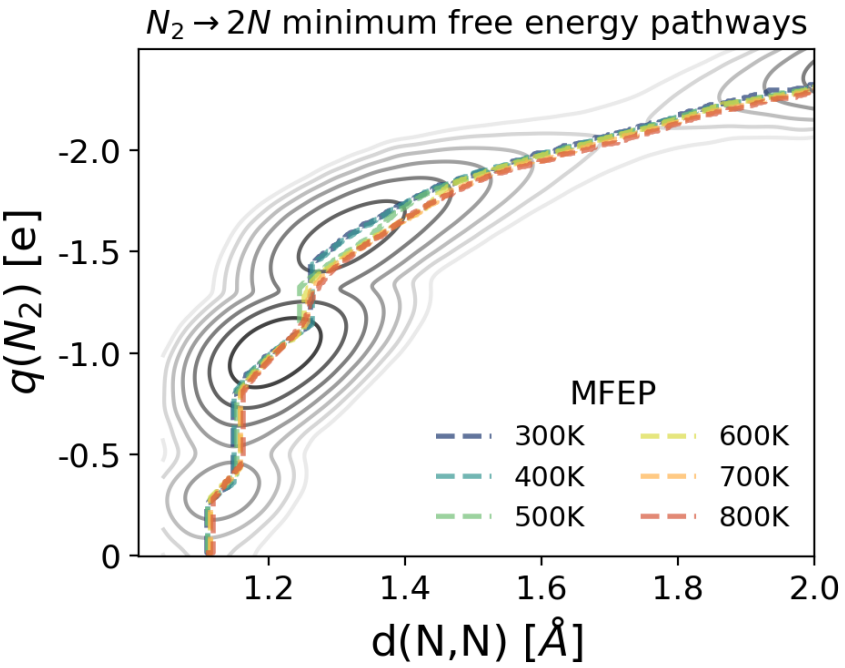

Fig. S14. Decomposition pathways versus temperature. The dashed lines represent the minimum free energy pathways calculated in the N-N distance and the $q(N_2)$ plane (dashed lines in Fig. S10), superimposed on the free energy at T=300K.

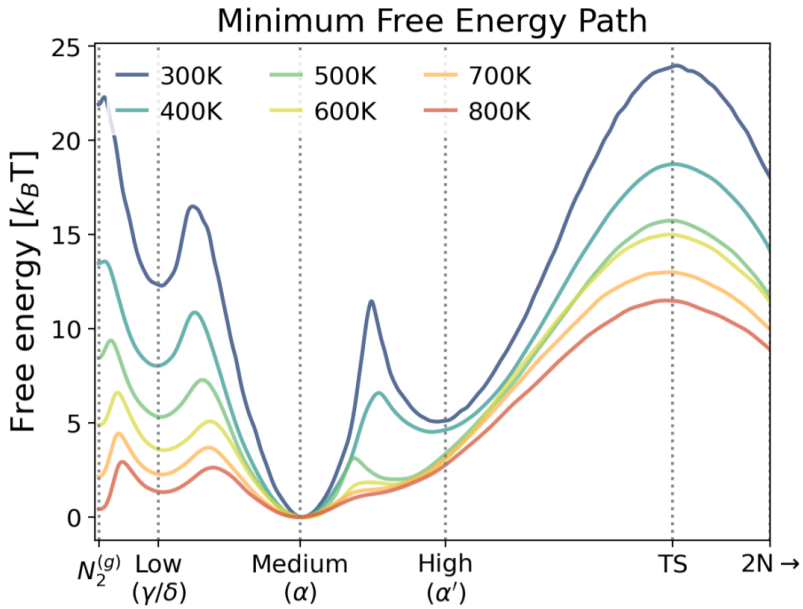

Fig. S15. Adsorption and decomposition free energy profile. Free energy along the minimum free energy pathways, as in Fig.6 in the manuscript, here rescaled by thermal energies $k_B T$.

4. Transition state characterization

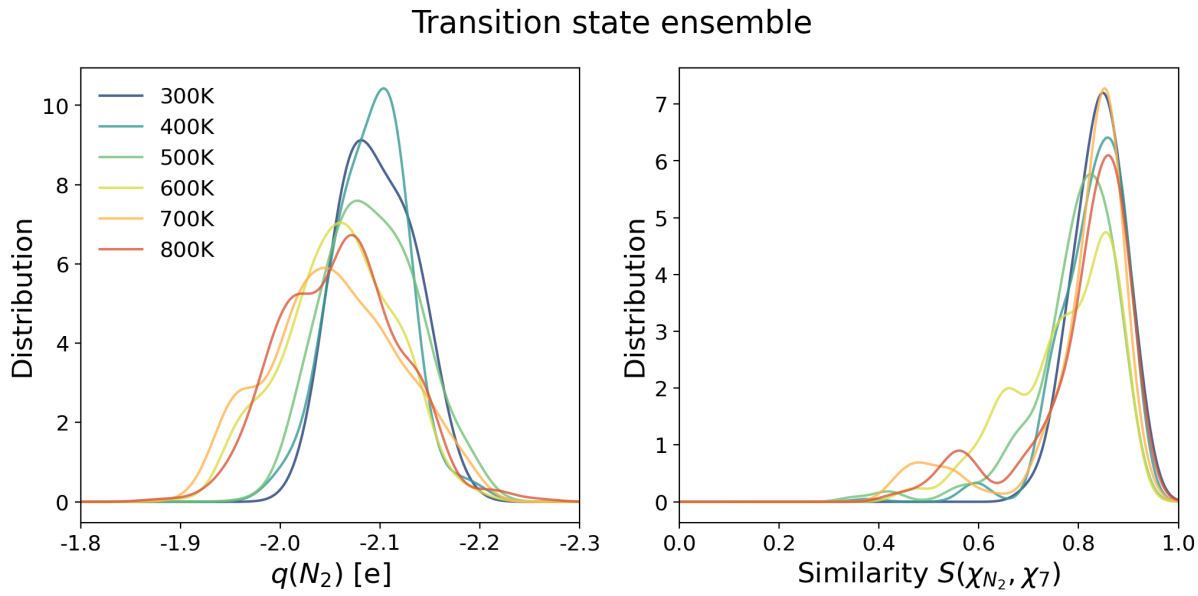

Fig. S16. Transition state ensemble. Distribution of the N_2 charge (left) and environment similarity between the environment of the Iron below (right) for the transition state ensemble configurations selected from the committor analysis. The distributions are normalized such that their integral sums to 1.

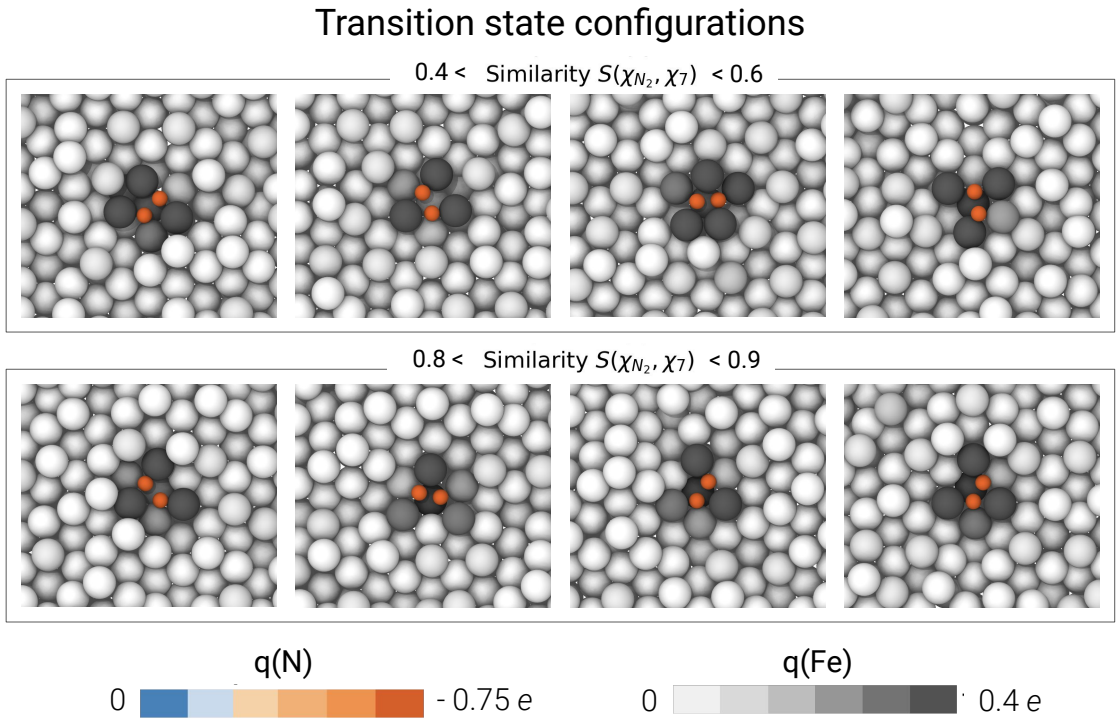

Fig. S17. Transition state snapshots. Configurations from the transition state ensemble at T=700K representative of the two peaks in Fig.7b. Snapshots of configurations with low similarity (top row) and high similarity (bottom) between the environment of the Fe atom on which N_2 is located and the reference site χ_7. The top row contains samples of the disordered active sites created by the dynamical roughening of the surface. Within each group, the snapshots are ordered by increasing similarity measures.

5. Temperature dependence of chemical reactivity

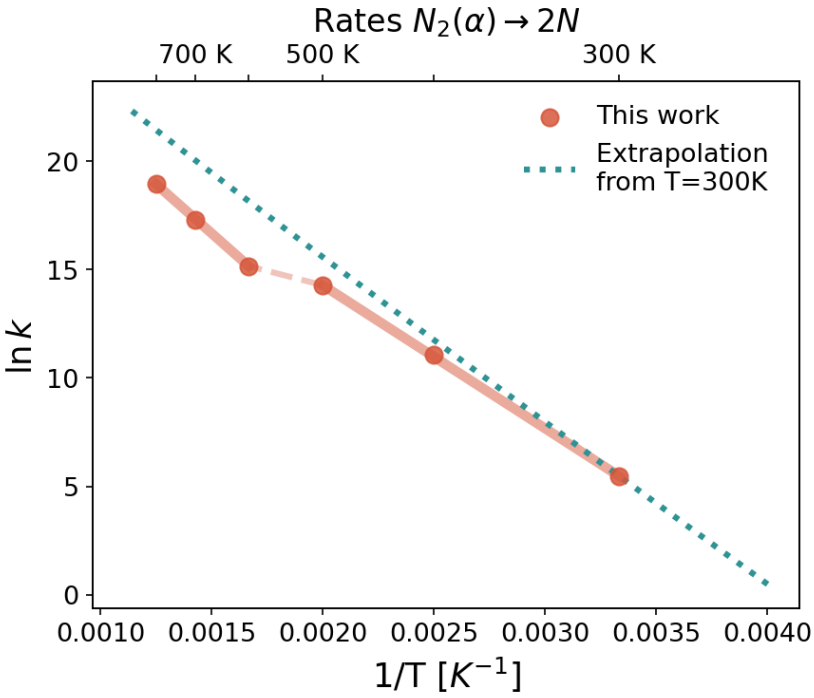

Fig. S18. Reaction rates and low-temperature extrapolation. The reaction rates computed with the Eyring-Polanyi equation are contrasted with an extrapolation from the low-temperature regime. This has been obtained by inserting in the Eyring-Polanyi formula the value of $\Delta G^{\dagger}$ estimated from the simulation at T=300 K (which is also close to previous T=0 K DFT calculations). This estimate obviously neglects any change in entropy. This extrapolation reproduces our data only up to T$\lesssim$ 500 K, while it overestimates the rates at higher temperatures, reflecting the existence of the two different regimes discussed in the paper. Above 500 K this temperature diffusion sets in and entropic contributions become important and must be taken explicitly into account. This is done automatically in our approach as the estimates of the free energy barriers from MD simulations already include entropy.