

Supporting Information

Nitrogen Reduction Reaction Energy and Pathway in Metal-zeolites: Deep Learning and Explainable Machine Learning with Local Acidity and Hydrogen Bonding Features

*Yuming Gu^{a, †}, Qin Zhu^{a, †}, Ziteng Liu^{a, †}, Cheng Fu^a, Jiayue Wu^a, Qiang Zhu^a,
Qingqing Jia^a, Jing Ma^{* a}*

^aKey Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, 210023, P. R. China

[†]These authors contribute equally

*E-mail: majing@nju.edu.cn

List of Contents

S1. Computational details.....	S2
S2. Construction of the ZA dataset.....	S4
S3. Multilevel attention graph convolutional neural network.....	S34
S4. Feature selection.....	S40
S5. Explainable machine learning models.....	S54
S6. Non-metal zeolites.....	S69
S7. Bandgap of metal-zeolites.....	S71
S8. Solvation effects.....	S72

S1. Computational details

Density functional theory (DFT) calculations of NRR pathways based on periodic boundary condition (PBC) model were performed by using the DMol³ module¹ in the Materials Studio software package². To test the performance of different functionals, PBE³, PW91⁴, and BLYP⁵⁻⁶ were applied for the rate limiting steps and bandgap calculations of the Ti-, Co-, and Nb-zeolites. The Grimme method⁷⁻⁸ was applied to take the van der Waals interaction into account for metal-zeolites (metal = Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Y, Zr, Nb, Mo, Ru, Rh, Pd, Ag, and Cd). A local basis cutoff of 4.0 Å in real space was employed. The energy convergence value between two consecutive steps was chosen as 2×10^{-5} hartree. The value of 4.0×10^{-3} hartree Å⁻¹ was set for the gradient, and the displacement value was allowed lower than 5×10^{-3} Å.

The relative energy (ΔE) of the intermediates in NRR was calculated by employing the computational hydrogen electrode (CHE) model, which used one-half of the chemical potential of hydrogen as the chemical potential of the proton-electron pair.⁹ The value of ΔE between two steps involved in the NRR process can be expressed as to:

$$\Delta E = E(\text{N}_{2-m}\text{H}_n@\text{zeolite}) + m E(\text{NH}_3) - E(\text{N}_2) - n/2 E(\text{H}_2) - E(\text{zeolite}) \quad (\text{S1}),$$

where $E(\text{N}_{2-m}\text{H}_n@\text{zeolite})$ was the energy of intermediates in the zeolites. $E(\text{NH}_3)$, $E(\text{N}_2)$, $E(\text{H}_2)$, and $E(\text{zeolite})$ were the energy of ammonia, dinitrogen, hydrogen, and substrate, respectively. The value, n is the number of H⁺/e⁻ pairs transferred, and m

$(m=0, 1)$ is the number of ammonia molecules released.

S2. Construction of the ZA dataset

Table S1. ZA dataset for experimental data

Zeolite	Ammonia production (mg/g) ^a
Ti-3A	0.4
Ti-4A	0.22
Ti-5A	0.55

^a taken from ref ¹⁰⁻¹¹

Table S2. ZA dataset for theoretical data

Metal	system	N_{HB}	χ	EA	IE	N_d	Z	R	M	N_N	N_H	Pathway	$(Si/Al)^{-1}$	PLD	R_{DLS}	Q	Q_f	ΔQ_{CT}	LA	d_N	d_{O1}	d_{O2}	d_{O3}	d_{O4}	d_{Si}	ΔE (eV)
Sc	Al-3-MFI-Sc-N ₂	0	1.36	0.1880	6.5616	1	21	162.00	44.96	2	0	1	0.03	5.277	0.1276	1.717	1.739	0.022	0.365	2.062	2.095	2.183	2.238	2.297	2.372	-0.97
	Al-3-MFI-Sc-NNH	1	1.36	0.1880	6.5616	1	21	162.00	44.96	2	1	1	0.03	5.277	0.1276	1.717	1.555	-0.162	0.365	2.103	2.114	2.171	2.181	2.238	2.831	-0.78
	Al-3-MFI-Sc-NNH ₂	2	1.36	0.1880	6.5616	1	21	162.00	44.96	2	2	1	0.03	5.277	0.1276	1.717	1.688	-0.029	0.365	2.069	2.135	2.145	2.192	2.234	2.828	-1.40
	Al-3-MFI-Sc-NNH ₃	3	1.36	0.1880	6.5616	1	21	162.00	44.96	2	3	1	0.03	5.277	0.1276	1.717	1.585	-0.132	0.365	1.924	2.119	2.191	2.224	2.274	2.851	-1.19
	Al-3-MFI-Sc-N	0	1.36	0.1880	6.5616	1	21	162.00	44.96	1	0	1	0.03	5.277	0.1276	1.717	1.588	-0.129	0.365	2.074	2.112	2.205	2.205	2.267	2.609	3.49
	Al-3-MFI-Sc-NH	0	1.36	0.1880	6.5616	1	21	162.00	44.96	1	1	1	0.03	5.277	0.1276	1.717	1.619	-0.098	0.365	1.941	2.099	2.215	2.222	2.279	2.859	0.60
	Al-3-MFI-Sc-NH ₂	1	1.36	0.1880	6.5616	1	21	162.00	44.96	1	2	1	0.03	5.277	0.1276	1.717	1.681	-0.036	0.365	1.995	2.088	2.197	2.203	2.265	2.704	-2.56
	Al-3-MFI-Sc-NH ₃	1	1.36	0.1880	6.5616	1	21	162.00	44.96	1	3	1	0.03	5.277	0.1276	1.717	1.772	0.055	0.365	2.058	2.129	2.188	2.203	2.219	2.683	-5.00
	Al-3-MFI-Sc-NHNH	1	1.36	0.1880	6.5616	1	21	162.00	44.96	2	2	2	0.03	5.277	0.1276	1.717	1.721	0.004	0.365	2.105	2.140	2.240	2.241	2.259	2.266	-1.23
	Al-3-MFI-Sc-NHNH ₂	1	1.36	0.1880	6.5616	1	21	162.00	44.96	2	3	2	0.03	5.277	0.1276	1.717	1.717	0.000	0.365	2.085	2.139	2.160	2.224	2.259	2.706	-2.47
	Al-3-MFI-Sc-NH ₂ NH ₂	2	1.36	0.1880	6.5616	1	21	162.00	44.96	2	4	2	0.03	5.277	0.1276	1.717	1.806	0.089	0.365	2.096	2.166	2.208	2.234	2.256	2.263	-4.01
	Al-3-MFI-Sc-NH ₂ NH ₃	2	1.36	0.1880	6.5616	1	21	162.00	44.96	2	5	2	0.03	5.277	0.1276	1.717	1.289	-0.428	0.365	2.152	2.219	2.257	2.367	2.408	2.847	-2.87

Note: To save the space, we only list DFT data for several metal-zeolites. If one needs the complete set of data, please write to majing@nju.edu.cn.

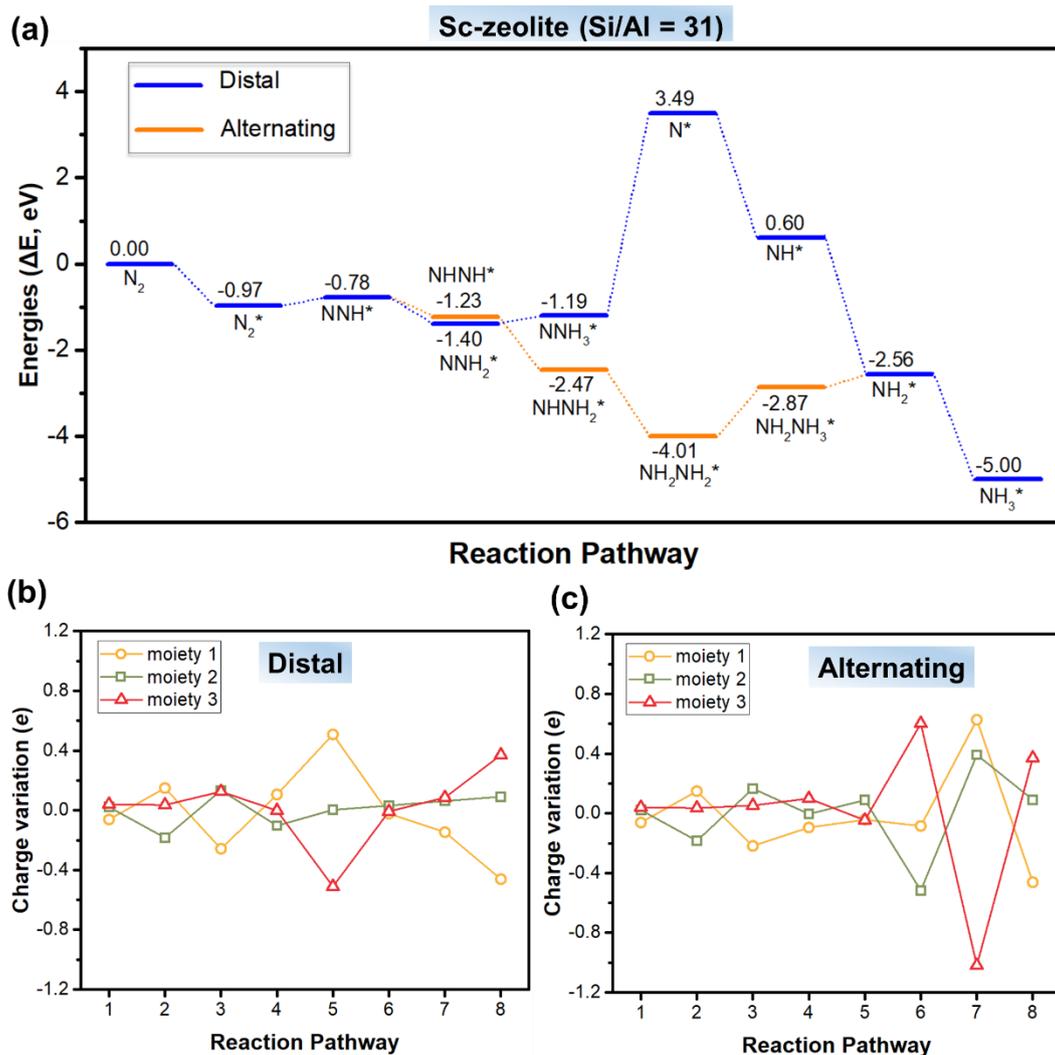


Fig. S1 (a) Energy diagrams for NRR in Sc-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

Each structure can be divided into three moieties: zeolite (moiety 1, substrate), metal (moiety 2, metal center), and the adsorbed $N_{2-m}H_n$ species (moiety 3, intermediates), respectively.

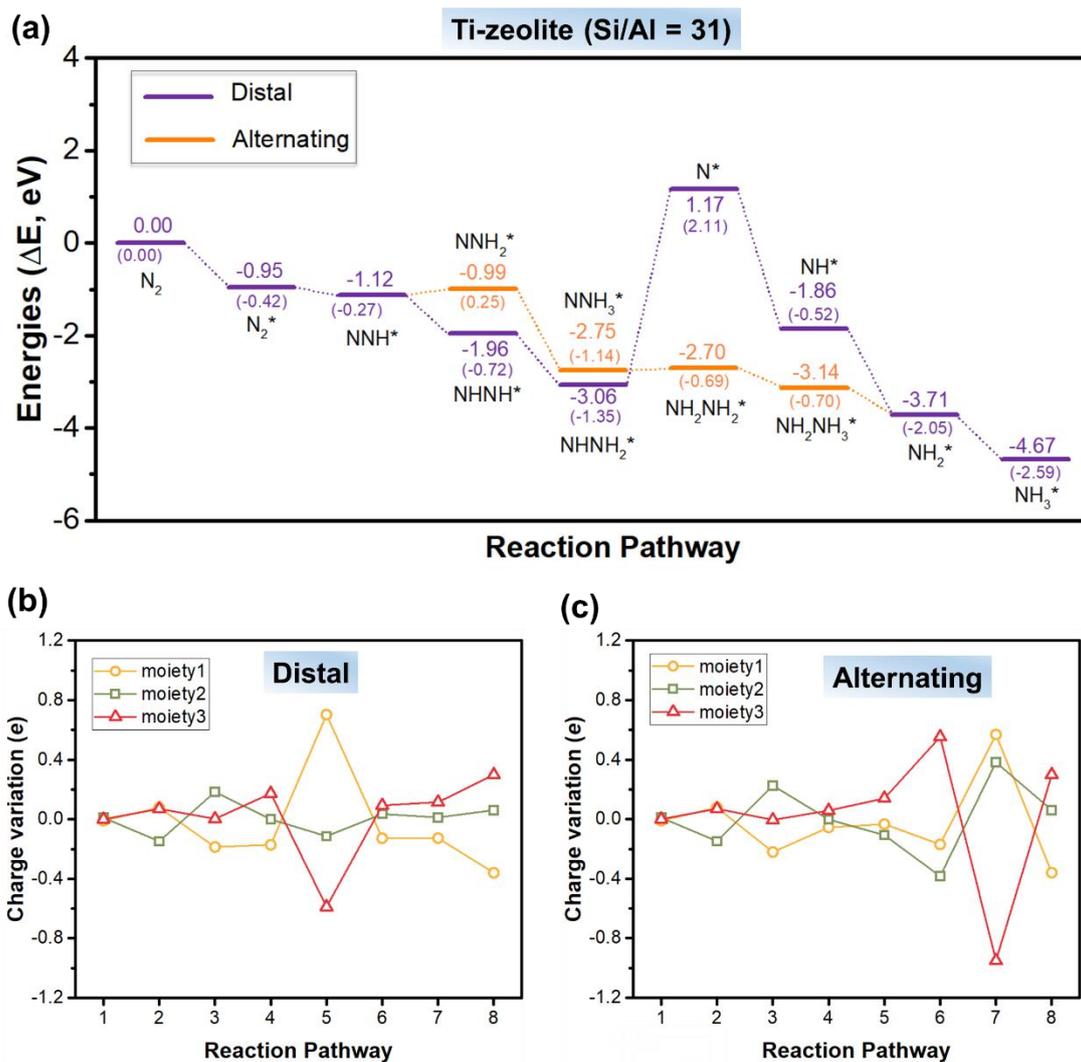


Fig. S2 (a) Energy diagrams for NRR in Ti-zeolite (Si/Al = 31) via distal and alternating pathways with free energy changes in parenthesis; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

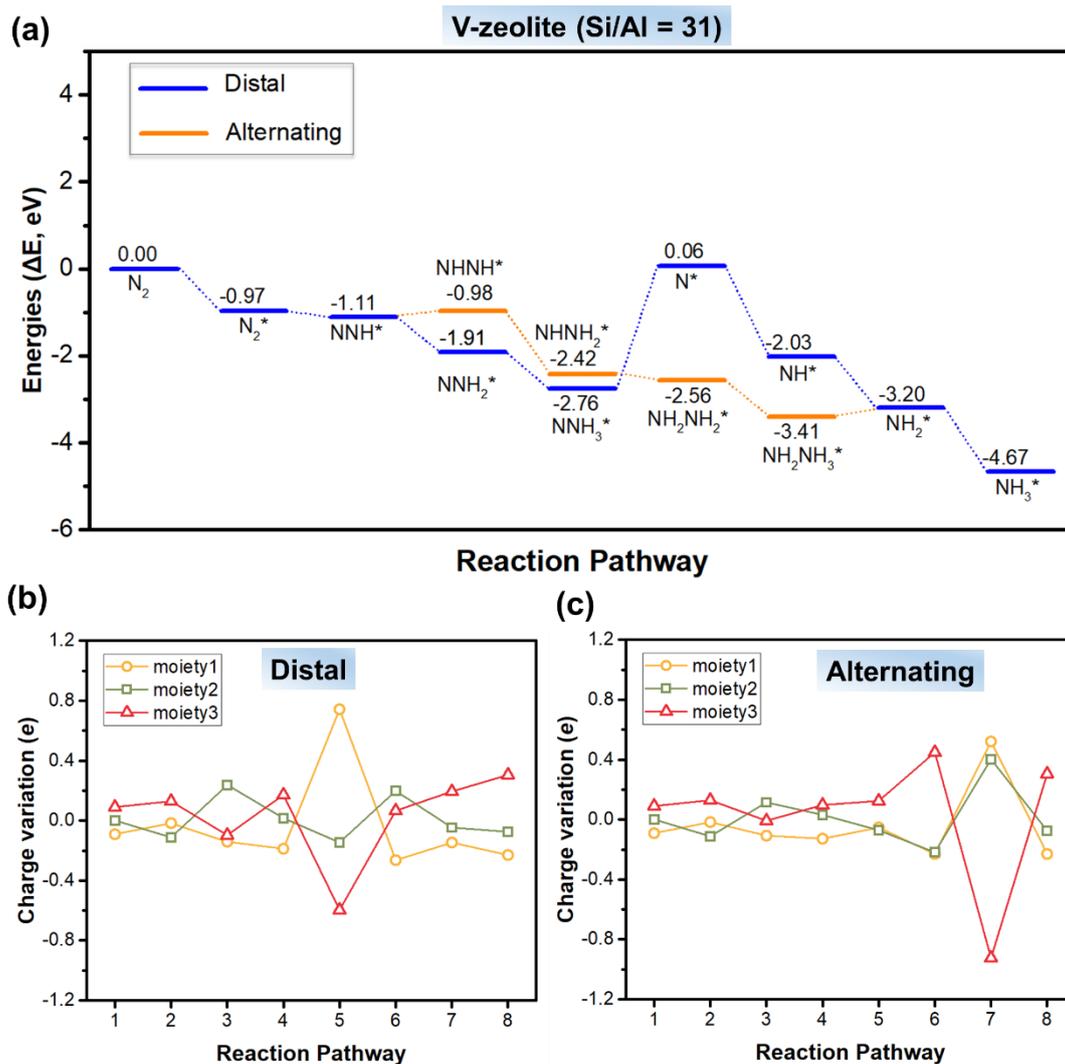


Fig. S3 (a) Energy diagrams for NRR in V-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

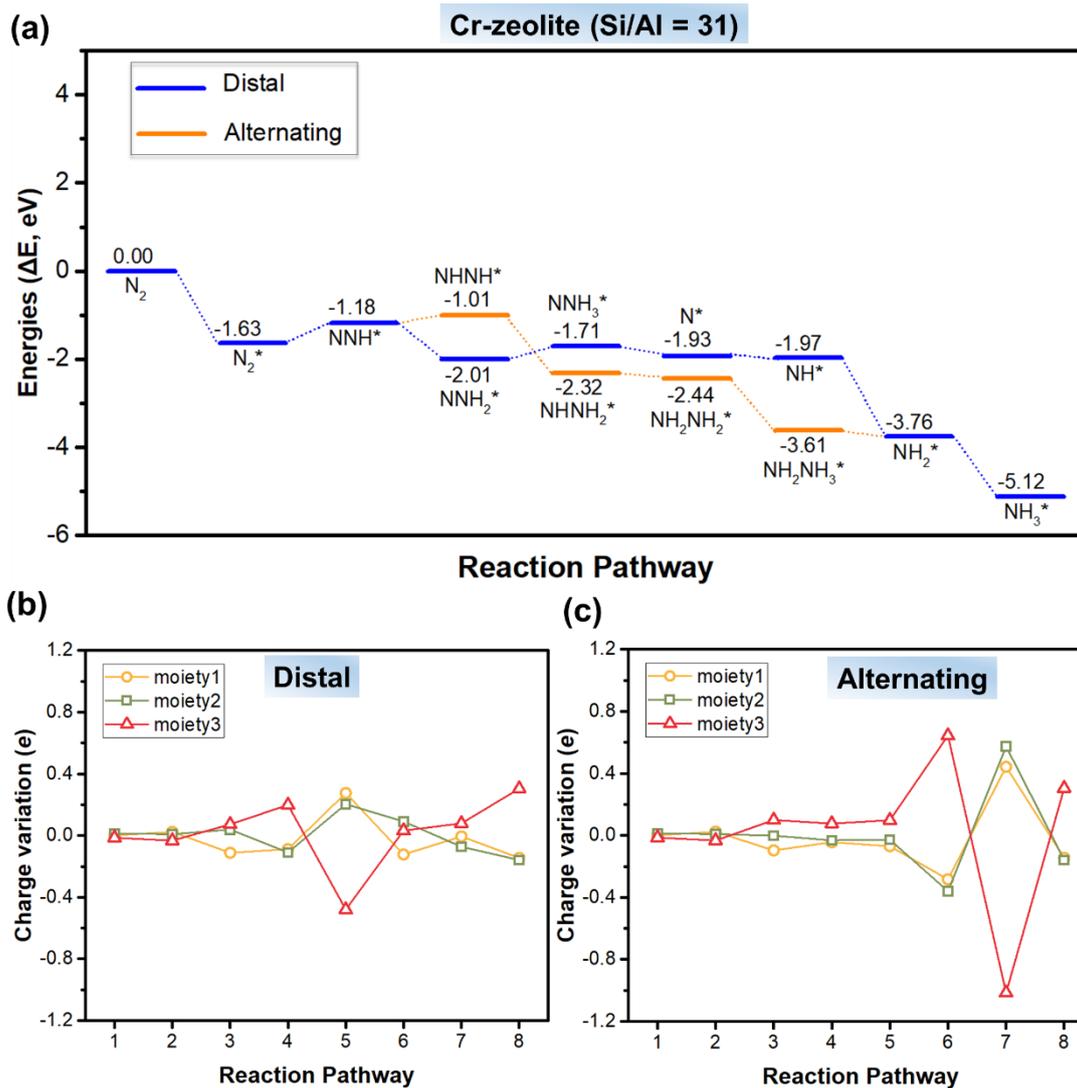


Fig. S4 (a) Energy diagrams for NRR in Cr-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

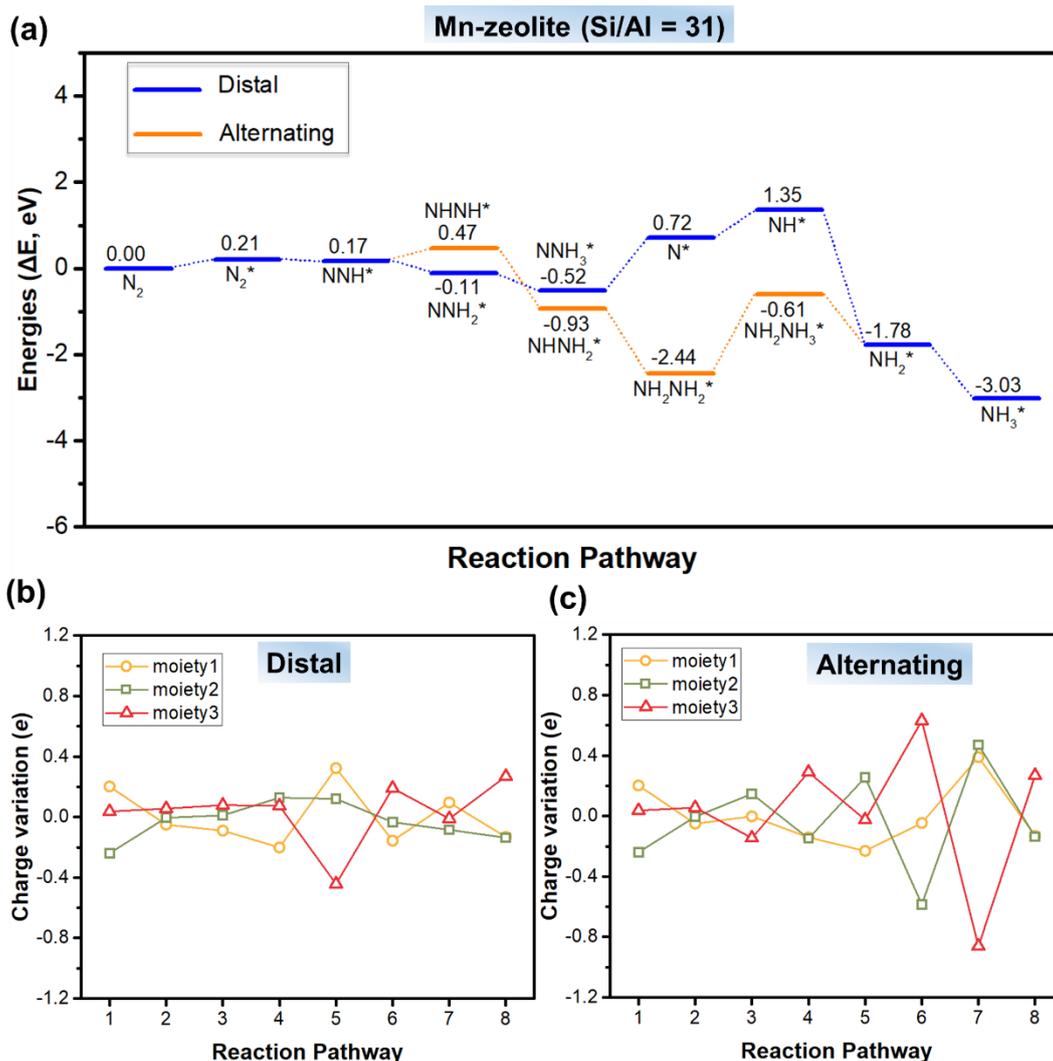


Fig. S5 (a) Energy diagrams for NRR in Mn-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

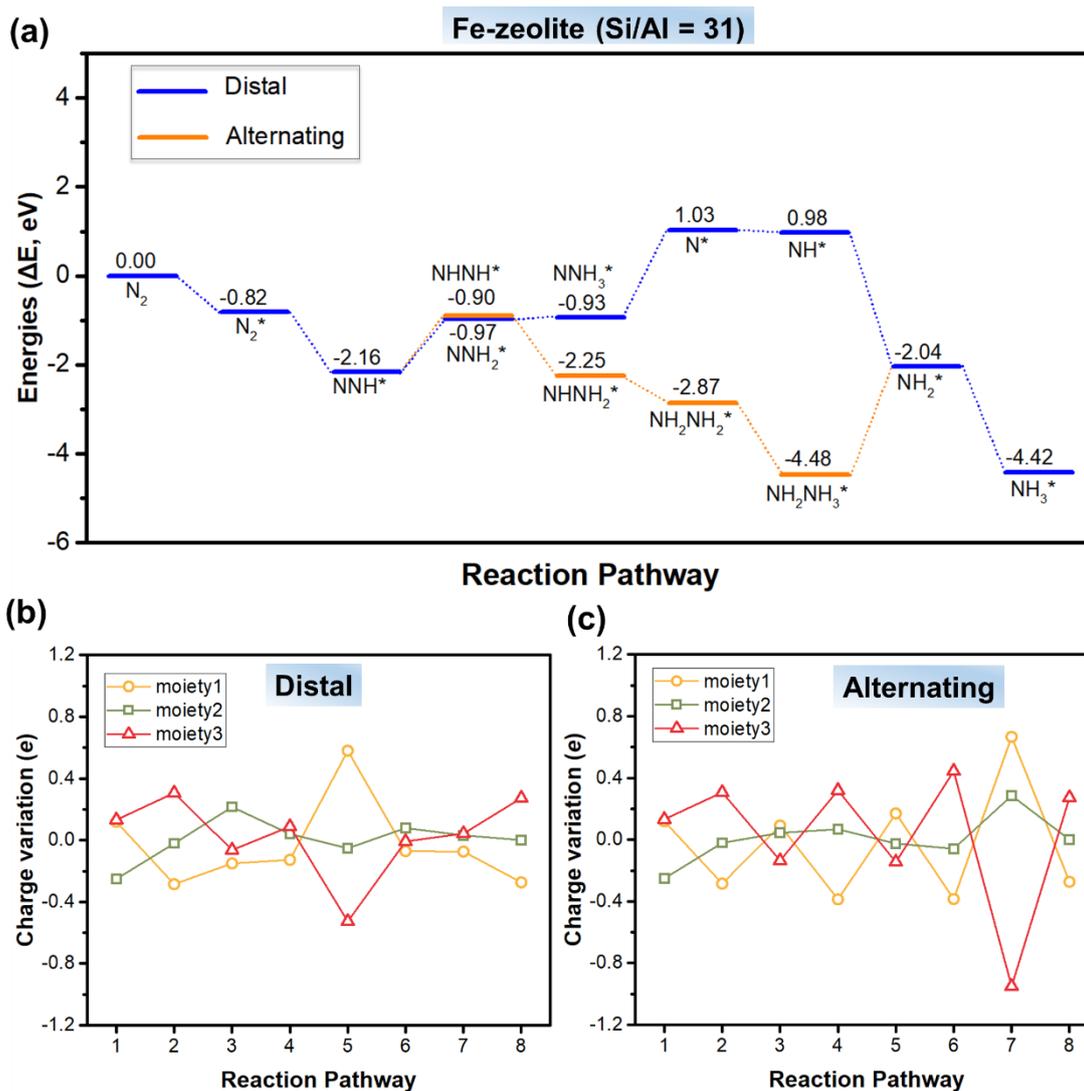


Fig. S6 (a) Energy diagrams for NRR in Fe-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

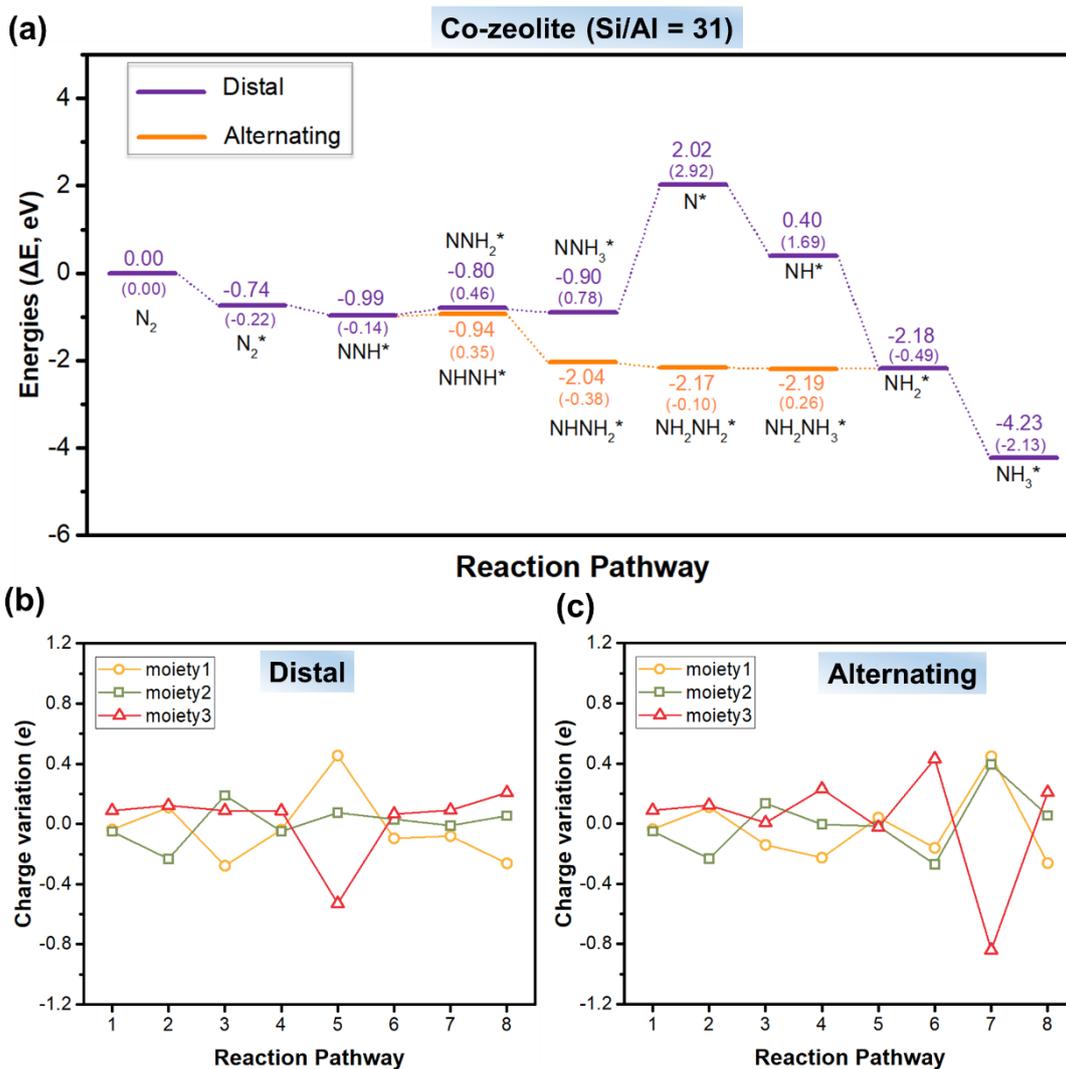


Fig. S7 (a) Energy diagrams for NRR in Co-zeolite (Si/Al = 31) via distal and alternating pathways with free energy changes in parenthesis; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

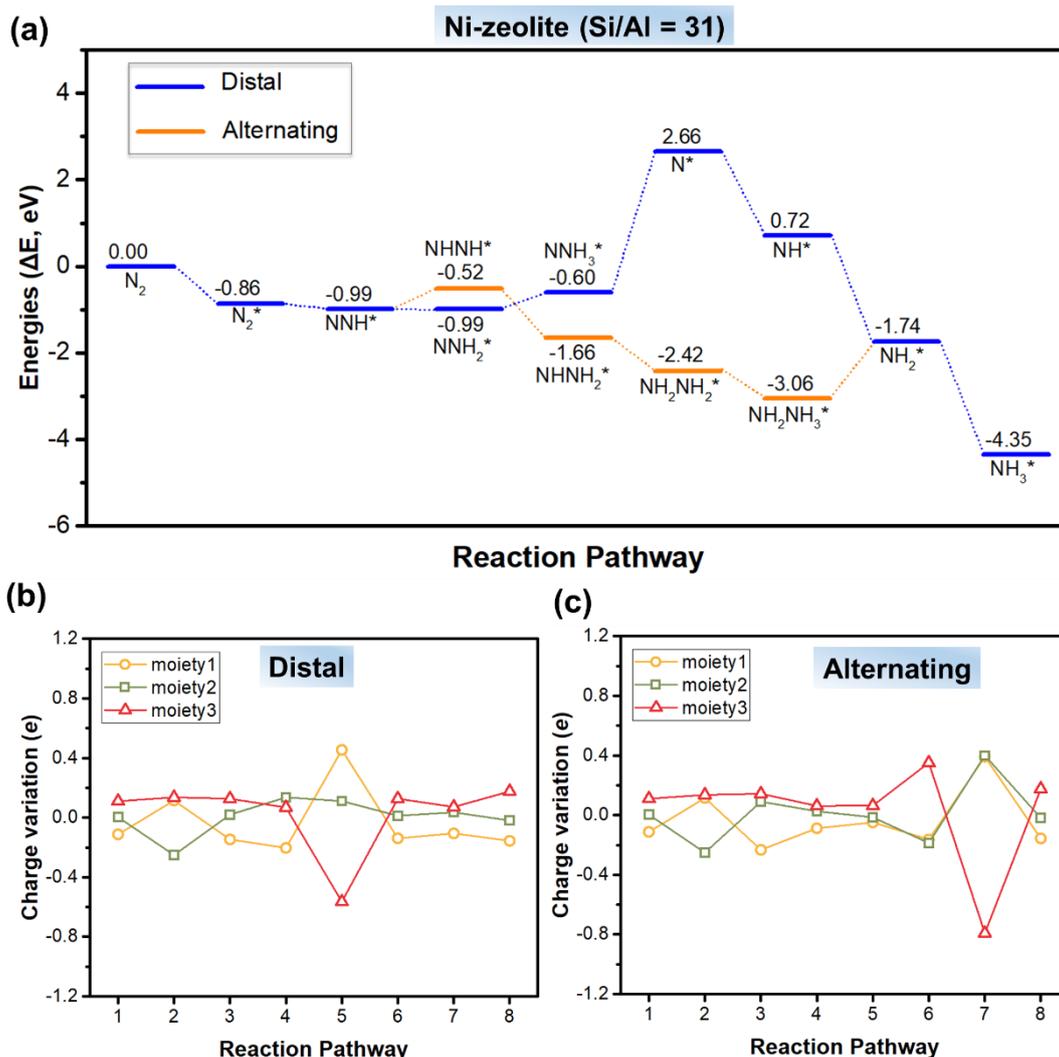


Fig. S8 (a) Energy diagrams for NRR in Ni-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

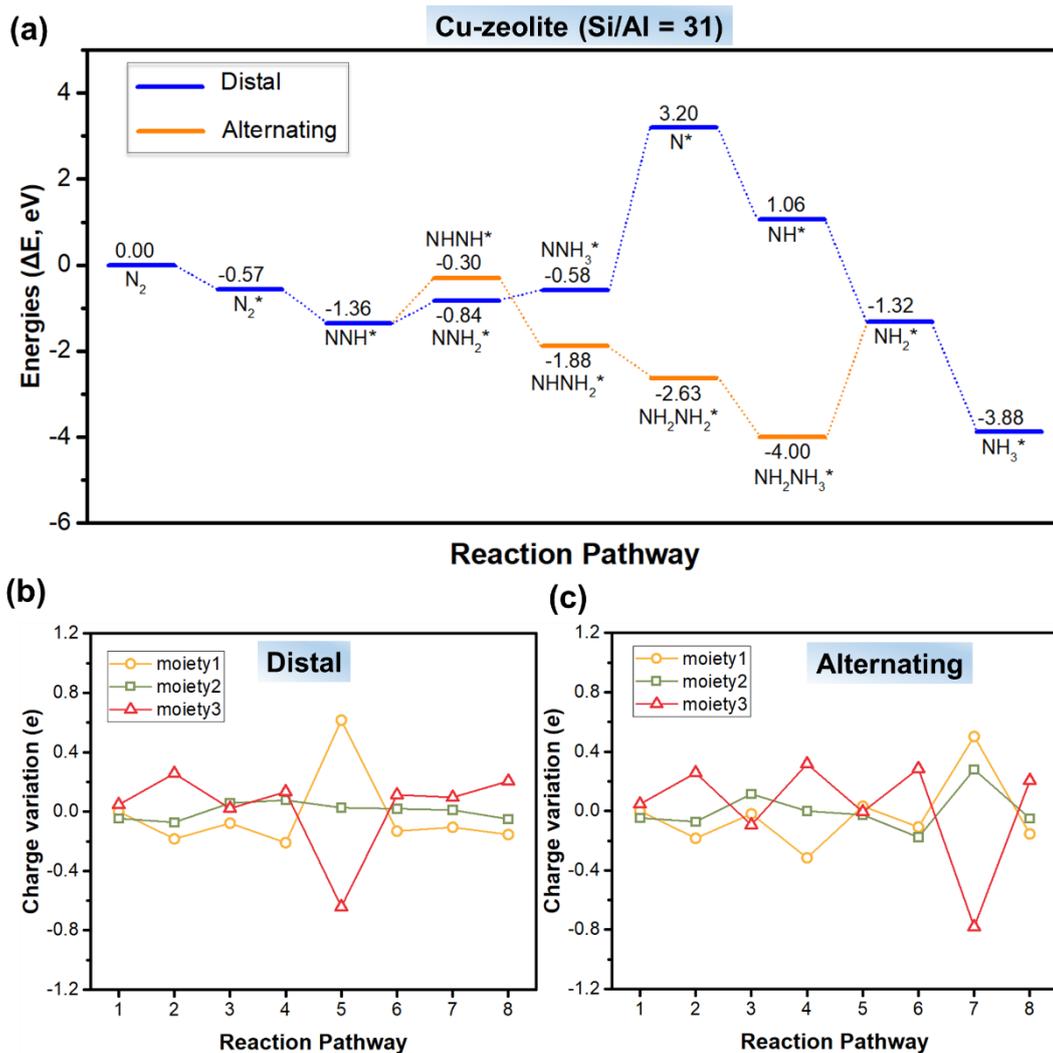


Fig. S9 (a) Energy diagrams for NRR in Cu-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

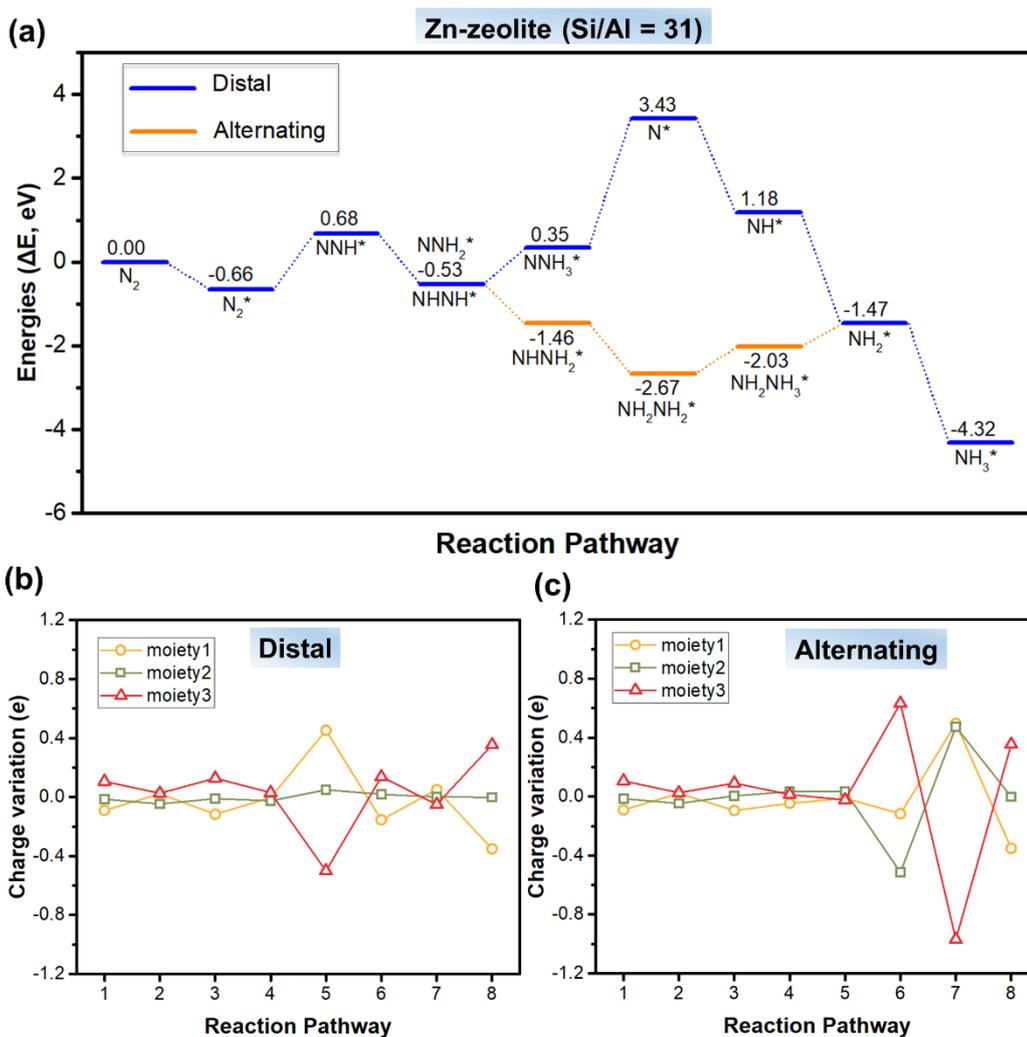


Fig. S10 (a) Energy diagrams for NRR in Zn-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

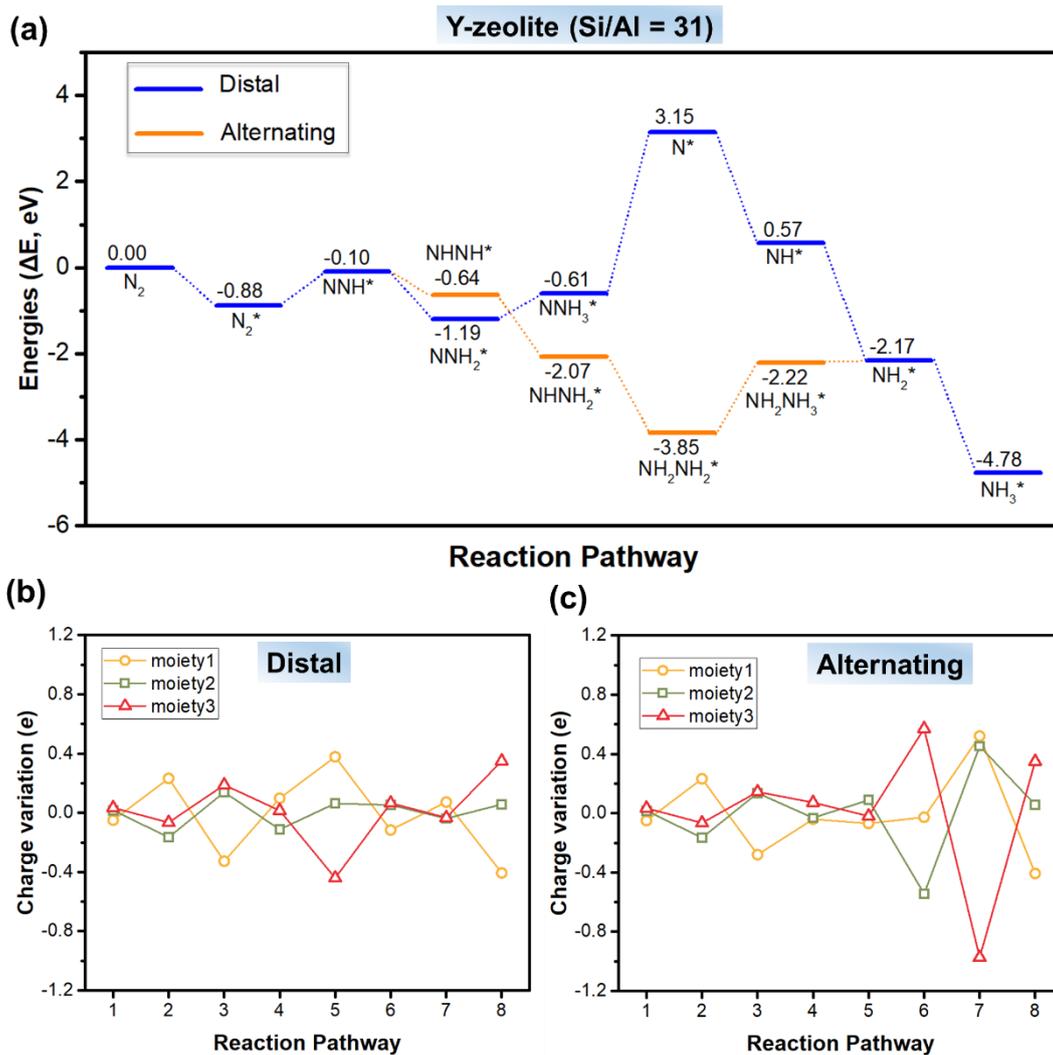


Fig. S11 (a) Energy diagrams for NRR in Y-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

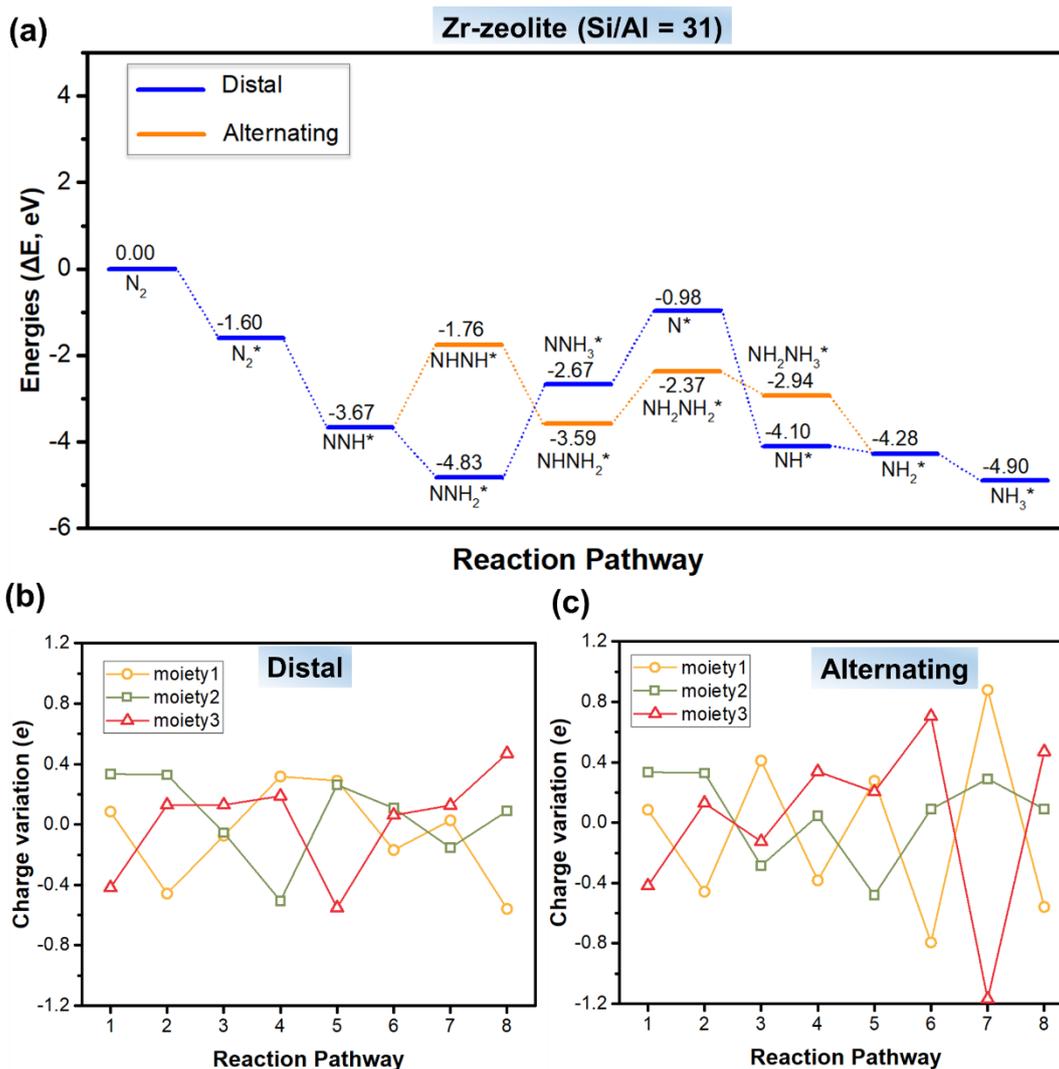


Fig. S12 (a) Energy diagrams for NRR in Zr-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

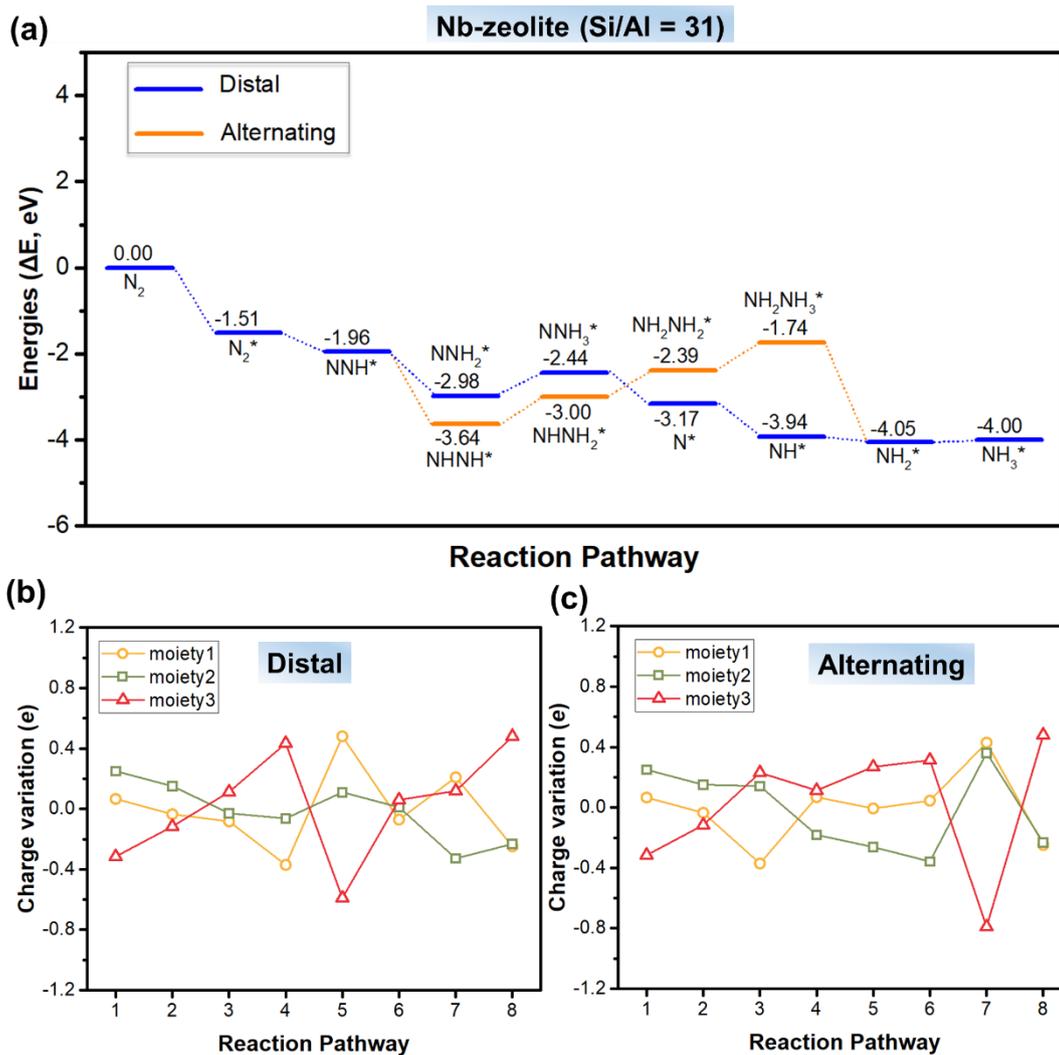


Fig. S13 (a) Energy diagrams for NRR in Nb-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

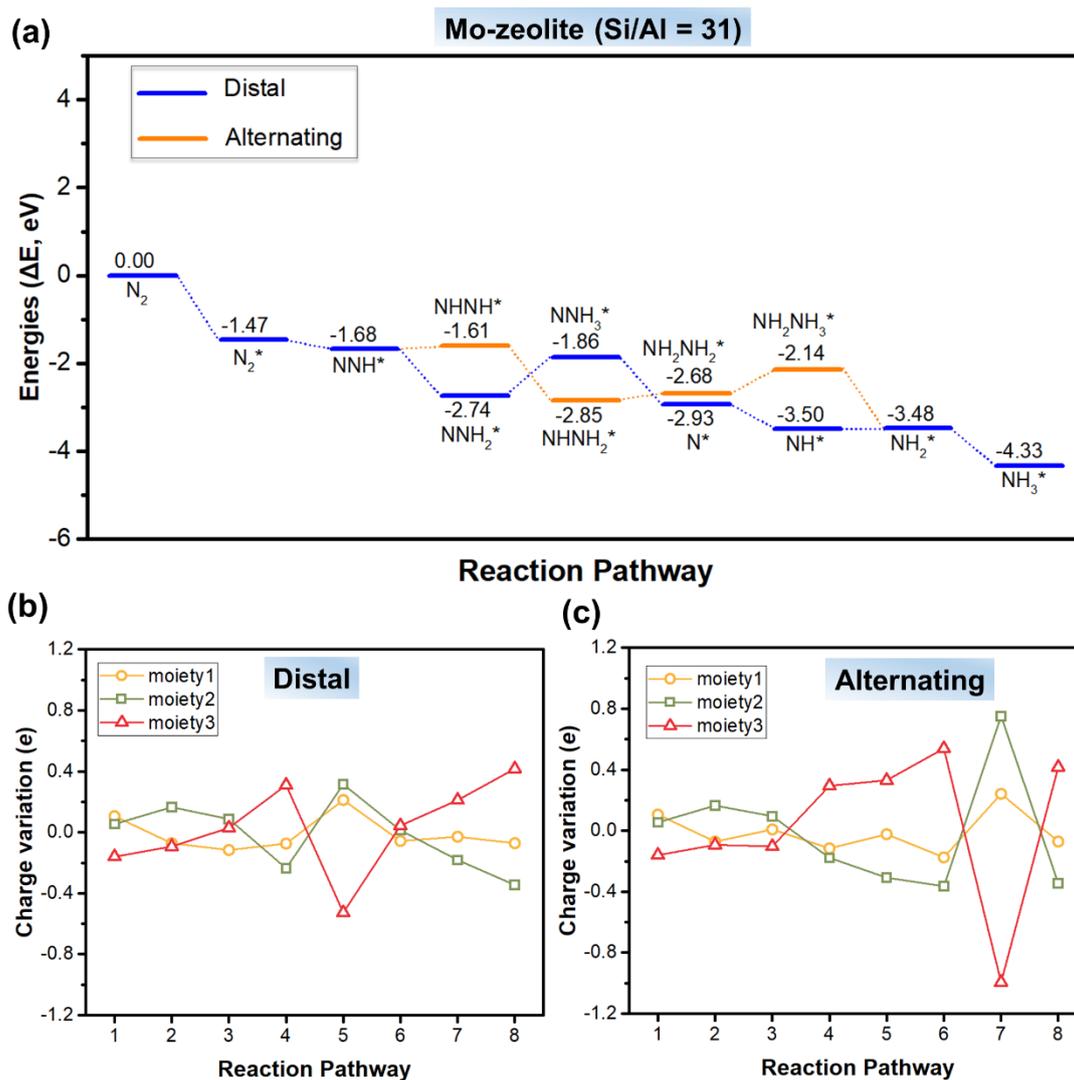


Fig. S14 (a) Energy diagrams for NRR in Mo-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

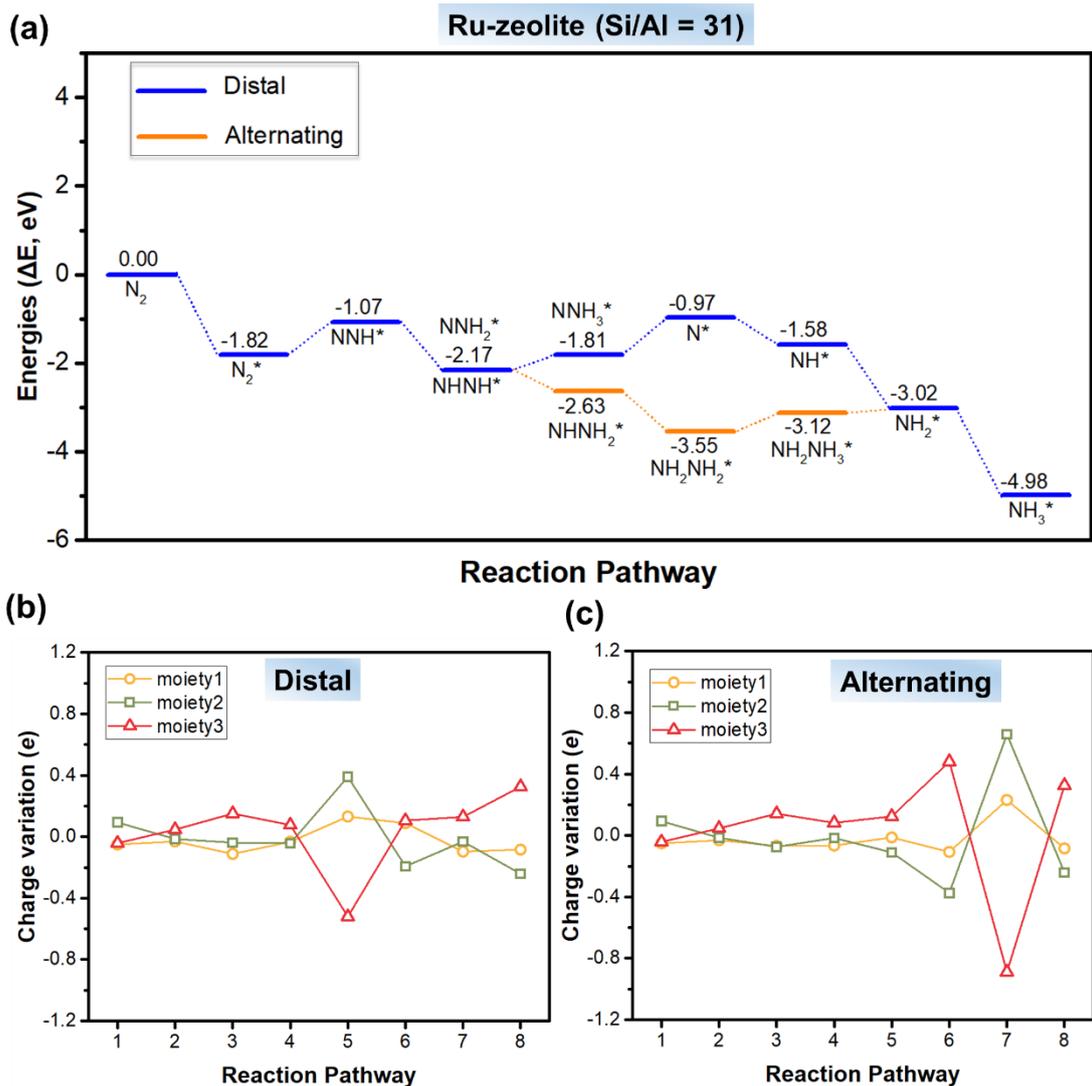


Fig. S15 (a) Energy diagrams for NRR in Ru-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

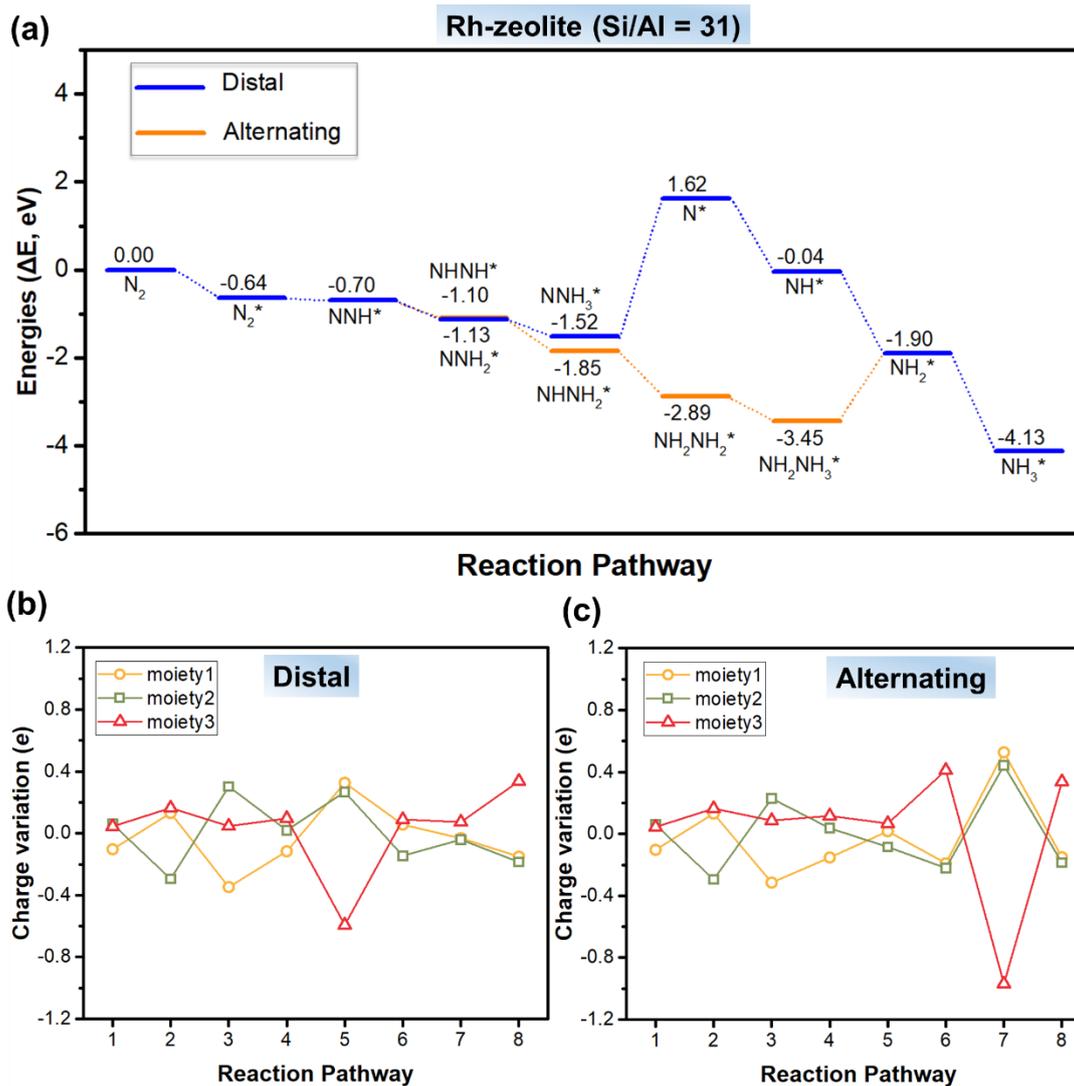


Fig. S16 (a) Energy diagrams for NRR in Rh-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

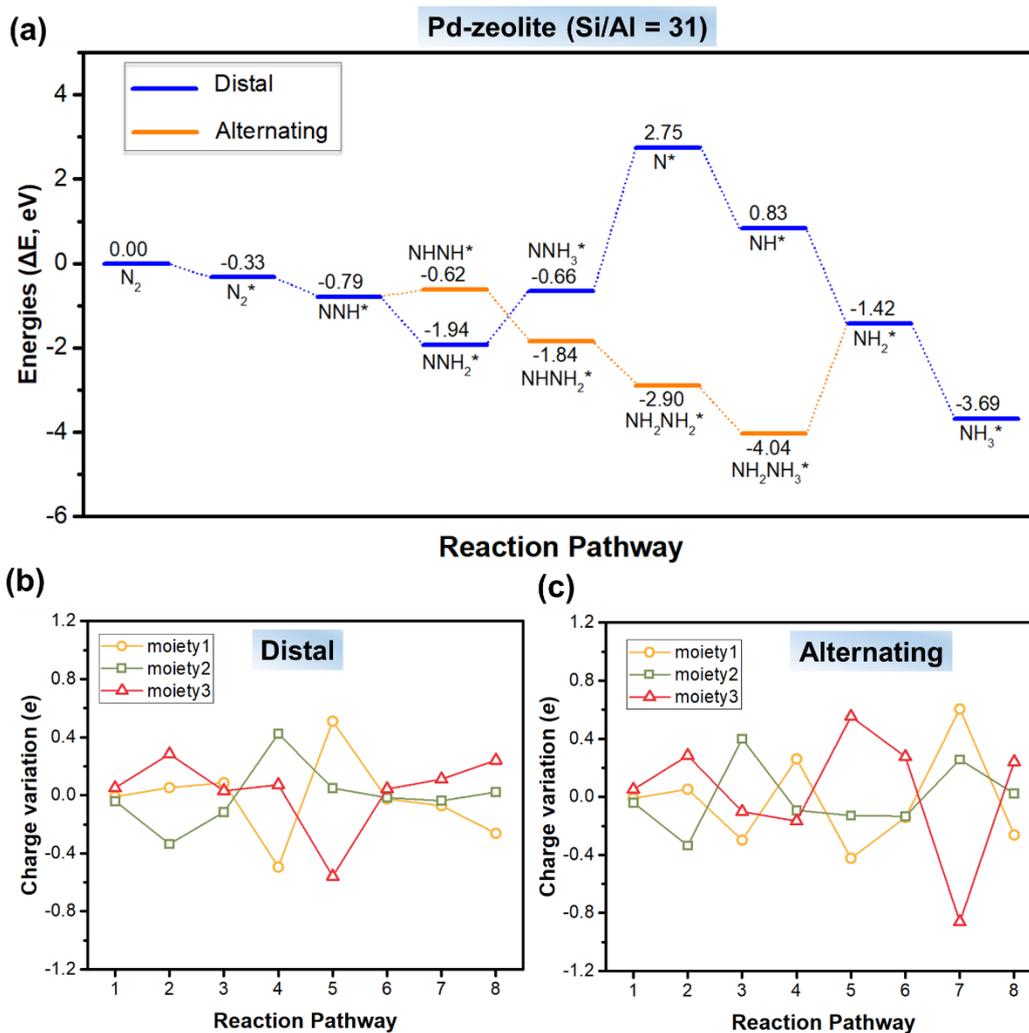


Fig. S17 (a) Energy diagrams for NRR in Pd-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

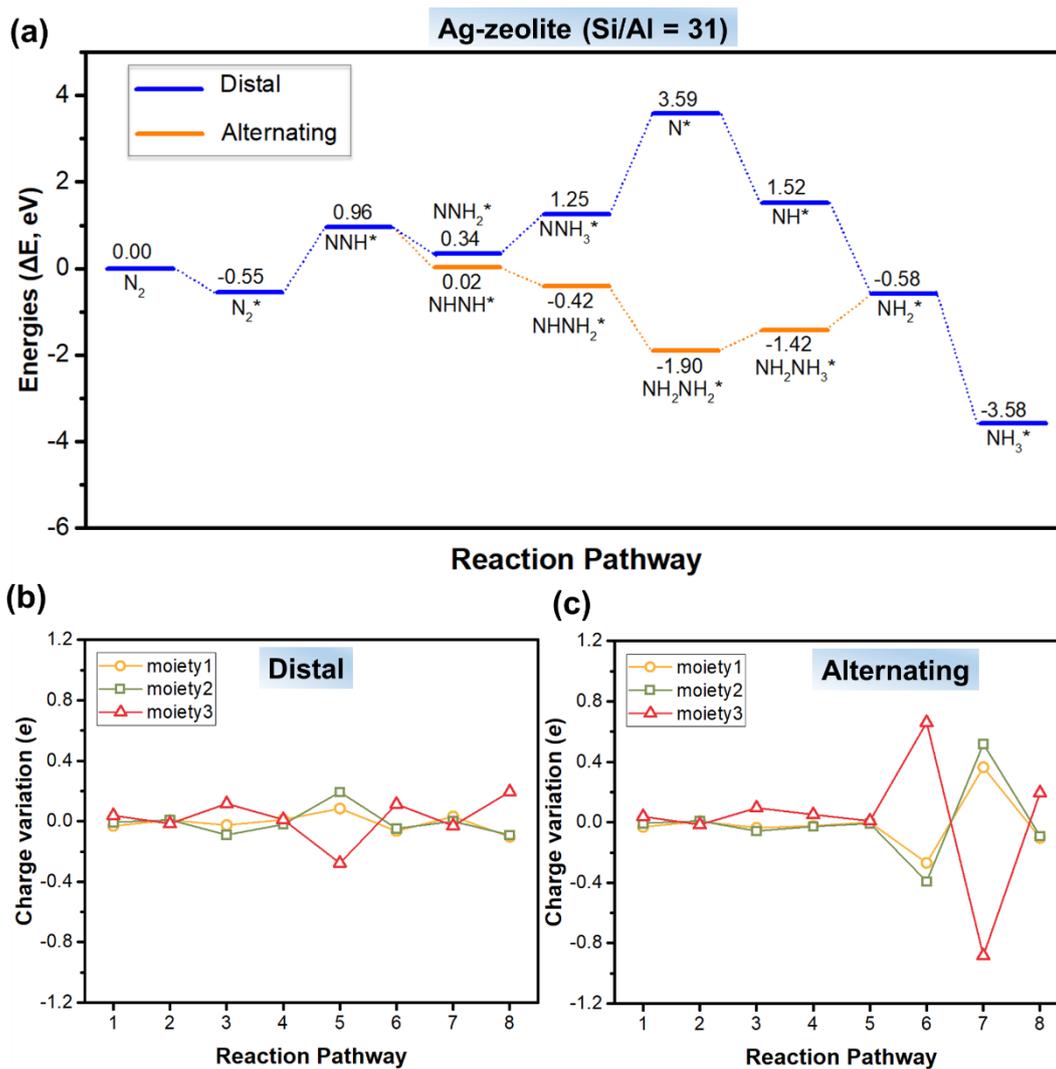


Fig. S18 (a) Energy diagrams for NRR in Ag-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

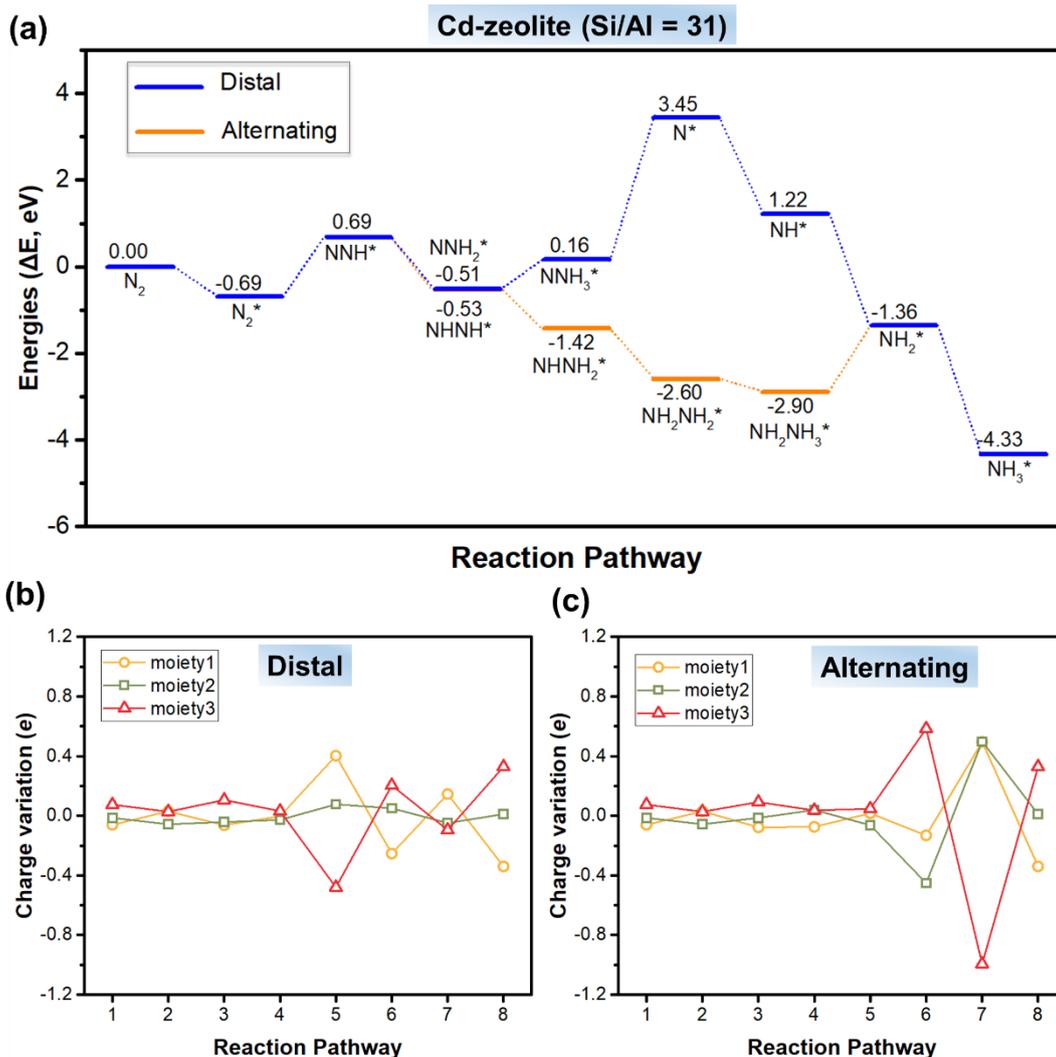


Fig. S19 (a) Energy diagrams for NRR in Cd-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

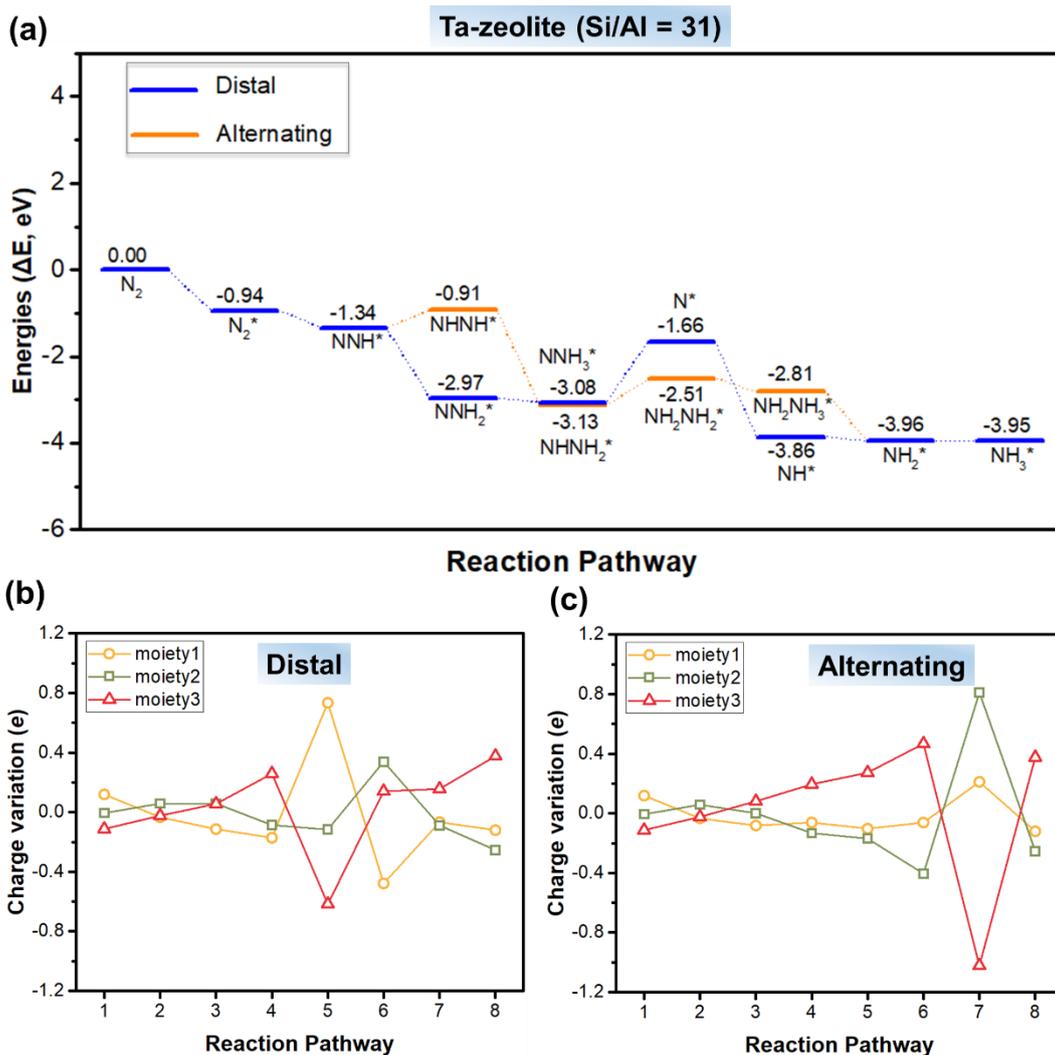


Fig. S20 (a) Energy diagrams for NRR in Ta-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

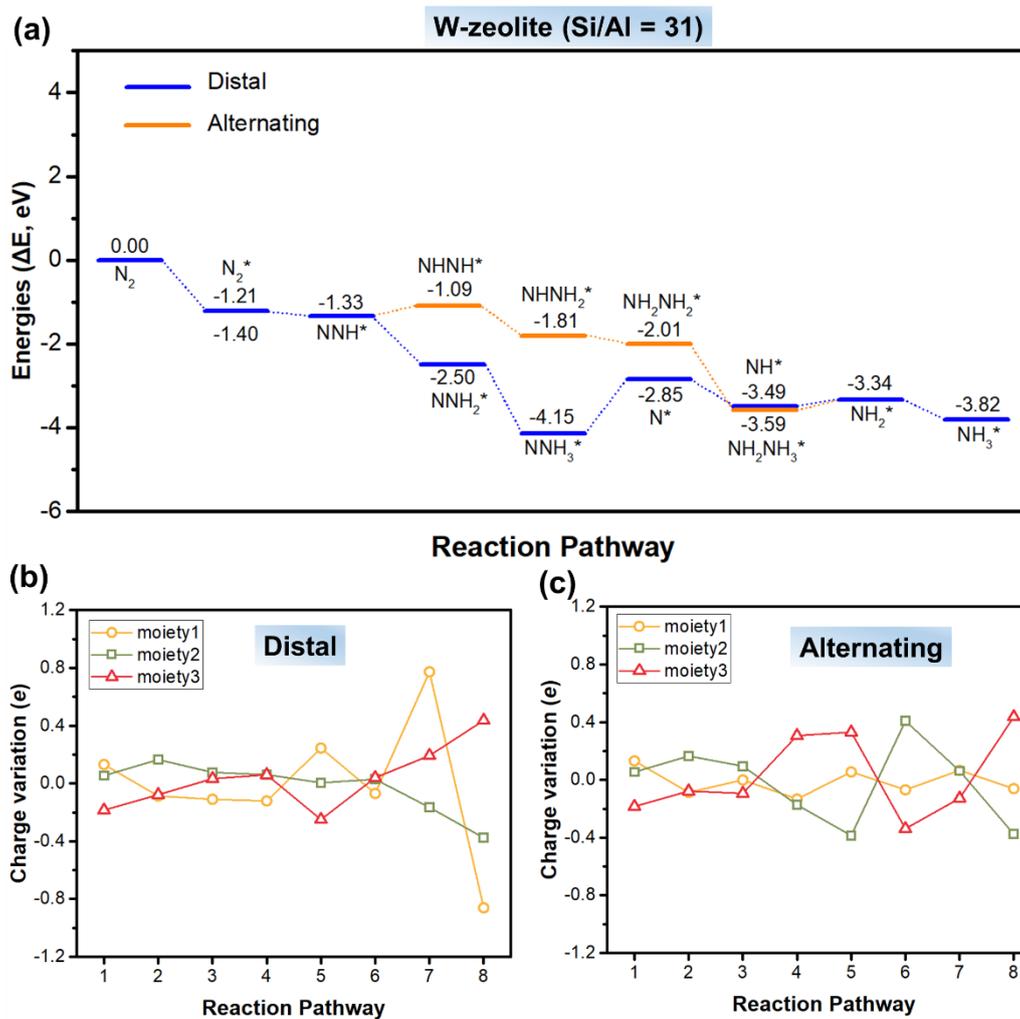


Fig. S21 (a) Energy diagrams for NRR in W-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

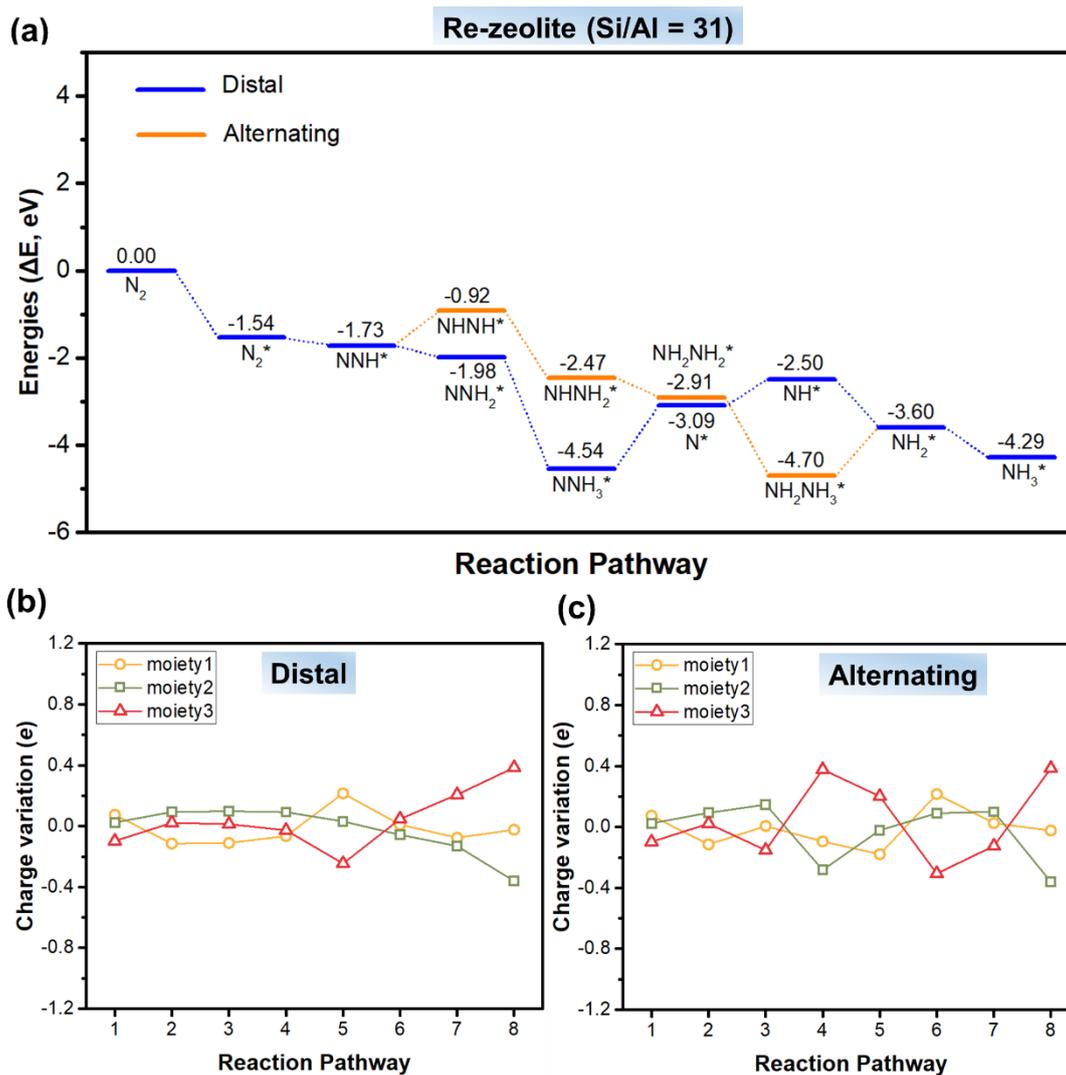


Fig. S22 (a) Energy diagrams for NRR in Re-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

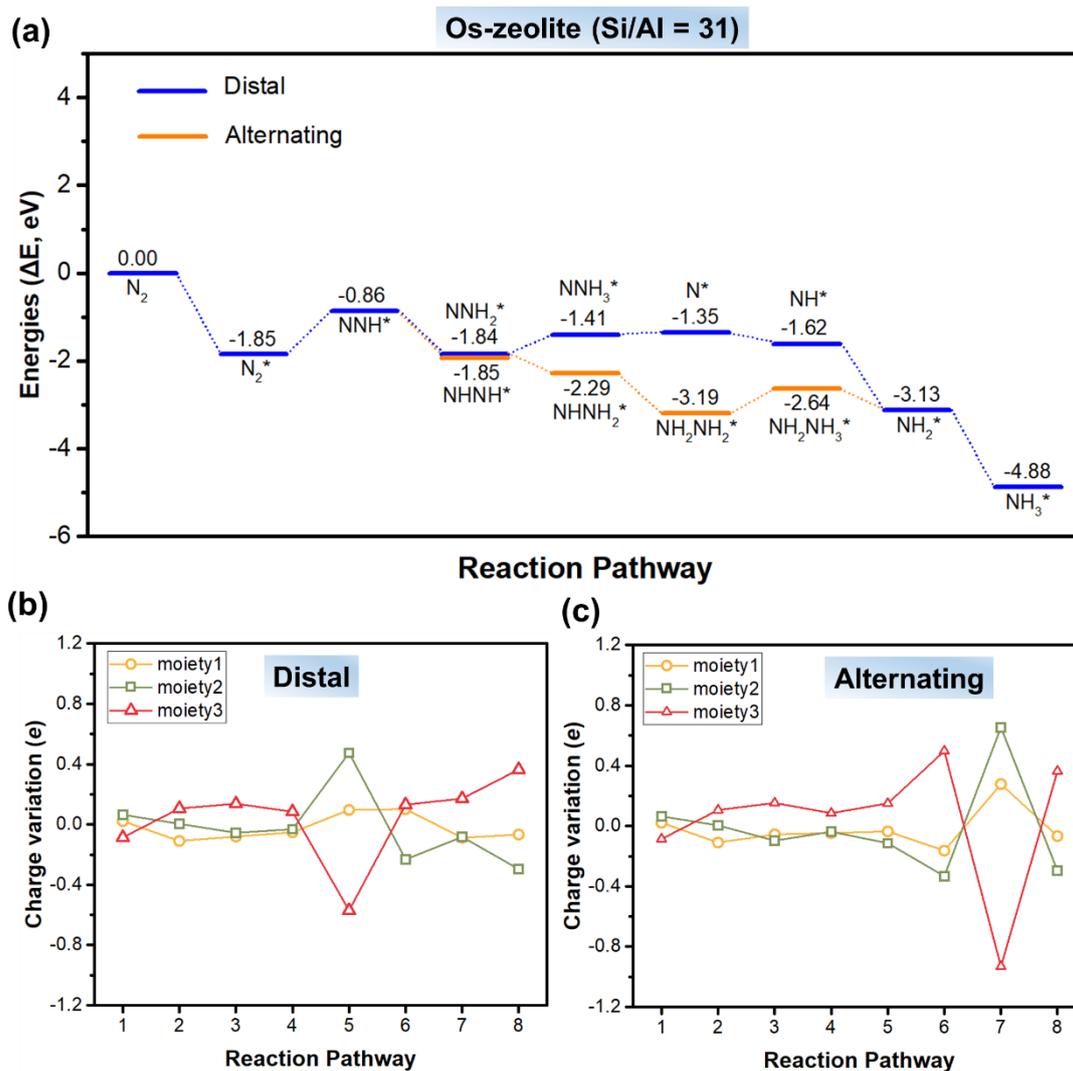


Fig. S23 (a) Energy diagrams for NRR in Os-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

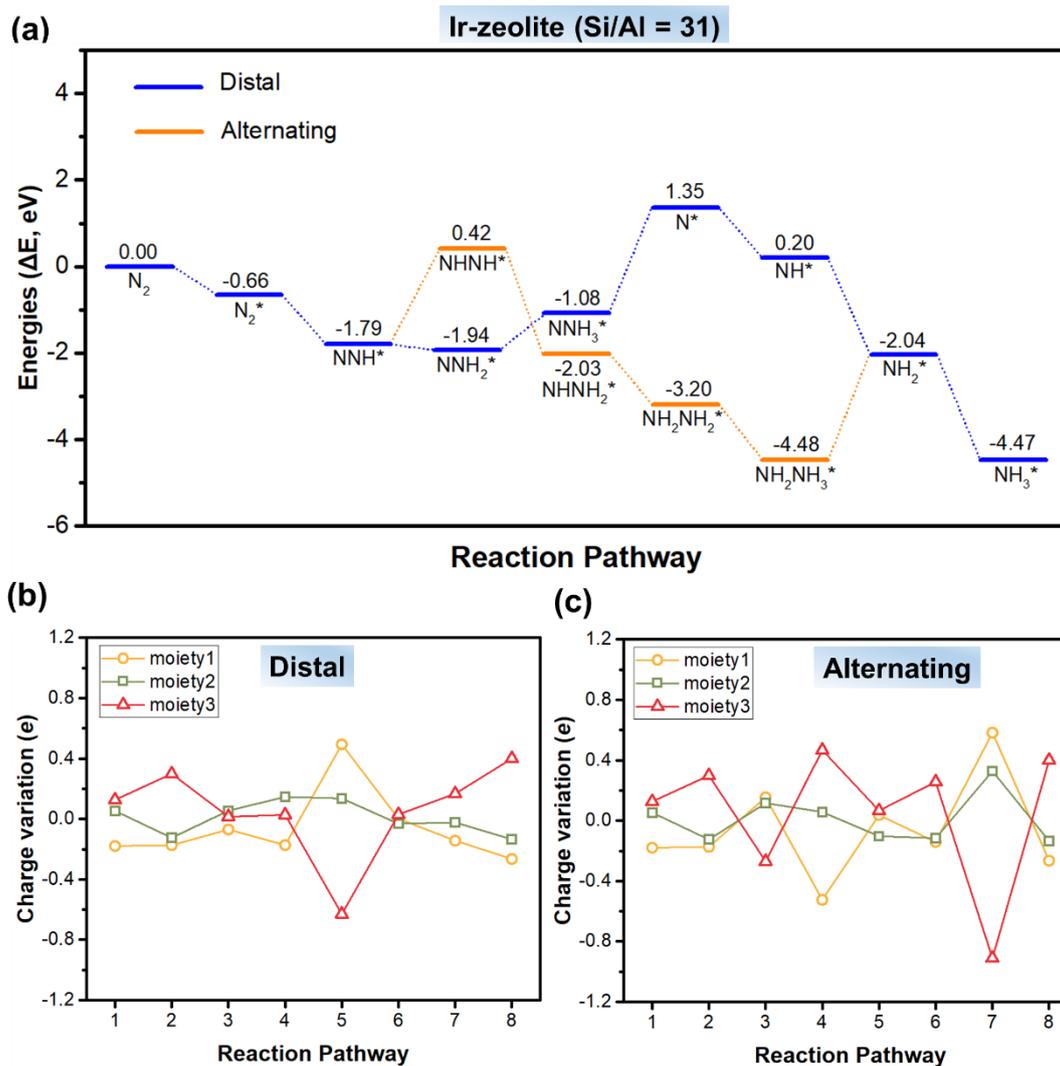


Fig. S24 (a) Energy diagrams for NRR in Ir-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

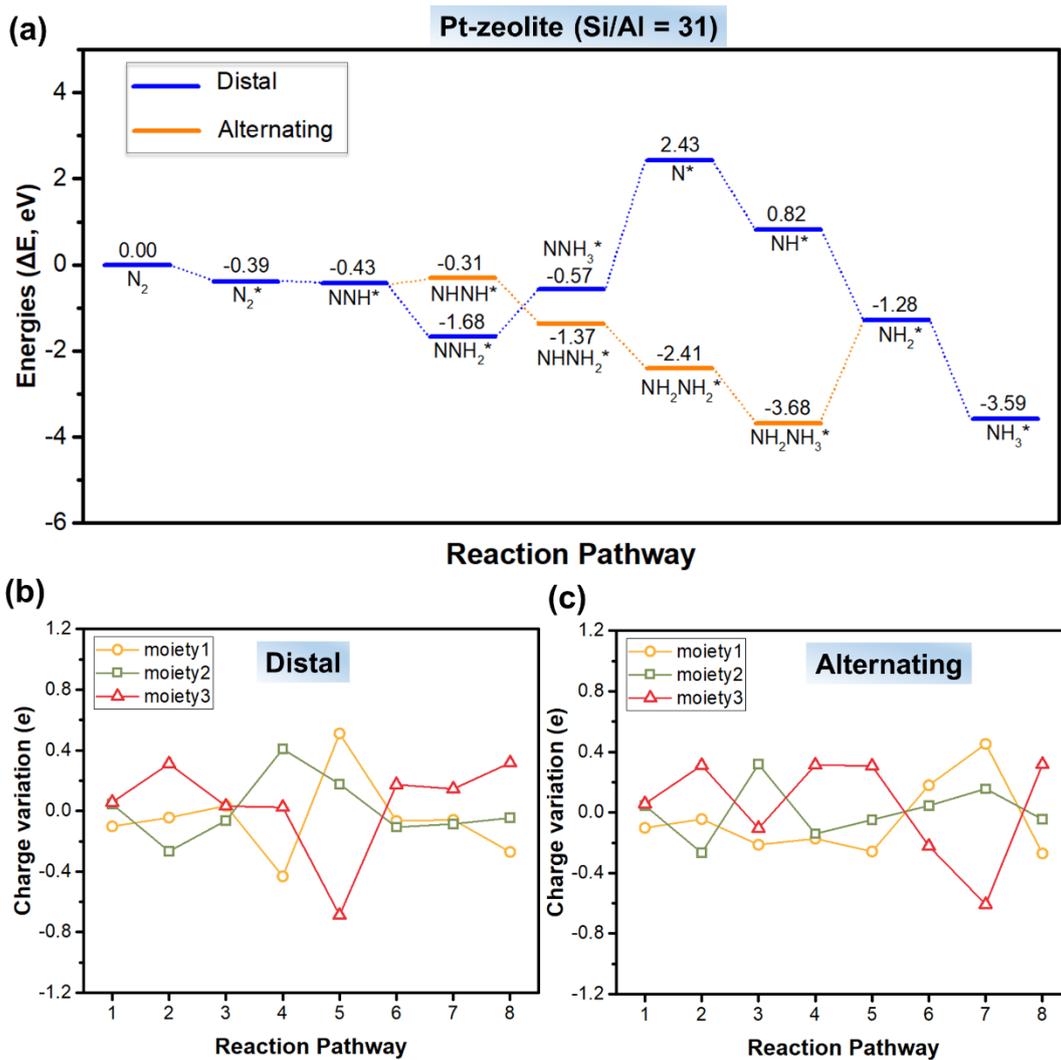


Fig. S25 (a) Energy diagrams for NRR in Pt-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

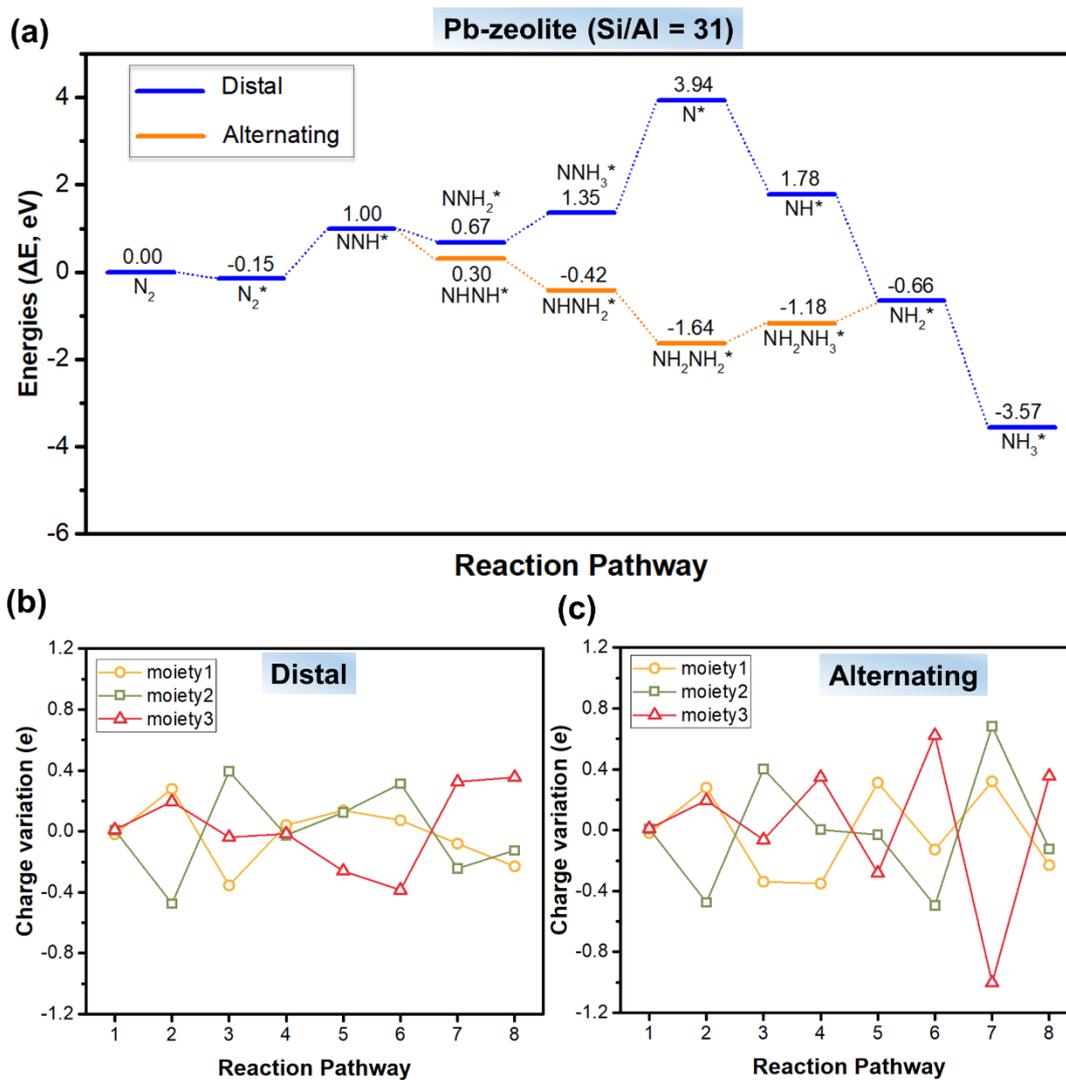


Fig. S26 (a) Energy diagrams for NRR in Pb-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

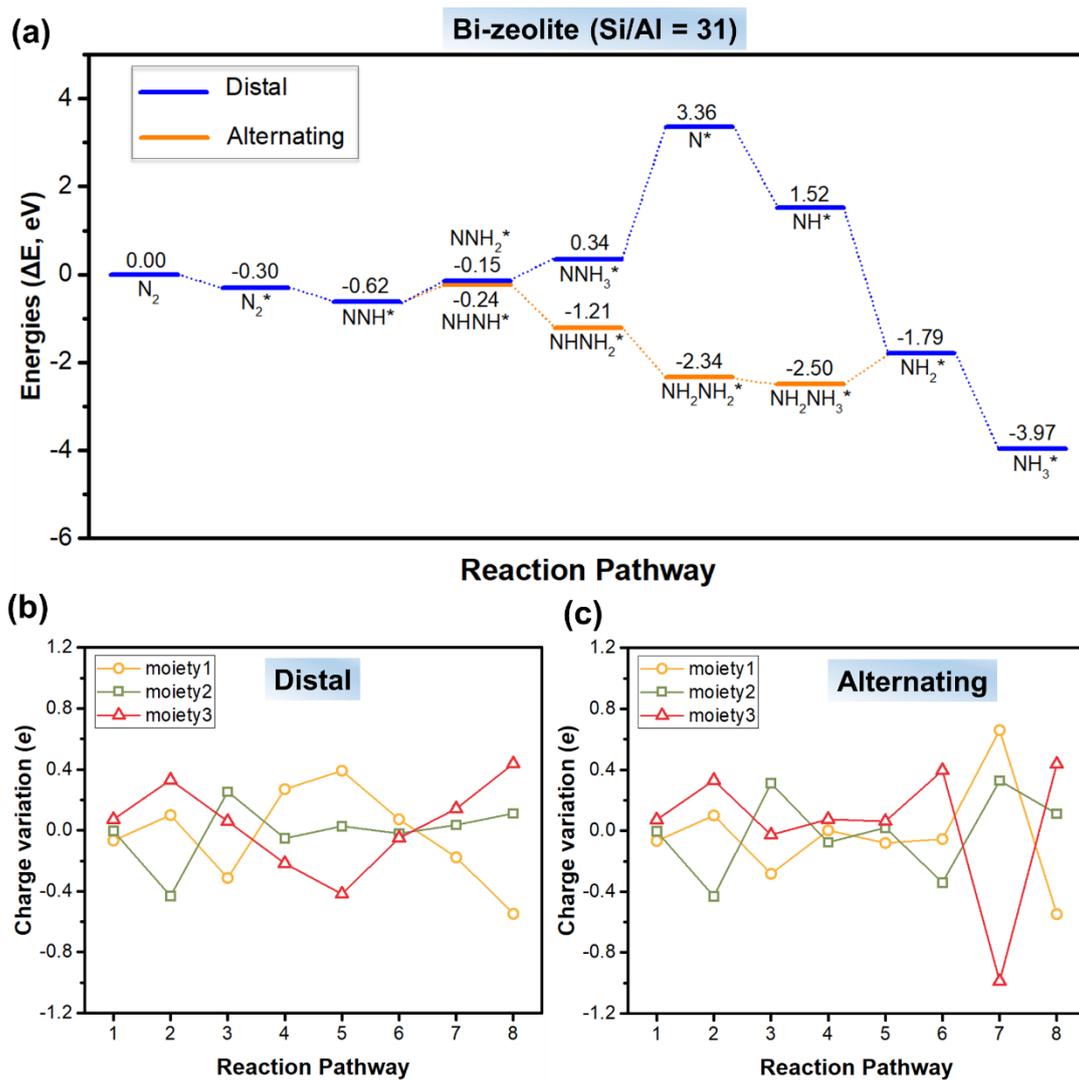


Fig. S27 (a) Energy diagrams for NRR in Bi-zeolite (Si/Al = 31) via distal and alternating pathways; charge variation of three moieties along the (b) distal and (c) alternating pathways of NRR.

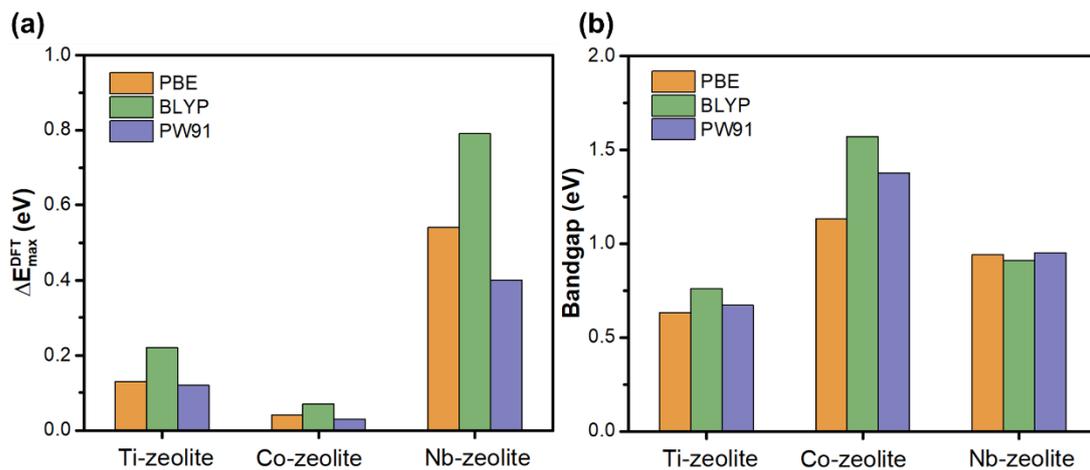


Fig. S28 (a) The calculated energy changes of rate limiting steps and (b) bandgap of metal-zeolites by using different DFT functionals.

Table S3. The calculated energy changes of rate limiting steps and bandgap of metal-zeolites by using different DFT functionals.

metal-zeolites	ΔE_{\max} (eV)			E_g (eV)		
	PBE	BLYP	PW91	PBE	BLYP	PW91
Ti-zeolite	0.13	0.22	0.12	0.63	0.76	0.67
Co-zeolite	0.04	0.07	0.03	1.13	1.57	1.37
Nb-zeolite	0.54	0.79	0.40	0.94	0.91	0.95

S3. Multilevel attention graph convolutional neural network

DeepMoleNet based on atom and atom pair as node and edge inputs could describe the molecular information to give the prediction of the catalytic systems automatically. These inputs are composed of real number vectors, one hot encoding related to atom and atom pair information. The atom nodes inputs are atomic type, atomic number, van der Waals radius, and atom node degree. The bond type, Gaussian expanded distance are for the description of the edge features. The hyper-parameters of DeepMoleNet are listed in **Table S4**, and ACSFs are calculated by Dscribe package.

12

The intermediate structures including 492 cluster models were employed in the graph convolution neural network, where 393 data were chosen as the training set, 49 for validation set, and the rest for the test set. Large-sized systems would be difficult to predict due to the large computation complexity, which may limited its applications. The local environment around active sites would make more contributions to the NRR than the long-ranged frameworks according to the chemical intuition.

Table S4. Hyper-parameters of DeepMoleNet

Category	Hyper-parameter	Value
Message Passing Function	Message passing steps T	6
	Edge network layers	3
	Edge network hidden dim	256
Update Function	Node hidden units dim	256
	Attention layer	2
Readout Function	Atom-wise NN	2
	Output NN hidden layer	3
	Output NN hidden units	256
Auxiliary Target/ACSFs	Radial functions	[1, 1], [1, 2], [1, 3]
	Angular functions	[1, 1, 1], [1, 2, 1], [1, 1, -1], [1, 2, -1]
Training	Initial learning rate	2×10^{-4}
	Scheduler	Cosine Annealing LR
	Optimizer	Adam
	Batch size	64
	Training epochs	5000

Table S5. R^2 of GCNN models on training, validation, and test sets.

Models	Training set	Validation set	Test set
Model A	0.91	0.90	0.88
Model B	0.86	0.87	0.87
Model C	0.90	0.92	0.90

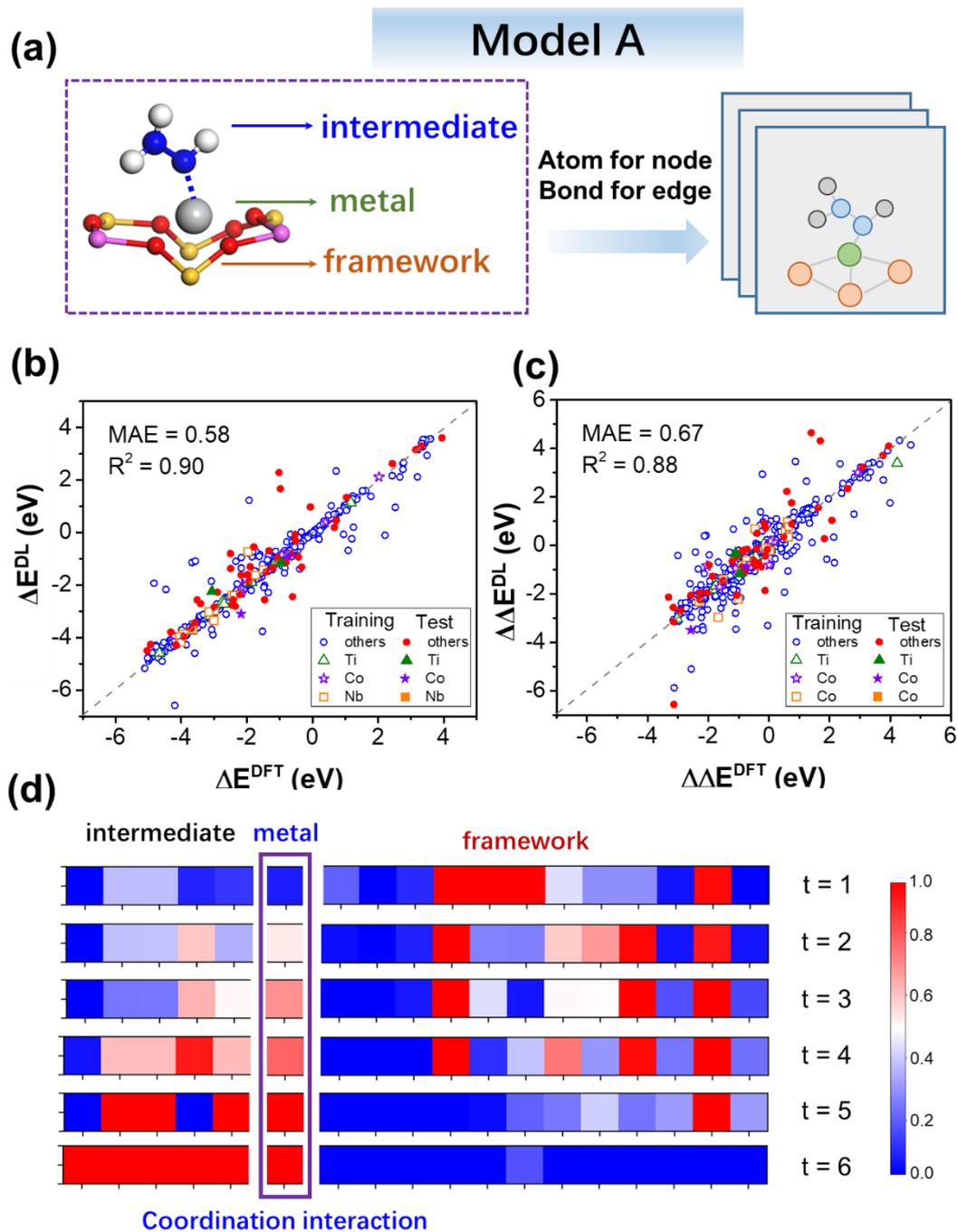


Fig. S29 (a) The structure of model A, including intermediate, metal and framework moieties; the prediction performance of (b) ΔE and (c) $\Delta\Delta E$; (d) illustration of the attention in message passing of model A. The color indicates the relative weight, in which red and blue colors mean the weight of 1 and 0, respectively.

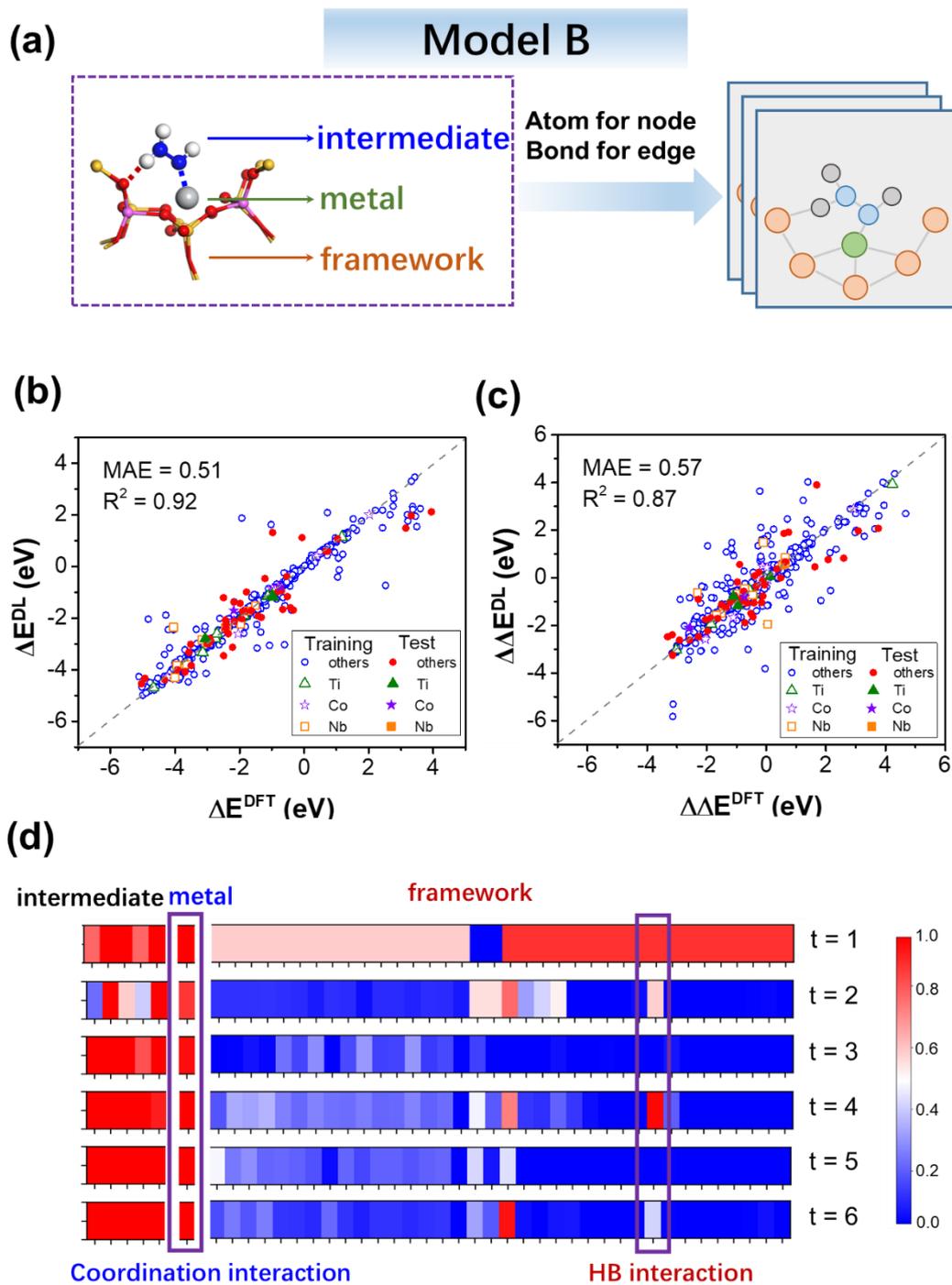


Fig. S30 (a) The structure of model B, including intermediate, metal and framework moieties; the prediction performance of (b) ΔE and (c) $\Delta\Delta E$; (d) illustration of the attention in message passing of model B. The color indicates the relative weight, in which red and blue colors mean the weight of 1 and 0, respectively.

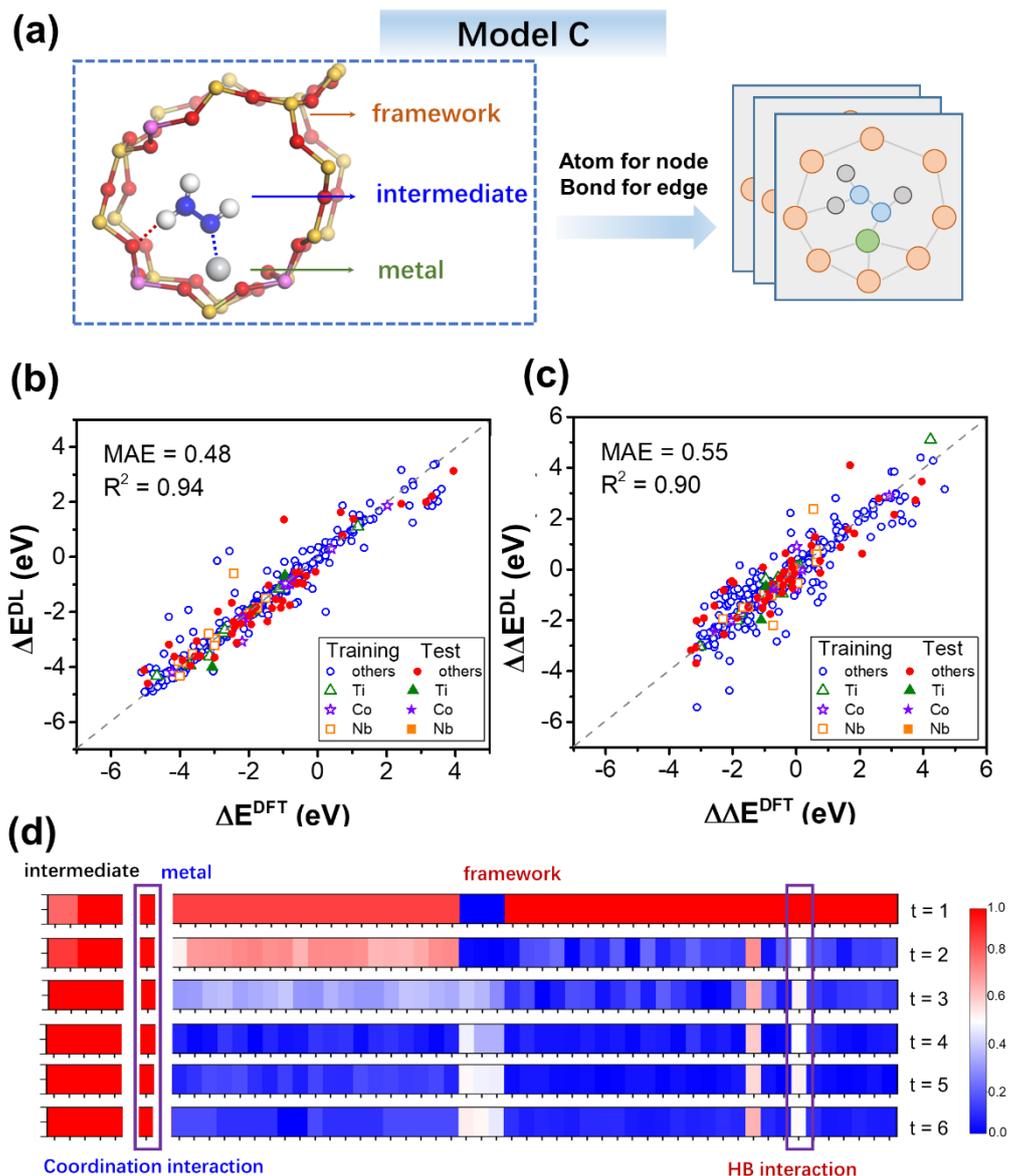


Fig. S31 (a) The structure of model C, including intermediate, metal and framework moieties; the prediction performance of (b) ΔE and (c) $\Delta \Delta E$; (d) illustration of the attention in message passing of model C. The color indicates the relative weight, in which red and blue colors mean the weight of 1 and 0, respectively.

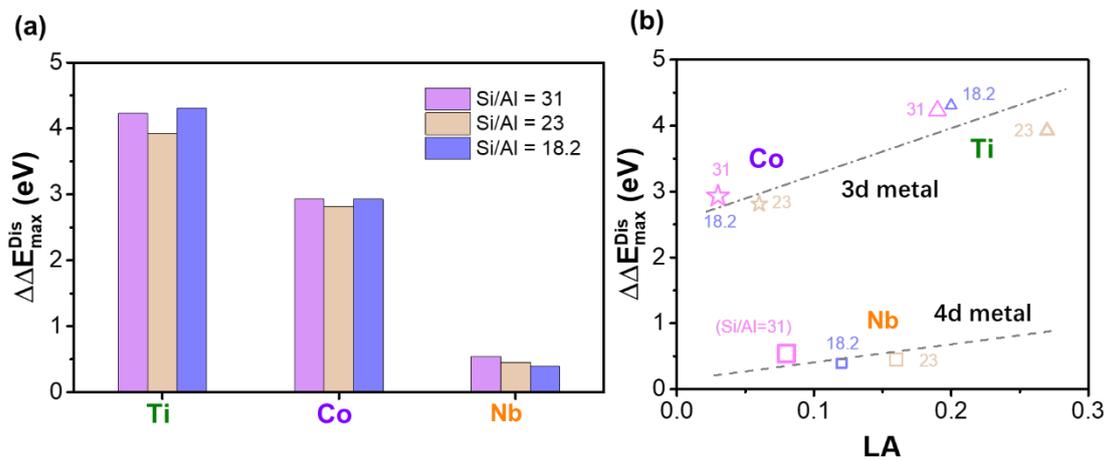


Fig. S32 (a) The values of $\Delta\Delta E_{\max}$ of metal-zeolites (Ti, Co, and Nb) along distal pathway with different Si/Al ratios (Si/Al = 31, 23, 18.2); (b) correlation between the $\Delta\Delta E_{\max}$ and LA descriptor.

S4. Feature selection

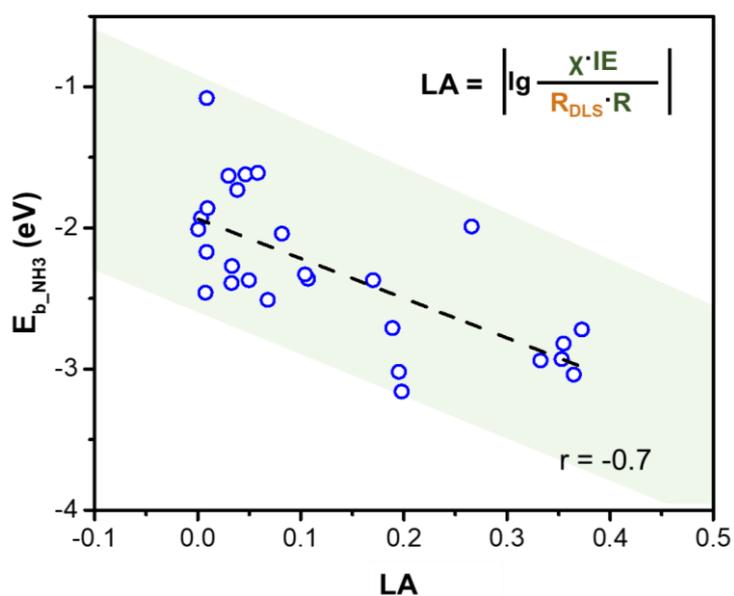


Fig. S33 The relationship between LA and binding energy of NH₃ in metal-zeolites.

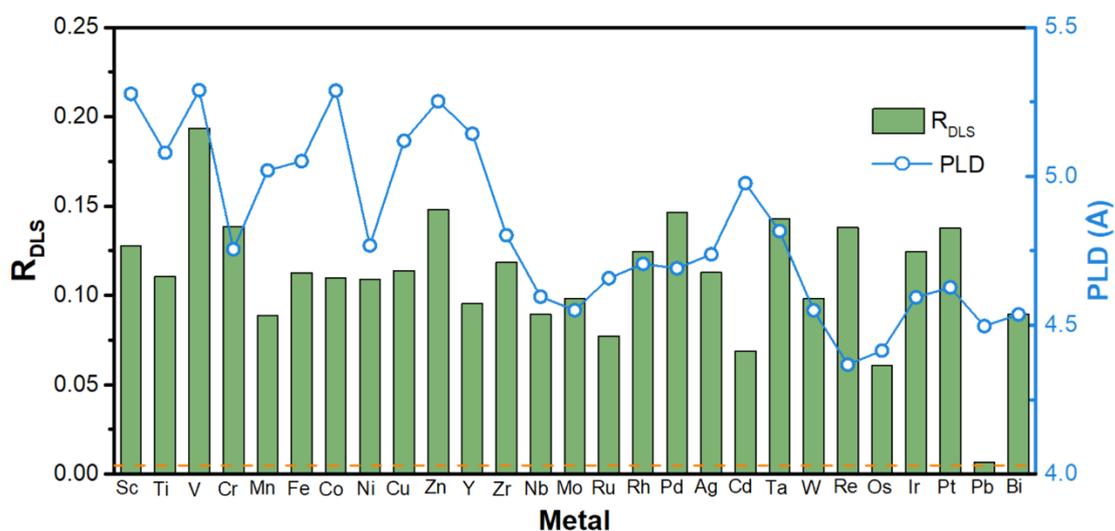


Fig. S34 R_{DLS} and pore largest diameter (PLD) of different metal-zeolites; for comparison, the R_{DLS} of siliceous zeolite is shown as a horizontal dashed line.

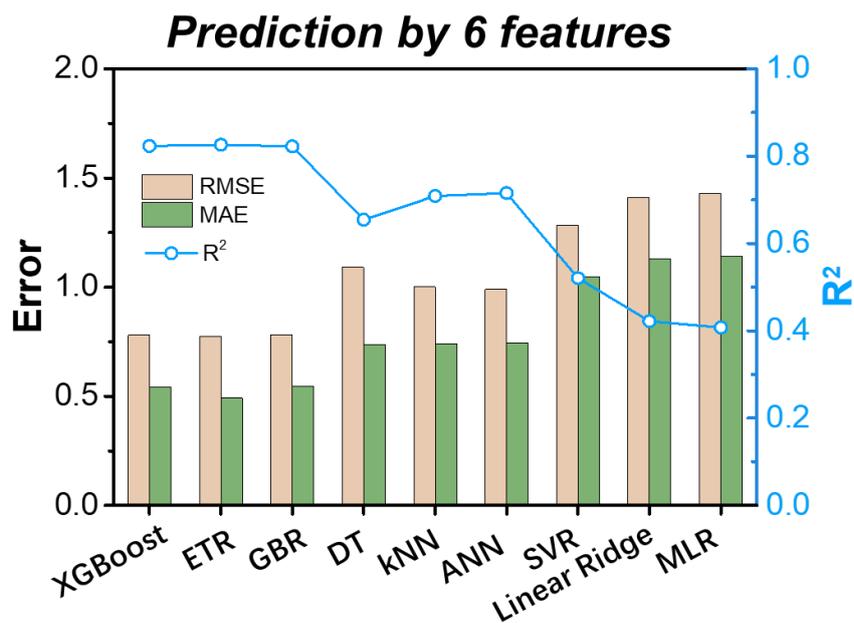


Fig. S35 Comparison of the performance of different algorithms of the explainable machine learning according to the root-mean-squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2) without PLD descriptors.

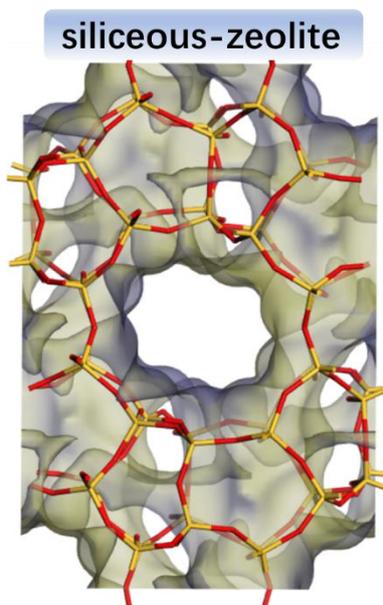
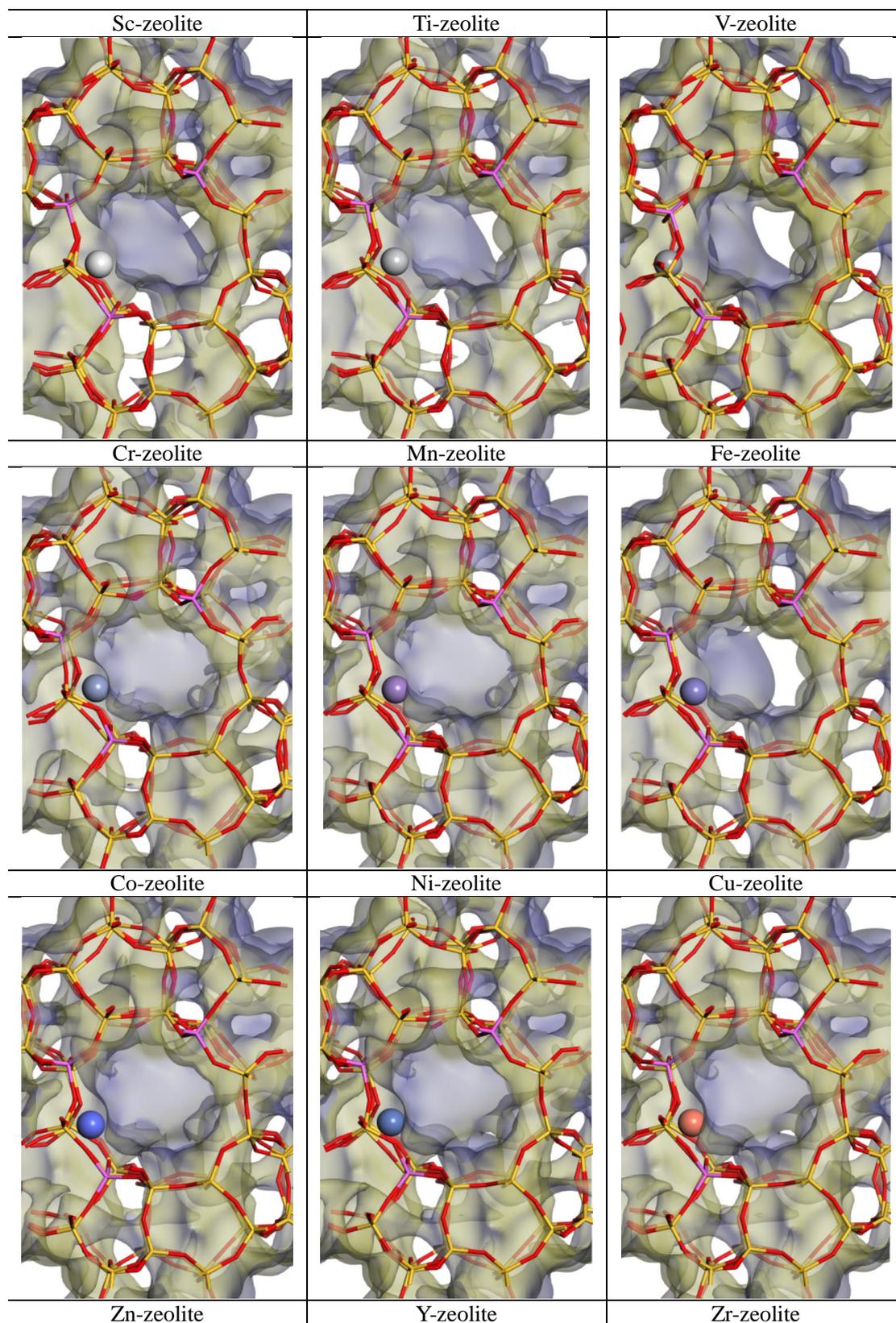
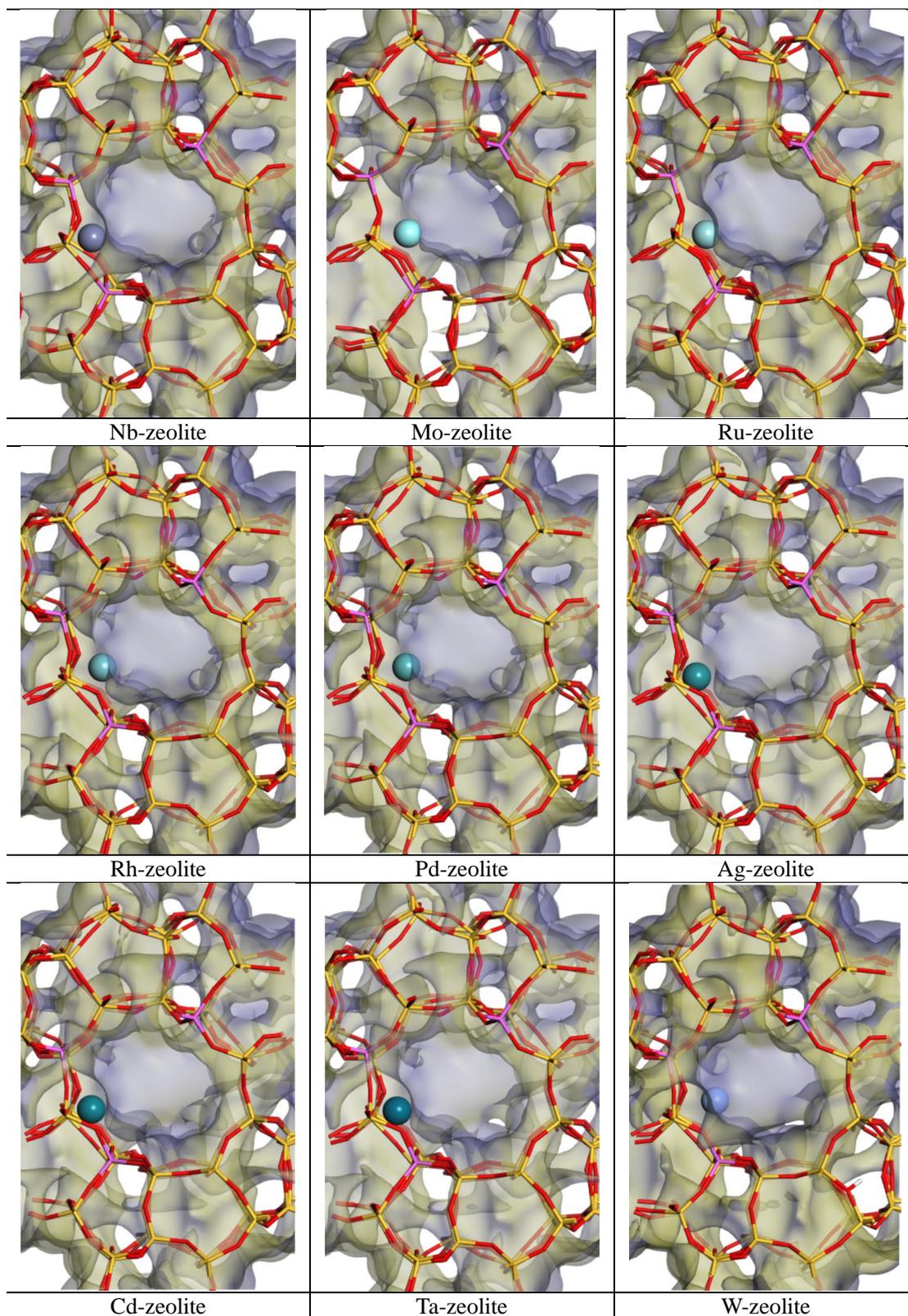
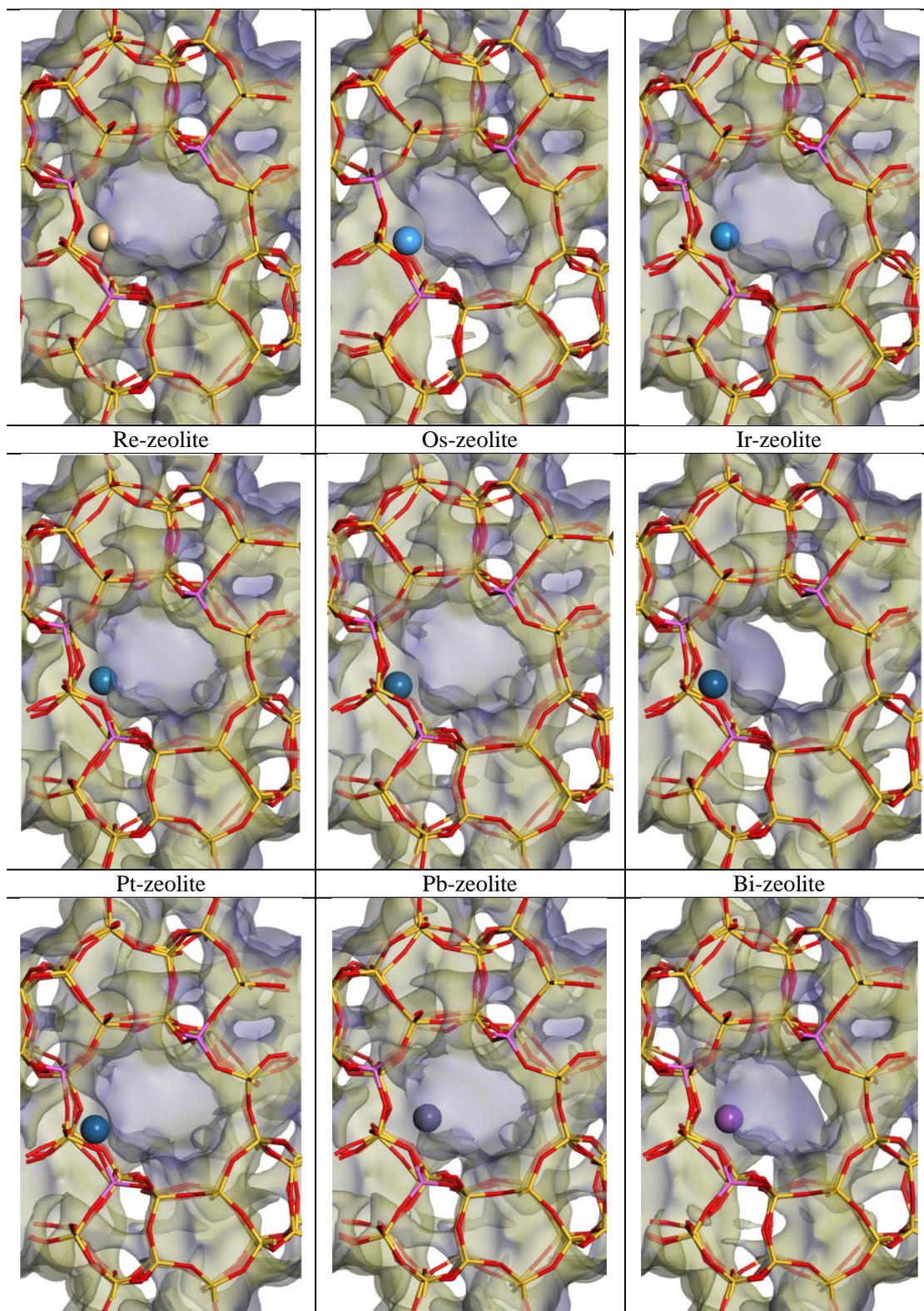


Fig. S36 The electrostatic potential of the siliceous zeolite.

Table S6. The electrostatic potential of metal-zeolites (metal = Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Y, Zr, Nb, Mo, Ru, Rh, Pd, Ag, Cd, Ta, W, Re, Os, Ir, Pt, Pb, and Bi).







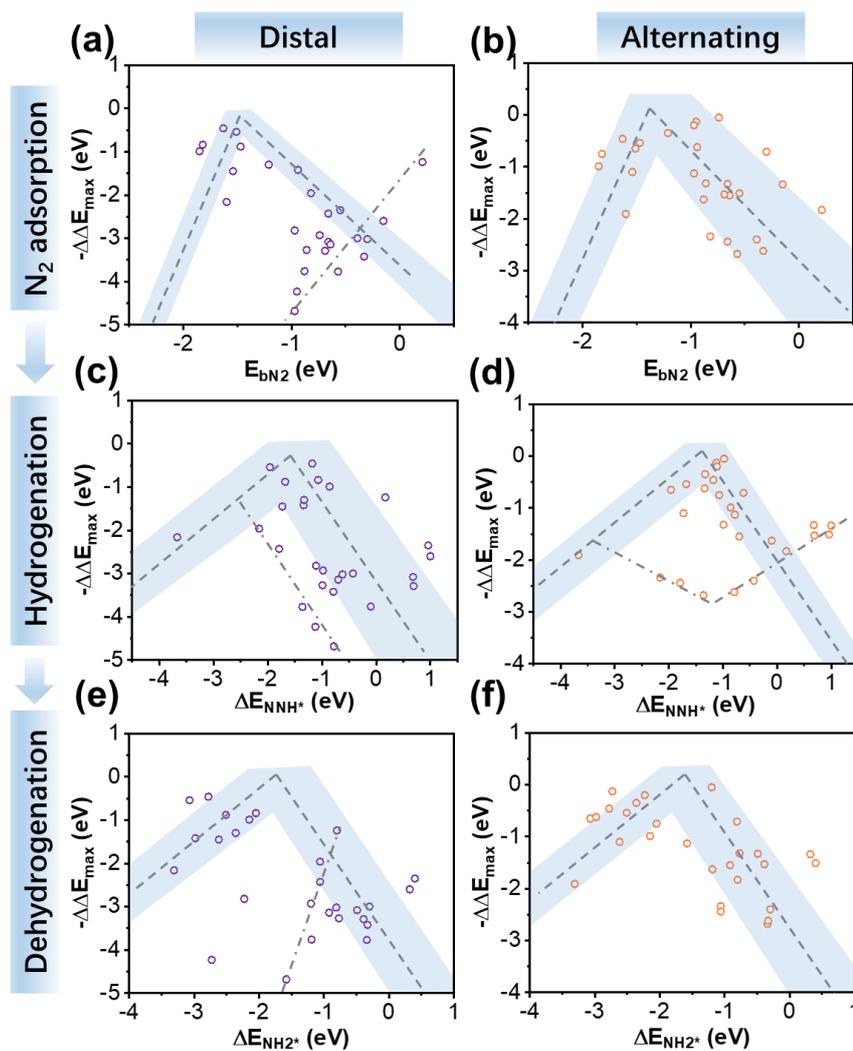


Fig. S37 Volcano-shaped relationship between the $\Delta\Delta E_{\max}$ and adsorption energy of (a-b) N₂ molecule, (c-d) NNH*, and (e-f) NH₂* intermediates along the distal and alternating pathways, respectively.

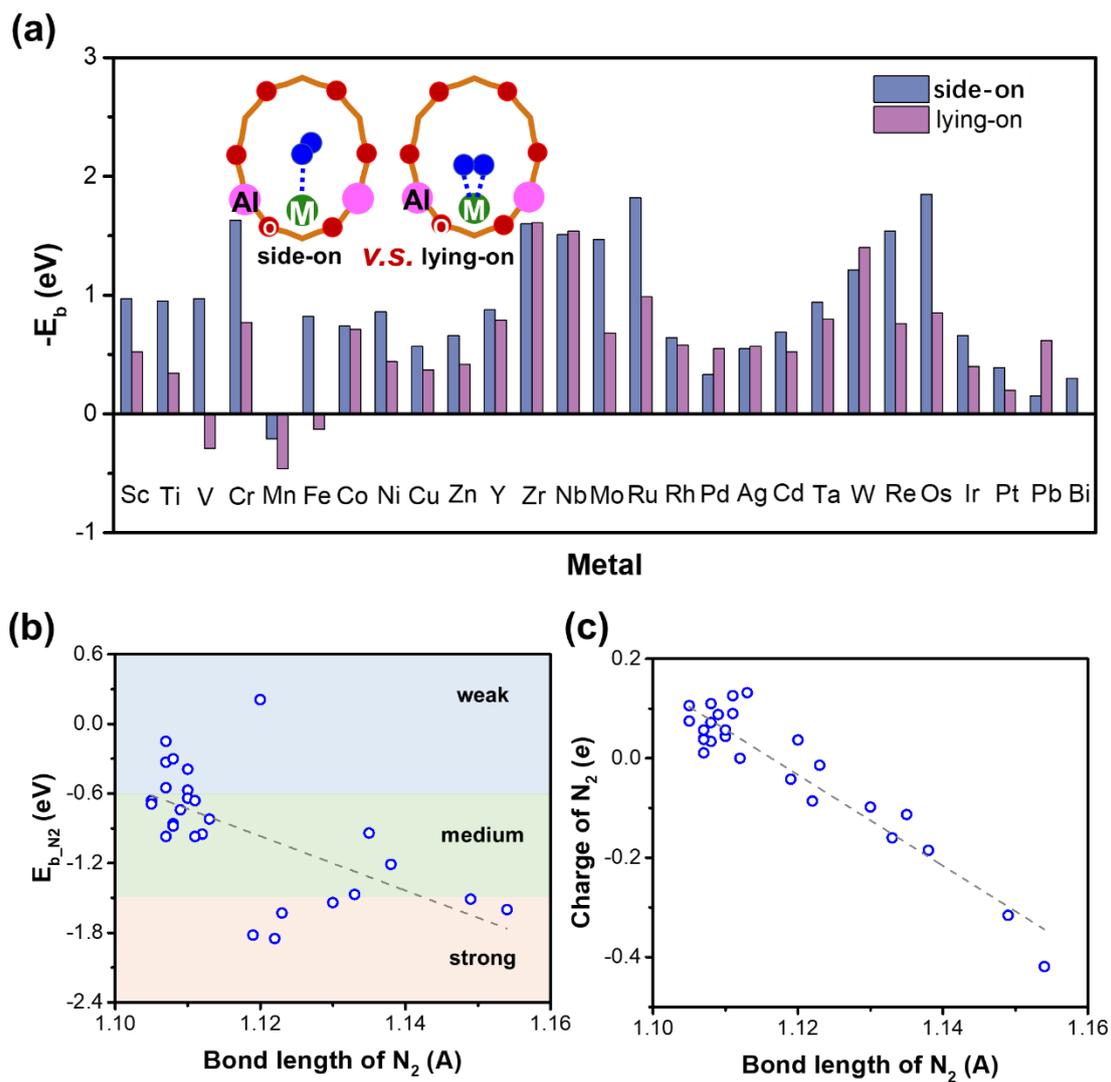


Fig. S38 (a) Side-on and lying-on adsorption modes of N_2 molecule in metal-zeolites; (b) the relationship between the N_2 binding energies and bond lengths of N_2 molecules; (c) the relationship between charges and bond lengths of N_2 molecules.

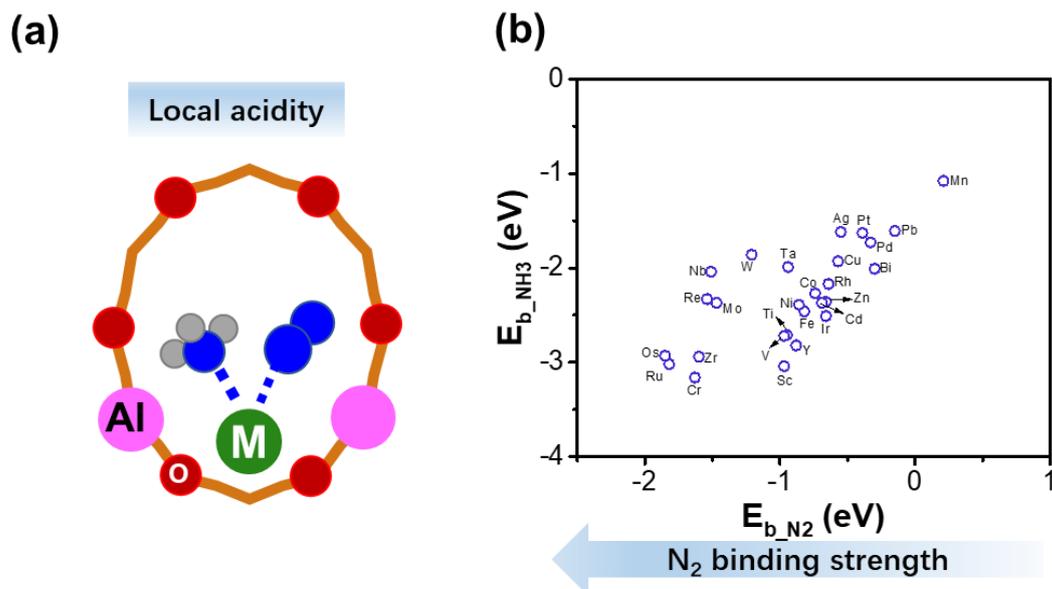
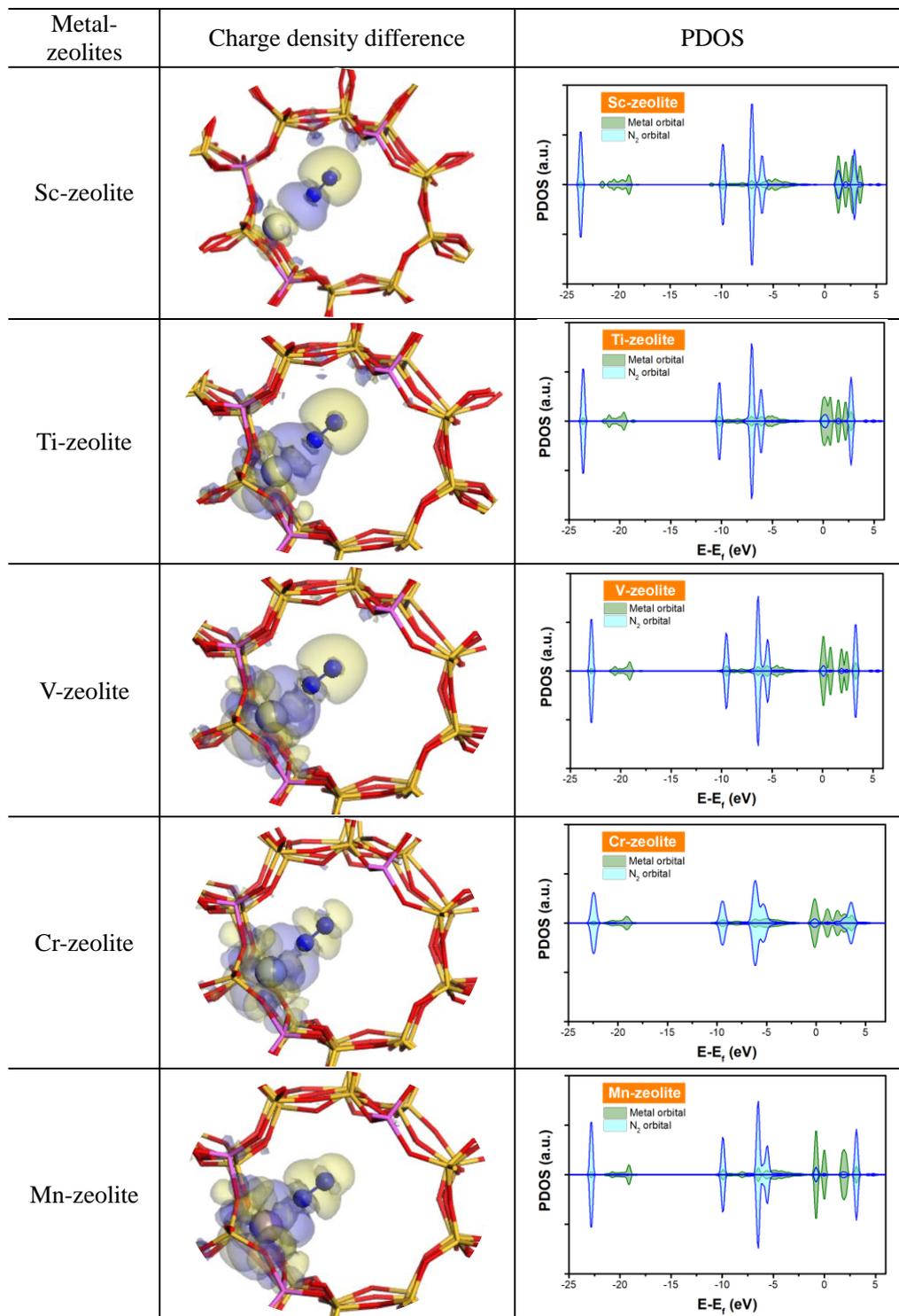
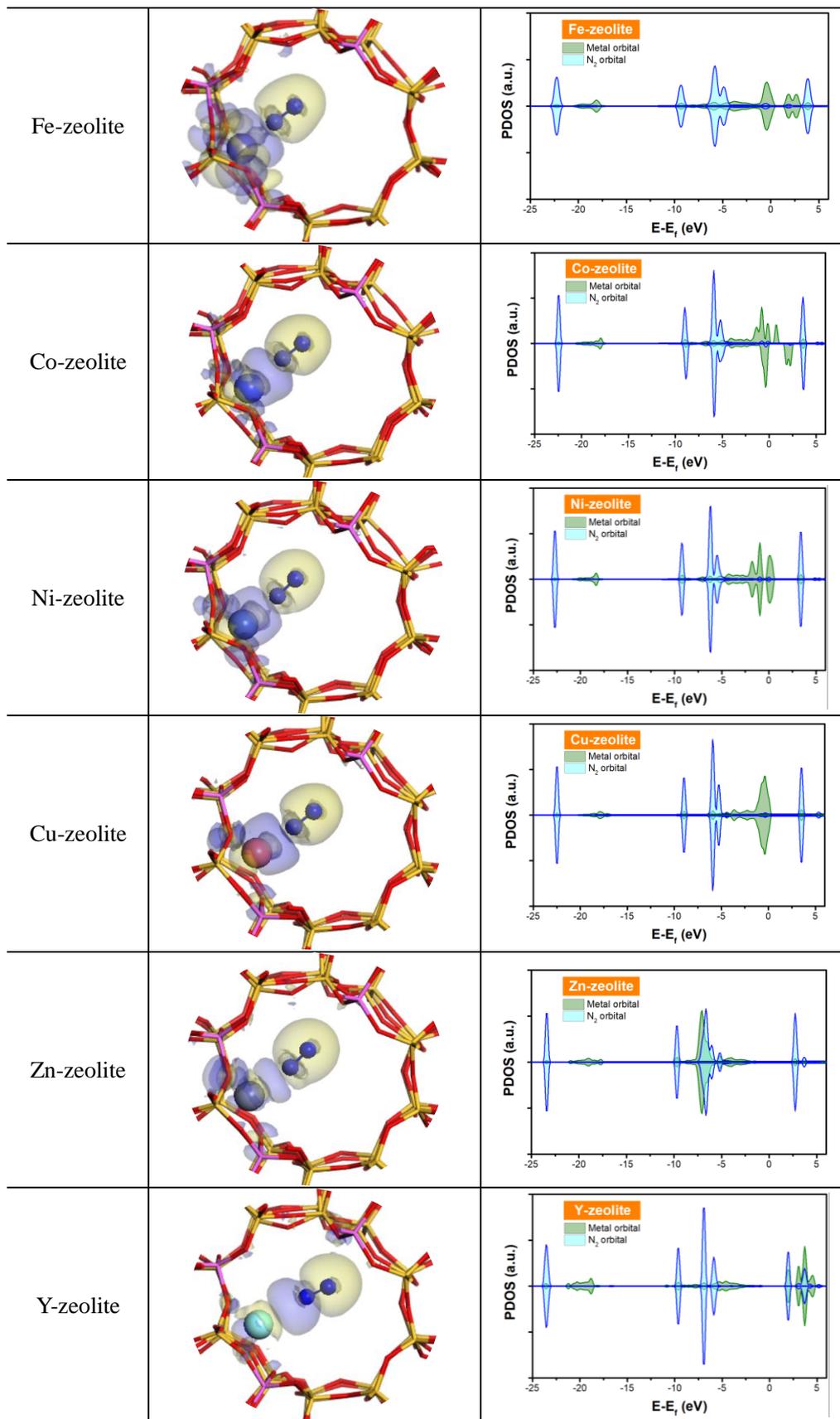
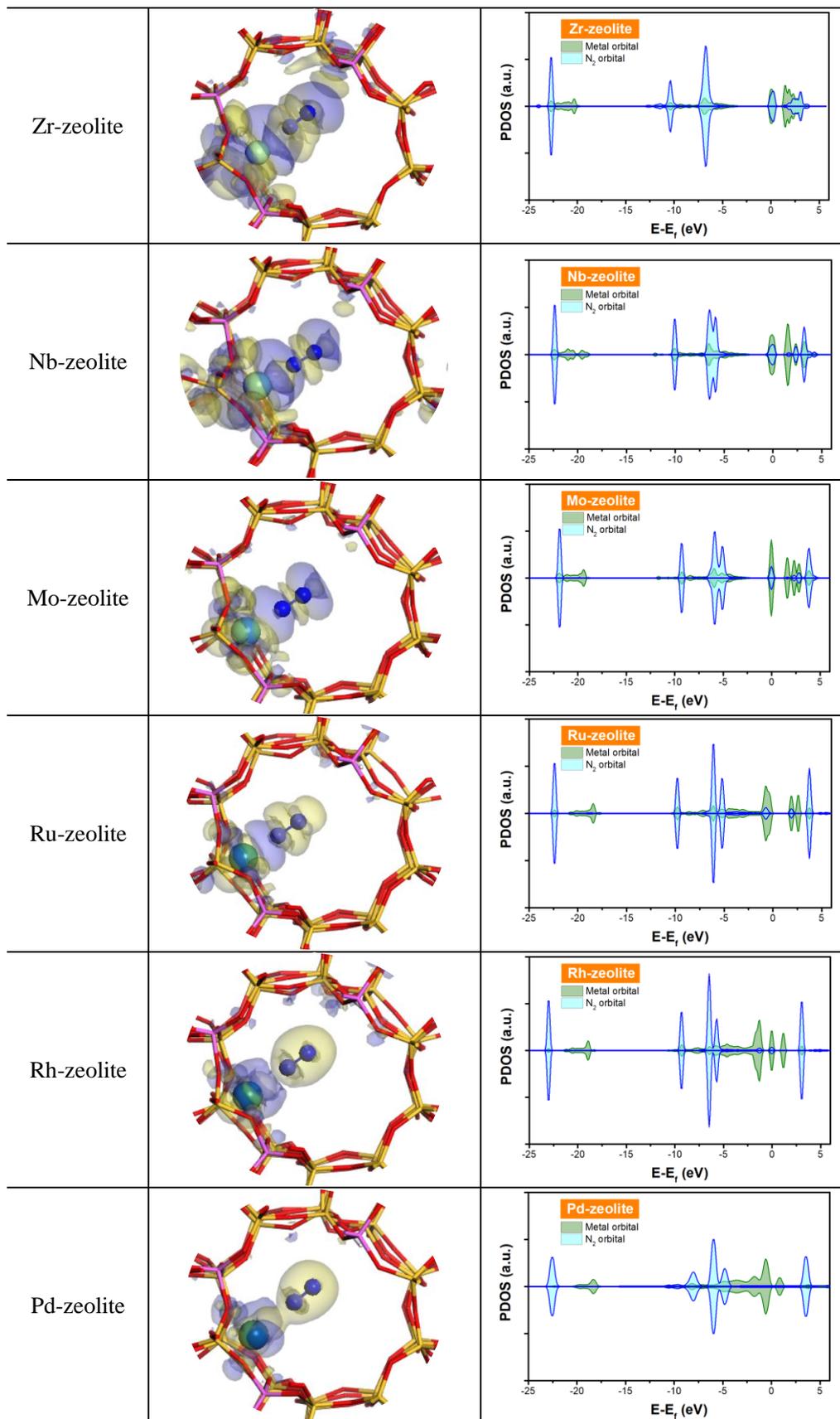


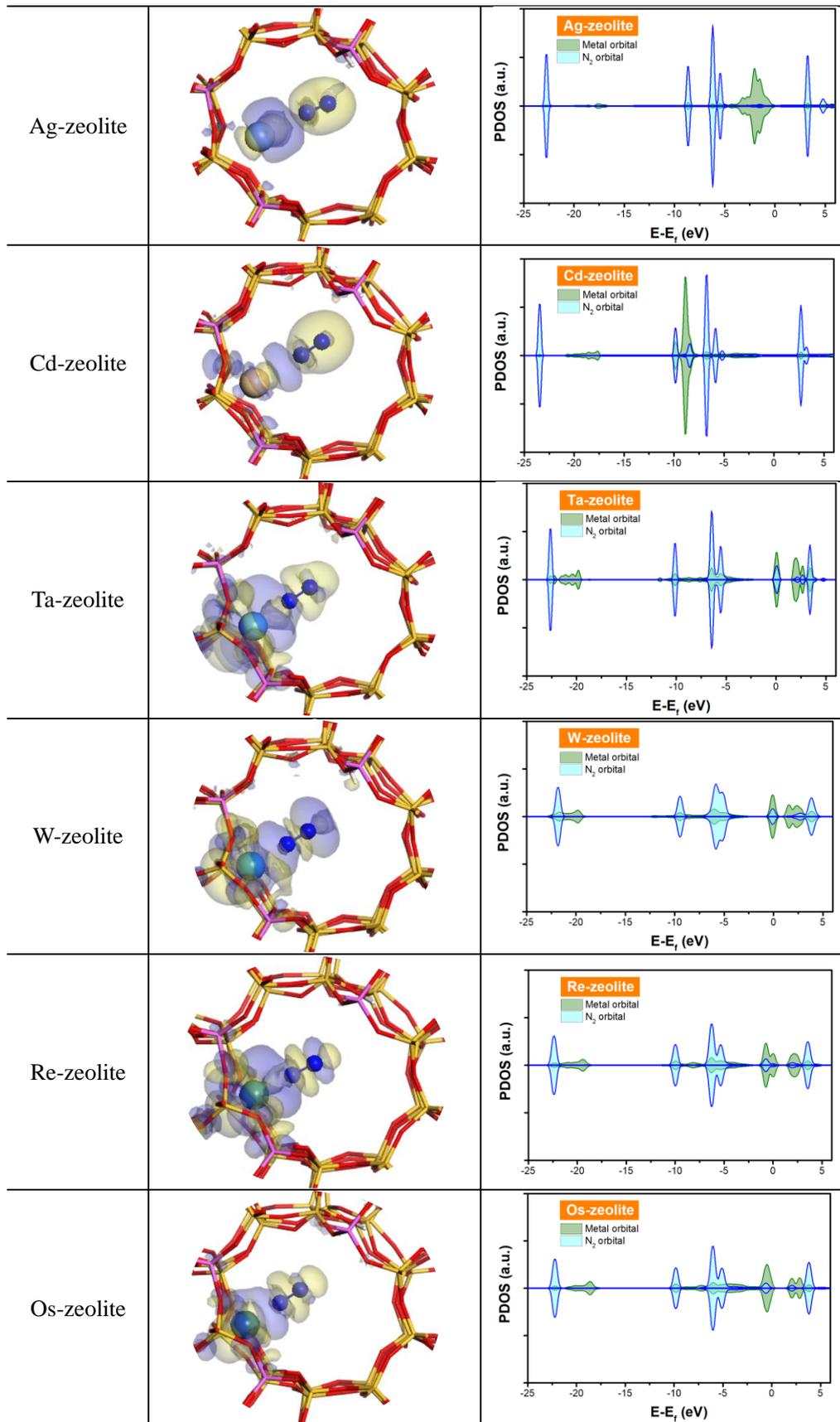
Fig. S39 (a) The index of local acidity from metal and framework; (b) the relationship between the local acidity and the binding strength of N_2 .

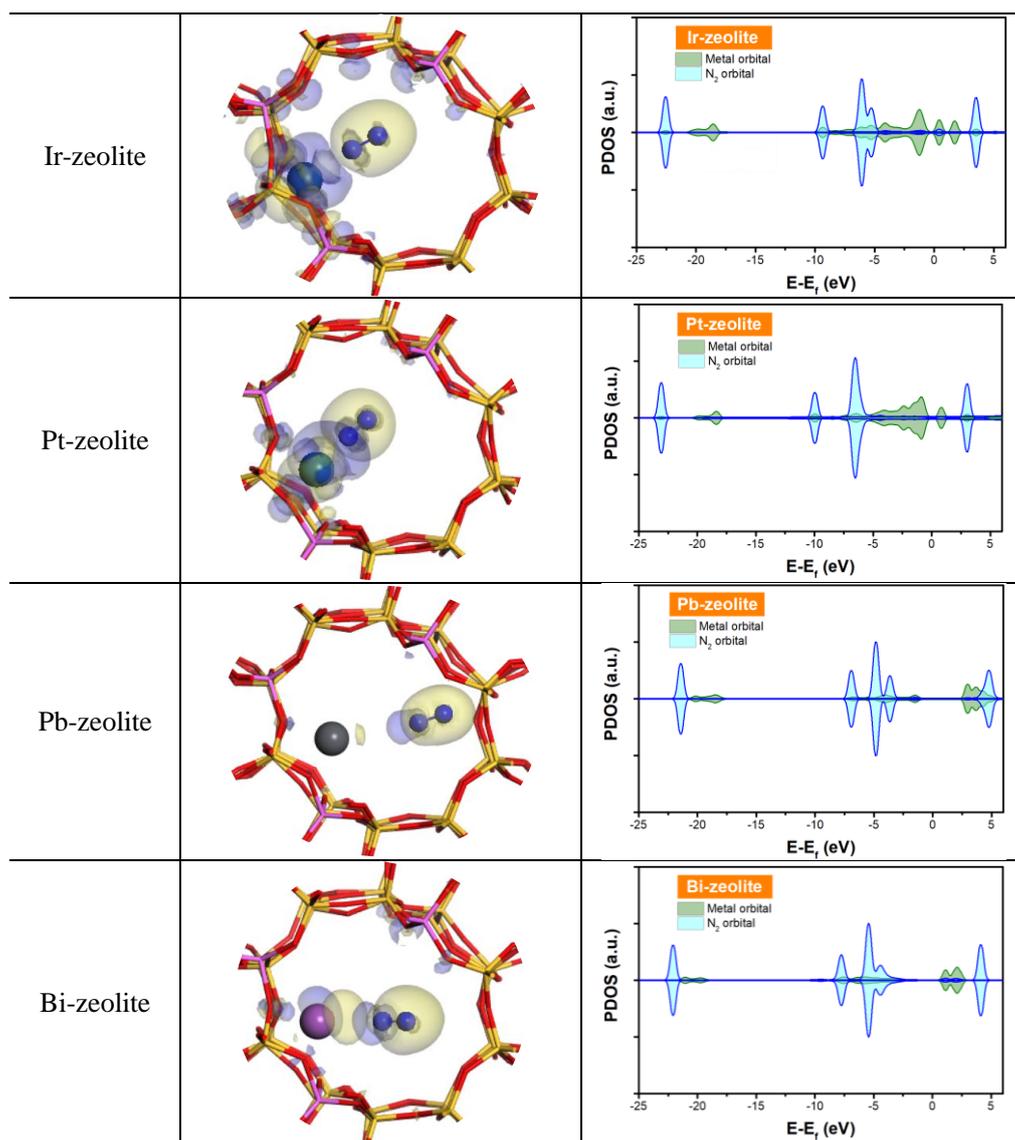
Table S7. Charge density differences and PDOS of N₂ adsorbed in metal-zeolites (metal = Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Y, Zr, Nb, Mo, Ru, Rh, Pd, Ag, Cd, Ta, W, Re, Os, Ir, Pt, Pb, and Bi).











The charge depletion and accumulation were colored by blue and yellow with isosurface of $0.003 \text{ e}/\text{\AA}^3$, respectively. Projected density of states (PDOS) of N_2 molecule in corresponding metal-zeolites, where the orbitals of metal center and N_2 molecule are depicted by green and blue colors, respectively.

For further insight on the effect of metal sites on NRR process, the plots of differential charge density of N_2 molecule adsorption and projected density of states (PDOS) in different metal-zeolites were shown in **Table S7**. The charge transfer could occur between N_2 and metal-zeolites to strengthen the adsorption ability. With the

adsorption of the N_2 molecule on the metal center, the orbitals of nitrogen shift significantly. The unoccupied orbitals of metal centers could accept electrons from the 2π and 3σ molecular orbitals of nitrogen molecules, which promoted the N_2 adsorption ability. Moreover, the occupied orbitals of metal sites back-donate electrons to the $2\pi^*$ orbital of nitrogen molecules, resulting in the elongation of $N\equiv N$ bond length. The $2\pi^*$ orbital shifted to the Fermi level to activate N_2 for hydrogenation in the following NRR steps.

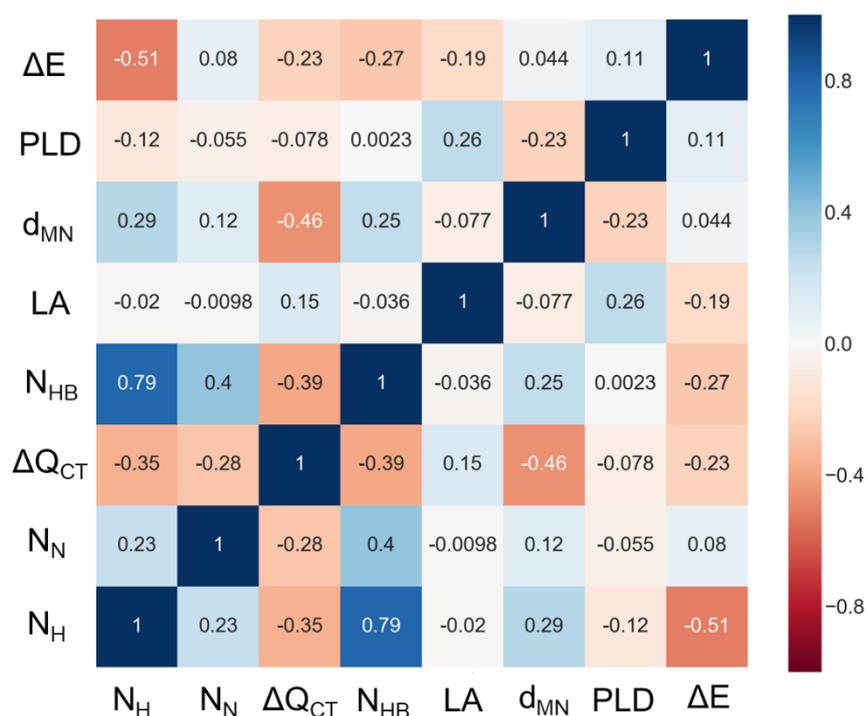


Fig. S40 Pearson correlation coefficient matrix of 7 features.

S5. Explainable machine learning models

S5.1 7-feature explainable machine learning models

Table S8. Parameters of the explainable machine learning models for predicting the pathway probability and relative energy

Parameter	Pathway probability	Relative energy
	prediction	prediction
learning_rate	0.25	0.115
max_depth	6	13
n_estimators	40	480
min_child_weight	1	8
subsamlpe	0.8	0.84
colsample_bytree	1	0.98
reg_alpha	0.35	0.9
reg_lambda	0.15	0.3

Table S9. The terms and units in SISO features

	terms	unit
SISO1	$\frac{N_H}{N_N}$	
	$PLD * \Delta Q_{CT}$	$\text{\AA} e$
SISO2	d_{MN}^2	\AA^2
	ΔQ_{CT}	e
	$e^{N_N} + e^{N_H}$	
SISO3	$\frac{\lg N_N}{N_N} N_{HB}$	
	PLD	\AA
	LA	

S5.2 A multiparameter explainable machine learning model

We choose 21 initial features to predict the relative energy (ΔE) during overall NRR in metal-zeolites, which can be classified into four groups, including metal site, coordination, framework, and pathway features, respectively. Pearson correlation coefficient matrix of initial features was shown in **Figure S41**.

Metal site features. The introduced metal atoms could be active sites for nitrogen reduction reaction because N_2 molecules would be activated by the “push-pull” hypothesis upon the adsorption in metal-zeolites. It is significant to select the descriptors of active sites, including electro-negativity of metal atoms (χ), electron affinity of metal atoms (EA), first ionization energy of metal atoms (IE), the number of d electrons of metal atoms (N_d), atomic number of metal atoms (Z), the molar mass of the metal atoms (M), atomic radius of metal atoms (R), and the Mulliken charge of the metal atoms before (Q) and during (Q_r) the NRR process, respectively.

Three features, including χ , EA, and IE, can describe the ability of acceptance and donation of electrons. For example, the more electronegative an element is, the more capable the atoms are of attracting electrons in the compound. Electron affinity (EA) is the energy released by a gaseous atom in its ground state to gain an electron to become a gaseous anion. And the first ionization energy (IE) of metal atom is the energy required for a gaseous atom in its ground state to lose an electron in its outermost shell. The lower the first ionization energy, the easier it is for an atom to

lose an electron and vice versa. To reveal the importance of the interaction between metal centers and the intermediates during the NRR process, N_d is selected as the descriptor to stand the number of valence electrons of metal atoms in the d orbitals. Three features (Z, M, and R) could describe the size of the metal center. Z is equal to the number of protons in nucleus, indicating the position of the element in the periodic table of the elements. M is the molar mass of the corresponding metal atoms in the zeolites and R is the atomic radius of metal atoms. To avoid the redundancy of the features, R is retained according to the Pearson correlation analysis (**Fig. S41**) in the 14-feature scheme. The charge transfer between the metal centers and intermediates plays an important role in the NRR process, so we take the Mulliken charge of metal (Q and Q_r standing for the charge before and during the NRR process, respectively) into account for the prediction.

Coordination features. The variation of coordination environment could adjust the chemical reactivity in reaction, which can be the local descriptors for the reaction prediction.¹³ The geometric distances between metal active sites and the nearest atoms were selected as the coordination environment descriptors. We found that the distance between metal atom and N atom of intermediates (d_{MN}) was shortest among the coordination descriptors, indicating the strongest coordination ability. The metal atom could be anchored in the framework by coordination with four O atoms in the 6-MR around the channel, which were denoted as d_{O1} , d_{O2} , d_{O3} , and d_{O4} with the increase of the length in turn. The d_{Si} was the nearest distance between the Si and the metal atom.

Framework features. The confinement of the metal atoms could be realized by the framework of zeolite, which can be immobilized by the small rings in the zeolites. Refinement distance least squares, R_{DLS} , is a local descriptor to describe the distortion of the framework in our previous work.¹⁴ It displayed the qualitative relationship between the local geometry distortion which has an influence of the adsorption of nitrogen molecules by changing the electrostatic environment. The larger distortion may give rise to a larger electrostatic interaction between the intermediates and the zeolites. Furthermore, the local electrostatic environment can also be tuned by the Si/Al ratio. To avoid the redundancy of the descriptors in the feature learning, we retain the R_{DLS} to give a description of the zeolite framework distortion.

Compared with the local descriptors focused on the specific catalytic sites, the global descriptors lay the emphasis on the whole channel structures and reaction pathways. The channel architecture of zeolites is important for the adsorption and diffusion properties in general. Here, the largest included sphere (PLD) was applied to describe the porosity of zeolites.

Pathway features. The formation of intermediates is complex during the NRR pathways. Different pathways of nitrogen reduction reaction have been studied over time, such as the Chatt type, including the distal and alternating reaction pathways.¹⁵⁻
¹⁹ Two possible associative NRR pathways, namely, distal and alternating are calculated to explore the performance of metal-zeolites as catalysts for the conversion of N_2 to NH_3 . In the distal pathway, hydrogen atoms will add to the nitrogen molecule

distal from the zeolite internal surface, firstly. After the distal NH_3 molecule has been released, the nitrogen atom close to the zeolite internal surface is reduced by hydrogen atoms attacking and desorbed as the second NH_3 molecule to finish the catalytic process. When following the alternating pathway, hydrogen atoms are added to the distal nitrogen atom firstly, and then to the proximal nitrogen atom alternately until the intermediate NH_2NH_2^* is formed. The first NH_3 molecule is released until the NH_2NH_2^* was hydrogenated. When the second NH_3 molecule is formed and desorbed, the alternating catalytic pathway is completed. We chose the pathway features to describe the distal pathway (pathway = 1) and alternating pathway (path = 2), total number of nitrogen atoms in intermediates (N_N), and number of hydrogen atoms in intermediates (N_H).

Although 21 initial features could give good prediction performance of the NRR process in metal-zeolites (MAE = 0.38 eV and $R^2 = 0.92$), as shown in **Fig. S42**, the number of the initial features seems to be redundant for the prediction for the NRR process as the optimal model. Can we reduce the number of the features to enhance the generalization ability of the prediction model? SISSO method can extract effective material descriptors from possibly correlated feature spaces in many applications, such as new tolerance factor of perovskite oxides and halides, and the activity and selectivity of the single atom catalysts in NRR.²⁰⁻²³ Here, the value of ΔE was selected as the target results to describe the activity of NRR in metal-zeolites. The SISSO approach could select 2.04×10^{11} candidate three-dimensional (3D) descriptors, where the dimensionality is defined as the number of iterations of complexity in the

descriptor space. The SISO features are obtained as follows.

$$SISO1 = \frac{d_{O1}}{d_{MN}} + \frac{N_H}{N_N} - (e^Q - e^{Q_r}) \quad (S2)$$

$$SISO2 = \frac{|2\chi - N_d|}{IE \times N_N \times (N_H + N_N)} \quad (S3)$$

$$SISO3 = \frac{\left| \frac{N_N}{d_{MN}} - \frac{N_H}{d_{O4}} \right|}{\sqrt{N_N \times R/d_{O3}}} \quad (S4)$$

Three 3D SISO features constructed by a combination of functional operations of 12 initial features and two framework features were then selected to give the prediction of NRR process by nine algorithms of feature learning in **Figs. S43** and **S44**. With the number of the features reduced, the GBR algorithm could still exhibit a good fitting ability (MAE = 0.39 eV, RMSE = 0.56 eV, and $R^2 = 0.91$) with the parameters in **Table S10**, indicating the feasibility of such a model for predicting the nitrogen fixation process in metal-zeolite systems. The importance of three SISO features and two framework features are 0.29, 0.25, 0.21, 0.13, and 0.12, respectively, as shown in **Fig. S43d**. The prediction performance of other algorithms, such as ANN, linear ridge, MLR, SVR, kNN, and Decisiontree, could get improved with the introduction of the SISO features obviously, indicating the importance of the feature establishment (**Fig. S45**).

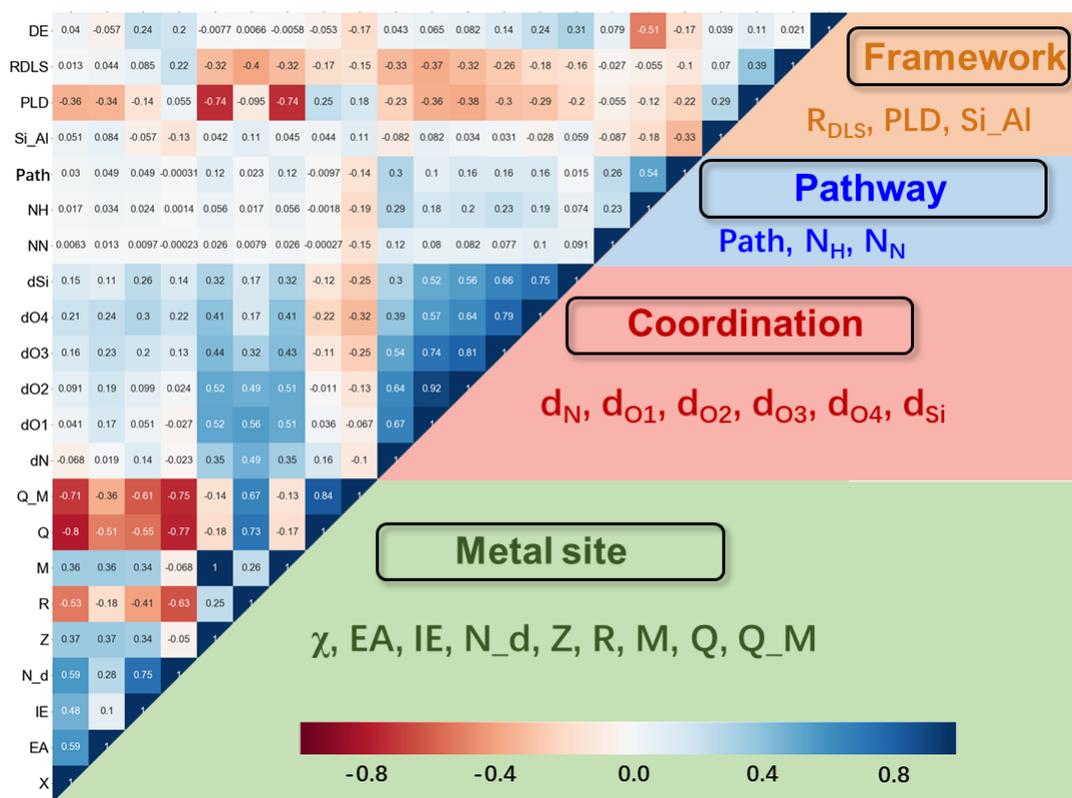


Fig. S41 Pearson correlation coefficient matrix of 21 features in multiparameter model, which were classified in framework (in orange color), reaction pathway (in blue), coordination (in red), and active site (in green) descriptors.

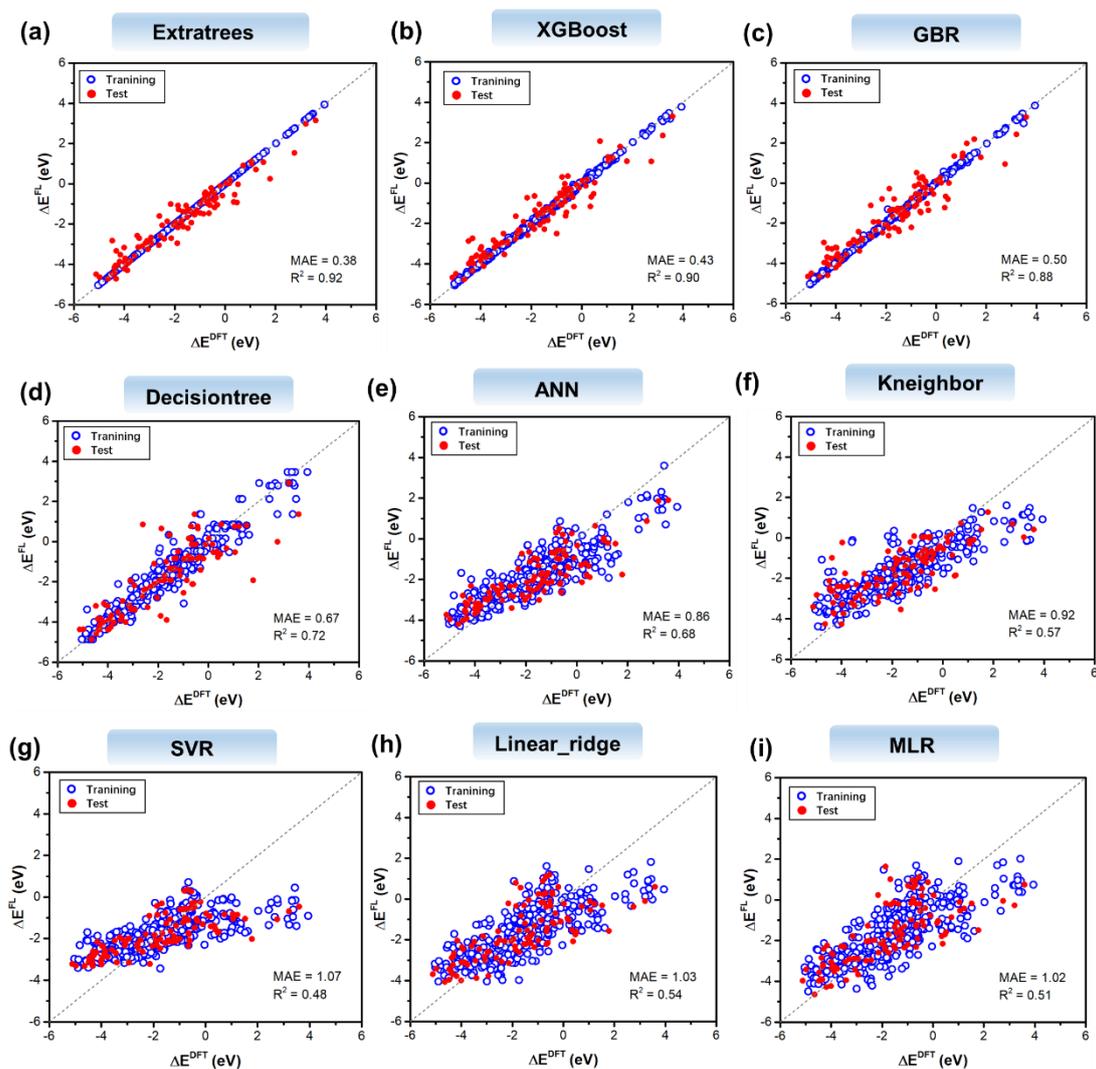


Fig. S42 Performance of different feature learning algorithms, including (a) Extratrees, (b) XGBoost, (c) GBR, (d) Decisiontree, (e) ANN, (f) Kneighbor, (g) SVR, (h) Linear_ridge, and (i) MLR with 21 initial features.

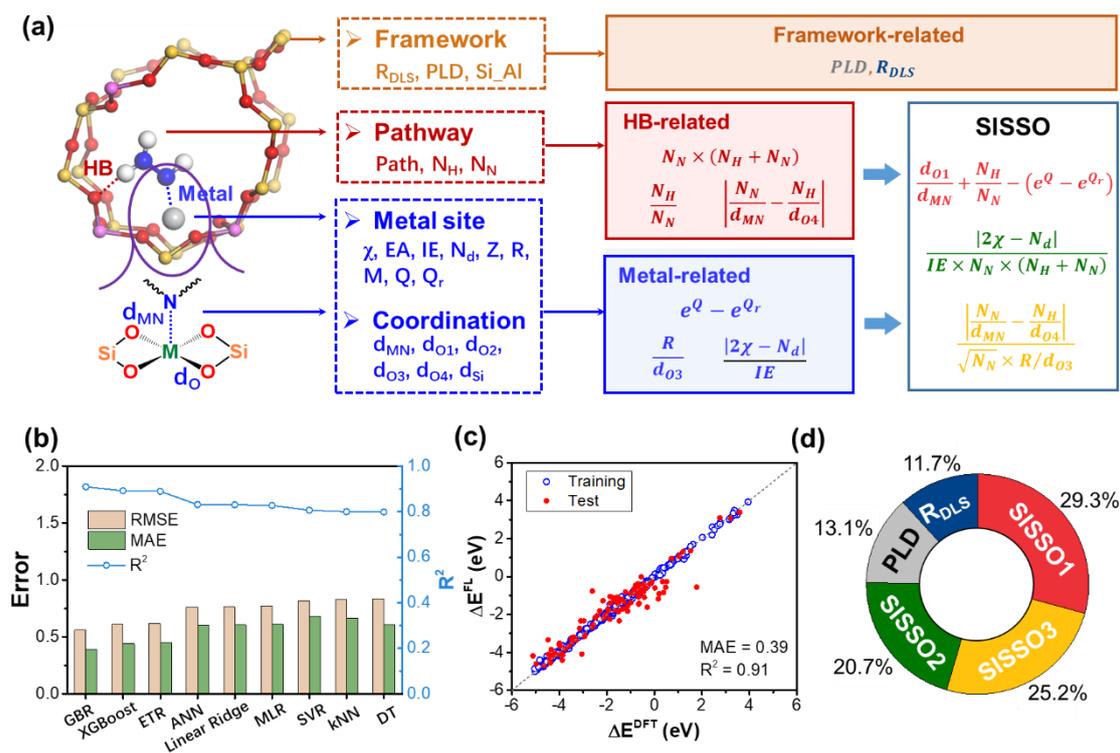


Fig. S43 (a) The selection process of three SISSO descriptors and two framework descriptors; (b) comparison of the performance of different algorithms of the FL according to the root-mean-squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2); (c) prediction of ΔE of NRR process by GBR algorithm via three SISSO 3D descriptors and two framework descriptors; (d) the feature importance of GBR models for predicted ΔE .

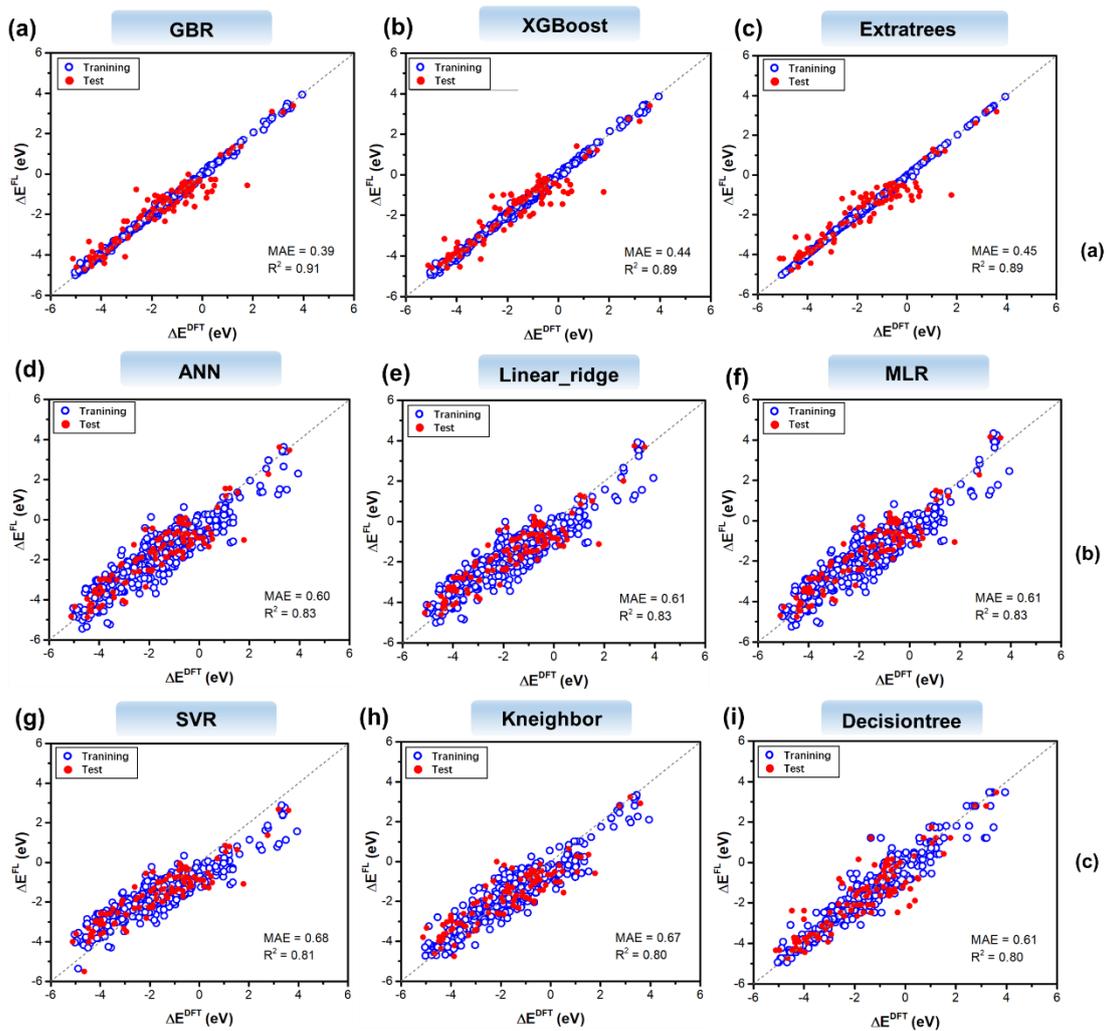


Fig. S44 Performance of different feature learning algorithms, including (a) GBR, (b) XGBoost, (c) Extratrees, (d) ANN, (e) Linear_ridge, (f) MLR (g) SVR, (h) Kneighbor, and (i) Decisiontree with 3 SISSO features and 2 zeolite framework features.

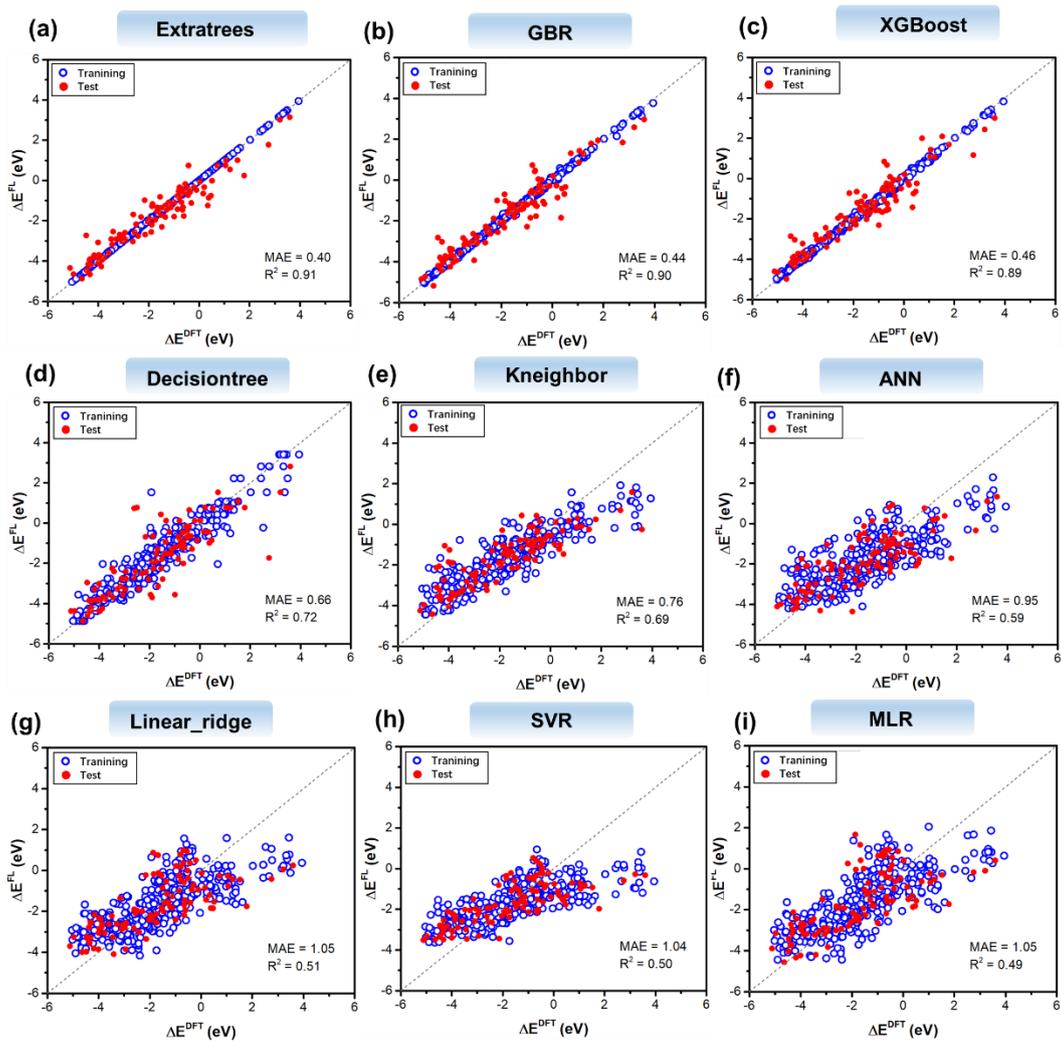


Fig. S45 Performance of different feature learning algorithms, including (a) Extratrees, (b) GBR, (c) XGBoost, (d) Decisiontree, (e) Kneighbor, (f) ANN, (g) Linear_ridge, (h) SVR, and (i) MLR with 14 initial features.

Table S10. Parameters of the multiparameter feature learning model

Parameter	value
learning_rate	0.115
max_depth	12
n_estimators	120
Subsample	0.78
alpha	0.9
min_samples_leaf	8
min_sample_split	5

Table S11. The bond length of N-N following the enzymatic pathway

Systems	Substrate	Active site	d_{NN} (Å)	E_{ad} (eV)	RLS (eV)	Ref.
W-zeolite	MFI	W	1.27	-1.40	0.47	this work
Mo-embedded BN monolayer	BN monolayer	MoN ₃	1.20	-0.87	0.35	<i>J. Am. Chem. Soc.</i> 2017 , <i>139</i> , 12480–12487
V@BN	BN monolayer	VN ₃	1.18	-0.73	0.41	<i>Nanoscale</i> , 2020 , <i>12</i> , 1541–1550
Tc@BN		TcN ₃	1.21	-1.15	0.59	
V/ β_{12} -BM	β_{12} -Boron monolayer	V	1.16	-0.63	0.28	<i>J. Phys. Chem. C</i> 2019 , <i>123</i> , 4274–4281
Mn/ β_{12} -BM		Mn	1.15	-0.52	0.83	
TiB ₂ monolayer	TiB ₂	Ti ₃	1.34	-2.54	0.58	<i>J. Mater. Chem. A</i> , 2019 , <i>7</i> , 25887–25893
NbB ₂ monolayer	NbB ₂	Nb ₃	1.34	-3.13	0.64	
meta-doped BP	phosphorene	meta-B ₂	1.22	-1.20	0.58	<i>J. Mater. Chem. A</i> , 2019 , <i>7</i> , 4865–4871
B ₂ @MoS ₂	MoS ₂	B ₂	1.24	-2.11	0.19	<i>Nanoscale</i> , 2019 , <i>11</i> , 18769–18778
Ru ₂ @GY	graphyne	Ru ₂	1.18	-0.35	0.43	<i>J. Phys. Chem. C</i> 2020 , <i>124</i> , 15295–15301
Mn ₂ @GDY	graphdiyne	Mn ₂	1.17	-0.55	0.52	<i>J. Phys. Chem. C</i> 2019 , <i>123</i> , 19066–19076
Fe ₂ @GDY		Fe ₂	1.19	-0.83	0.56	
Co ₂ @GDY		Co ₂	1.17	-0.75	0.54	
B@g-CN	g-CN	B	1.21	-0.86	0.31	<i>Nano Lett.</i> 2019 , <i>19</i> , 6391–6399
B/C ₂ N	C ₂ N	B	1.21	-0.58	0.18	<i>Phys. Chem. Chem. Phys.</i> , 2019 , <i>21</i> , 12346–12352
B _{int} -doped C ₂ N layer	C ₂ N	B	1.18	-0.77	0.15	<i>J. Mater. Chem. A</i> , 2019 , <i>7</i> , 2392–2399
B@C ₂ N	C ₂ N	B	1.21	-0.98	0.45	<i>Nanotechnology</i> , 2019 , <i>30</i> , 335403
B ₂ @C ₂ N		B ₂	1.21	-2.11	0.35	
Mn ₂ @C ₂ N	C ₂ N	Mn ₂	1.21	-0.65	0.23	<i>Small Methods</i> 2019 , <i>3</i> , 1800291
Fe ₂ -N ₆ @G	graphene	Fe ₂ N ₆	1.20	-0.51	0.46	<i>J. Phys. Chem. Lett.</i> 2020 , <i>11</i> , 6320–6329
Ru ₂ -N ₆ @G		Ru ₂ N ₆	1.21	-1.56	0.36	
Ir ₂ -N ₆ @G		Ir ₂ N ₆	1.20	-0.46	0.53	
Mo ₁ -N ₁ C ₂	graphene	MoN ₁ C ₂	1.18	-1.19	0.40	<i>J. Phys. Chem. C</i> 2018 , <i>122</i> , 16842–16847
Mo/SeG	graphene	Mo/Se	1.18	-1.13	0.41	<i>Phys. Chem. Chem. Phys.</i> , 2019 , <i>21</i> , 14583–14588
Mo@BCN-5	BN/graphene	Mo	1.19	-0.97	0.58	<i>J. Mater. Chem. A</i> , 2019 , <i>7</i> , 15173–15180
MoN ₃ @555-777 graphene	graphene	MoN ₃	1.22	-1.14	0.65	<i>Phys. Chem. Chem. Phys.</i> , 2020 , <i>22</i> , 9322–9329
GDY-2B(S ₂ S ₂)	graphdiyne	B ₂	1.28	-0.85	0.28	<i>Phys. Chem. Chem. Phys.</i> , 2021 , <i>23</i> , 17683–17692
GDY-2B(S ₄ A ₁)		B ₂	1.21	-0.32	0.78	

Table S12. NH₃ desorption free energies and yields in experiment

Systems	NH ₃ desorption free energies (eV)	Yield (μg h ⁻¹ mg _{cat} ⁻¹)	FE	Ref.
boron carbide nanosheet	1.73	26.57	15.95% (-0.75 V)	<i>Nat. Commun.</i> , 2018 , 9, 3485.
boron phosphide nanoparticles	1.23	26.42	12.7% (-0.60 V)	<i>J. Mater. Chem. A</i> , 2019 , 7, 16117.
iron phthalocyanine	0.94	137.95	10.5% (-0.30 V)	<i>ACS Catal.</i> , 2019 , 9, 7311–7317.
NiPS ₃ nanosheets	1.32	118	>17% (-0.40 V)	<i>Chem. Eng. J.</i> , 2022 , 430, 132649.
Fe-SnO ₂	1.16	82.7	20.4 % (-0.30 V)	<i>Angew. Chem. Int. Ed.</i> , 2020 , 59, 10888–10893.

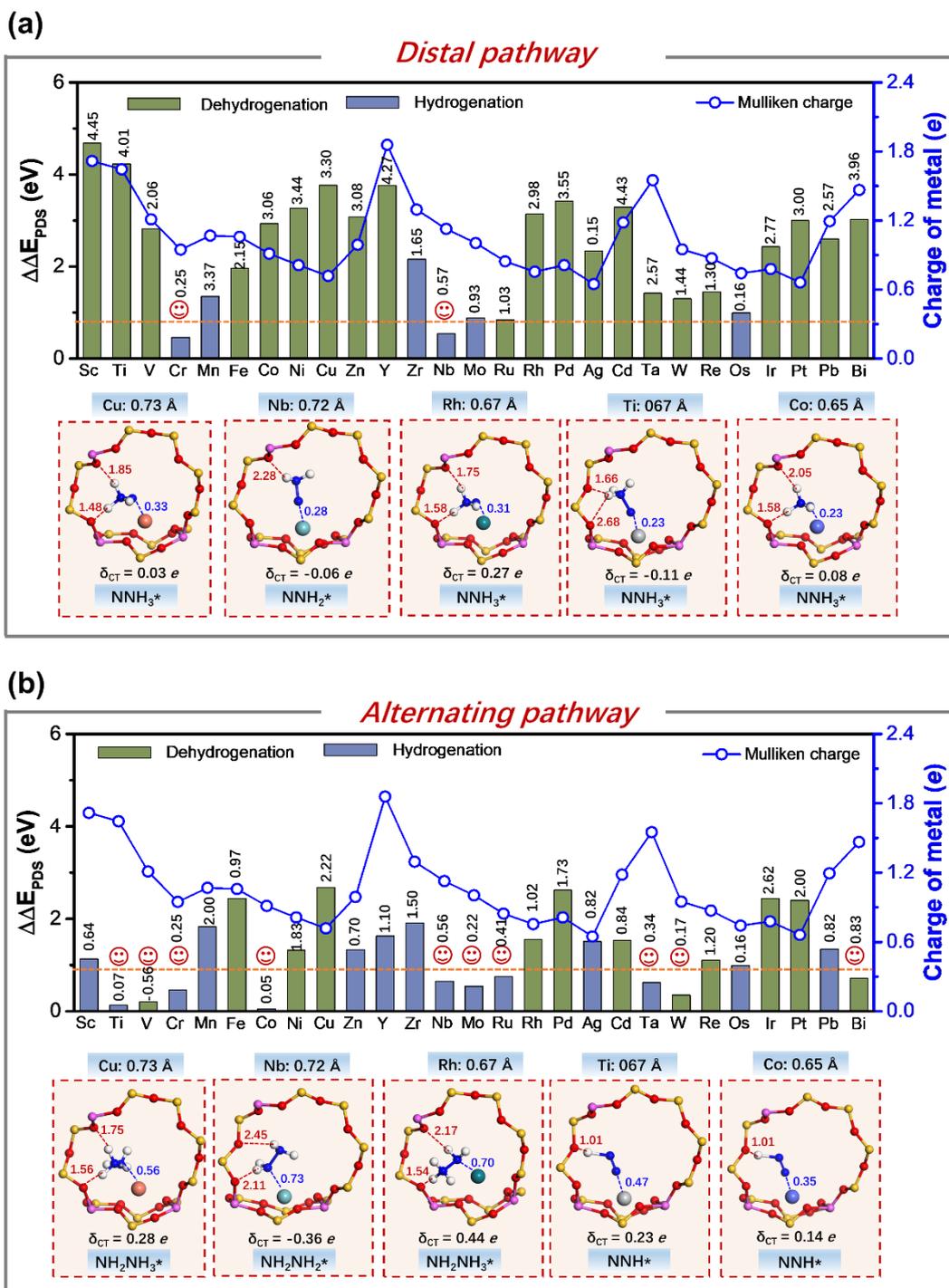


Fig. S46 Summary of $\Delta\Delta E_{RLS}$ values obtained by DFT and feature learning and Mulliken charges on 27 metal-zeolite catalysts along (a) distal and (b) alternating pathways. Color codes represent the product at the RLS of each catalyst; the structures of key intermediates in Cu-, Nb-, Rh-, Ti-, and Co-zeolites.

S6. Non-metal zeolites

In silicalite MFI zeolites, the distal pathway with the hydrogen bonding interaction was calculated for the reduction of N_2 to NH_3 , as shown in **Fig. S47a**. The adsorption of nitrogen can be reach -0.30 eV, indicating that the first step of nitrogen fixation can be realized. However, the first hydrogenation occurs by a proton coupled with an electron attacking the nitrogen in framework to form the N_2H^* intermediates, in which the energy is uphill by 2.18 eV. In the following step, the H atom consecutively attacks the same N atom to form the NNH_2^* and NNH_3^* species, with the 0.04 eV downhill and 1.68 eV uphill in energy, respectively. The first NH_3 molecule is released with the energy uphill of 0.13 eV. The other hydrogenation steps, including reducing species N^* to NH^* , NH^* to NH_2^* and NH_2^* to NH_3^* are downhill by -1.24 , -1.71 , and -3.09 eV, respectively. However, the too much energy input from N_2^* to N_2H^* indicates the difficulty of the nitrogen reduction in silicate zeolites. In **Fig. S47b**, the distance between the nitrogen atoms elongated from 1.11 Å to 1.43 Å with the H atom attacking the distal nitrogen atom until the first NH_3 molecule desorption, indicating that the reduction could take place in the zeolite under the optimized conditions. In **Fig. S47c**, the extent of charge variation along the NRR pathway is tiny, indicating that the charge transfer between intermediates and the zeolite framework is small. Hence, a more efficient zeolitic catalyst should be designed rationally to realize the nitrogen reduction reaction.

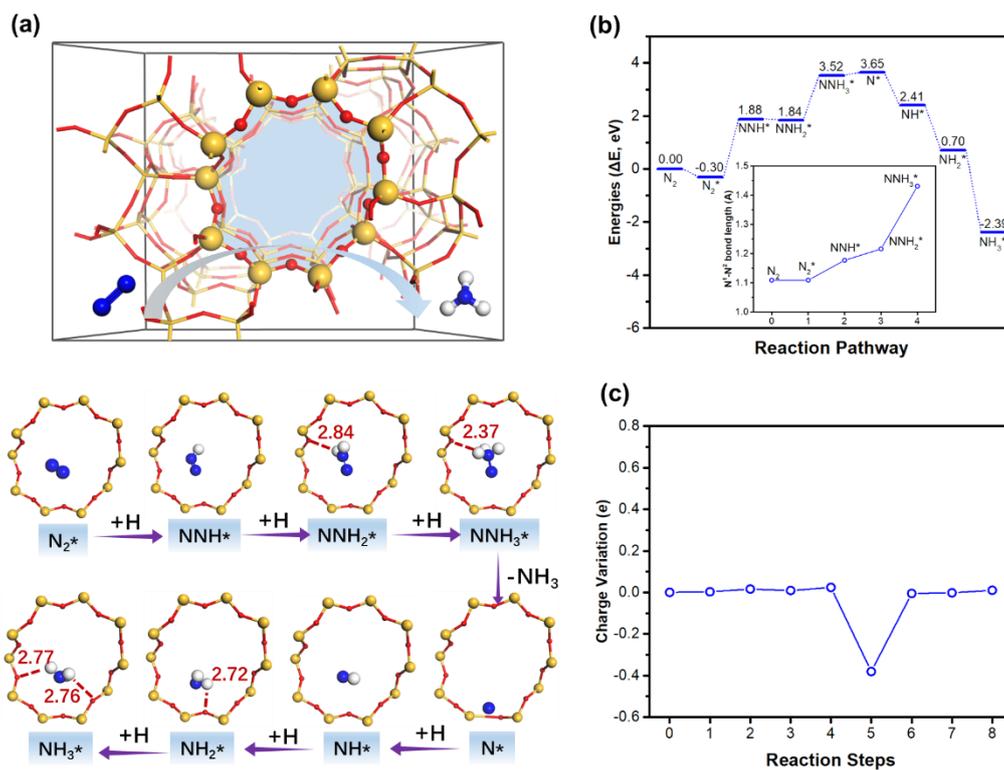


Fig. S47 (a) Schematic illustration of the distal reaction pathway during the N_2 reduction; (b) energy diagrams for NRR in silicalite MFI zeolite via distal pathway, the N^1-N^2 bond distance variation during N_2 fixation process also presented in inset; (c) charge variation of intermediates along the distal pathway of NRR.

S7. Bandgap of metal-zeolites

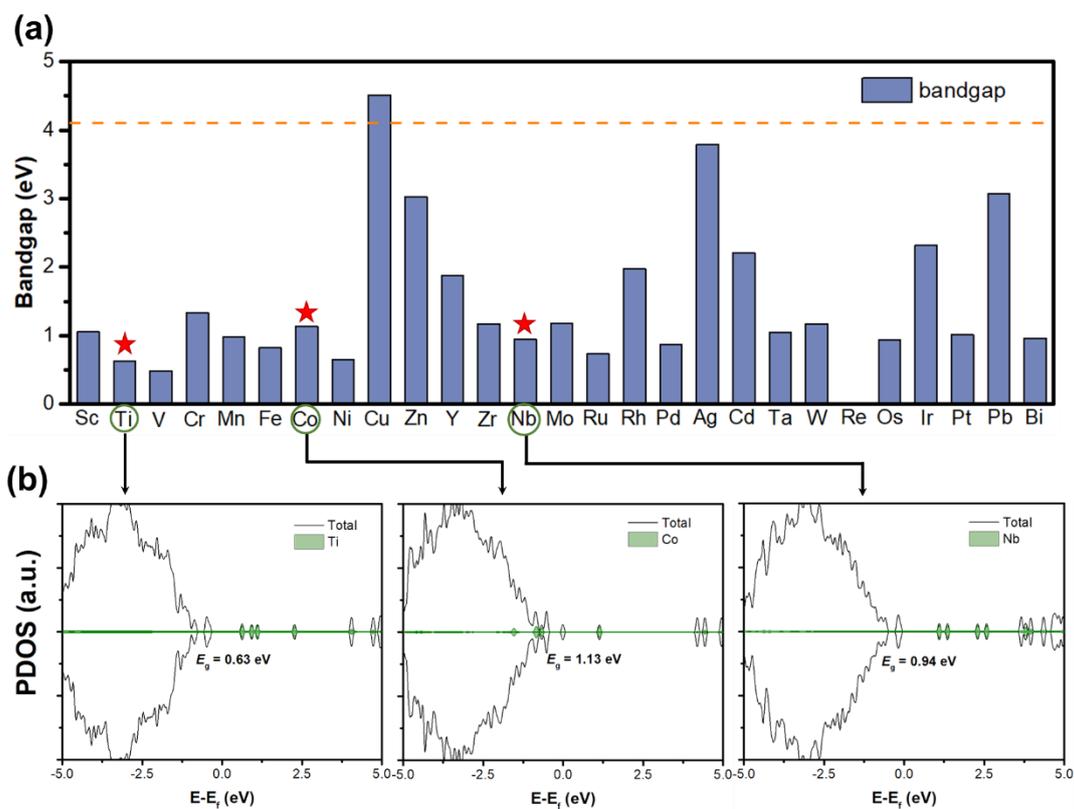


Fig. S48 (a) The bandgap of different metal-zeolites; for comparison, the Al-containing zeolite (Si/Al = 31) shown as a horizontal dashed line; (b) the PDOS of Ti-, Co-, Nb-zeolite, respectively.

S8. Solvation effects

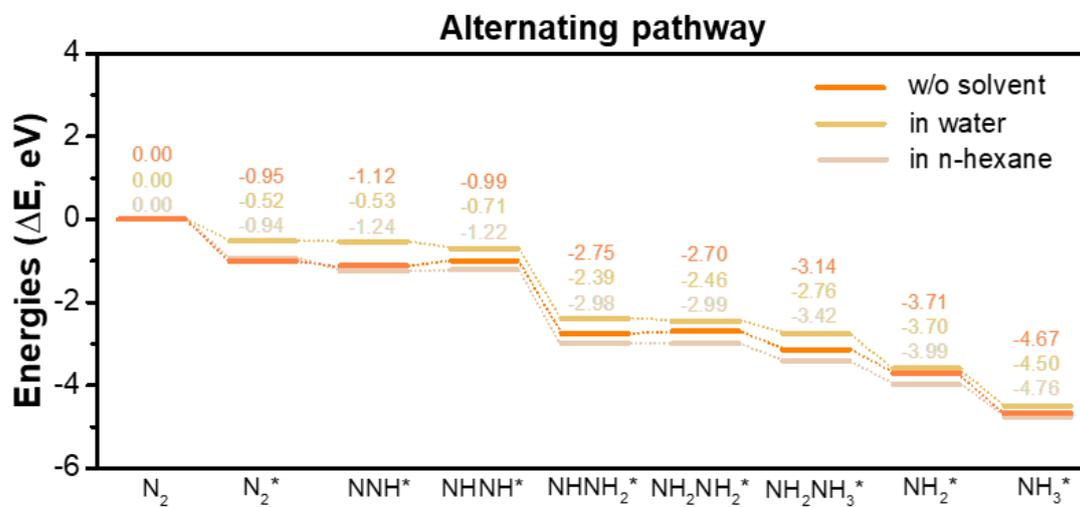


Fig. S49 The energy profiles for Ti-zeolite during alternating pathway by applying water, n-hexane solvents, and without solvation effects.

References

1. Delley, B., From molecules to solids with the DMol³ approach. *J. Chem. Phys.* **2000**, *113*, 7756–7764.
2. Materials Studio, version 7.0; Accelrys Inc.: San Diego. **2013**.
3. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
4. Perdew, J. P.; Wang, Y., Accurate and Simple Analytic Representation of the Electron-gas Correlation Energy. *Phys. Rev. B* **1992**, *45*, 13244–13249.
5. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti Correlation-energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785–789.
6. Becke, A. D., Density-functional Exchange-energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.
7. Grimme, S., Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27* (15), 1787–99.
8. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.
9. Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L., Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode. *J. Phys. Chem. B* **2004**, *108*, 17886–17892.
10. Khan, F.; Yue, P., Photoassisted Water Cleavage and Nitrogen Fixation over Titanium-Exchanged Zeolites. *Ind. Eng. Chem. Prod. Res. Dev.* **1983**, *22*, 238–241.
11. Khan, F.; Yue, P., Photochemical Synthesis of Ammonia over Zeolites. *J.C.S. CHEM. COMM.* **1981**, *745*, 1049–1050.
12. Himanen, L.; Jäger, M. O. J.; Morooka, E. V.; Federici Canova, F.; Ranawat, Y. S.; Gao, D. Z.; Rinke, P.; Foster, A. S., DSCRIBE: Library of Descriptors for Machine Learning in Materials Science. *Comput. Phys. Commun.* **2020**, *247*, 106949.
13. Berwanger, J.; Polesya, S.; Mankovsky, S.; Ebert, H.; Giessibl, F. J., Atomically Resolved Chemical Reactivity of Small Fe Clusters. *Physical Review Letters* **2020**, *124*, 096001.
14. Gu, Y.; Liu, Z.; Yu, C.; Gu, X.; Xu, L.; Gao, Y.; Ma, J., Zeolite Adsorption Isotherms Predicted by Pore Channel and Local Environmental Descriptors: Feature Learning on DFT Binding Strength. *J. Phys. Chem. C* **2020**, *124*, 9314–9328.
15. Thompson, N. B.; Green, M. T.; Peters, J. C., Nitrogen Fixation via a Terminal Fe(IV) Nitride. *J. Am. Chem. Soc.* **2017**, *139* (43), 15312–15315.
16. Hoffman, B. M.; Lukoyanov, D.; Yang, Z. Y.; Dean, D. R.; Seefeldt, L. C., Mechanism of Nitrogen Fixation by Nitrogenase: The Next Stage. *Chem. Rev.* **2014**, *114* (8), 4041–62.
17. Seefeldt, L. C.; Hoffman, B. M.; Dean, D. R., Mechanism of Mo-dependent Nitrogenase. *Annu. Rev. Biochem.* **2009**, *78*, 701–22.
18. Hendrich, M. P.; Gundersen, W.; Behan, R. K.; Green, M. T.; Mehn, M. P.; Betley, T. A.; Lu, C. C.; Peters, J. C., On the Feasibility of N₂ Fixation via a Single-site Fe^IFe^{IV} Cycle Spectroscopic Studies of Fe^I(N₂)Fe^I, Fe^{IV}≡N, and Related Species. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (46), 17107–17112.
19. Chatt, J.; Dilworth, J. R.; Richards, R. L., Recent Advances in the Chemistry of Nitrogen Fixation. *Chem. Rev.* **1978**, *78*, 589–625.
20. Zhao, W.; Chen, L.; Zhang, W.; Yang, J., Single Mo₁(W₁, Re₁) Atoms Anchored in Pyrrolic-N₃

Doped Graphene as Efficient Electrocatalysts for the Nitrogen Reduction Reaction. *J. Mater. Chem. A* **2021**, *9*, 6547-6554.

21. Ouyang, R.; Ahmetcik, E.; Carbogno, C.; Scheffler, M.; Ghiringhelli, L. M., Simultaneous Learning of Several Materials Properties from Incomplete Databases with Multi-task SISSO. *J. Phys.: Mater.* **2019**, *2*, 024002.

22. Bartel, J.; Sutton, C.; Goldsmith, B. R.; Ouyang, R.; Musgrave, C. B.; M., G. L.; Scheffler, M., New Tolentance Factor to Predict the Stablility of Perovskite Oxides and Halides. *Sci. Adv.* **2019**, *5*, eaav0693.

23. Ouyang, R.; Curtarolo, S.; Ahmetcik, E.; Scheffler, M.; Ghiringhelli, L. M., SISSO: A Compressed-sensing Method for Identifying the Best Low-dimensional Descriptor in an Immensity of Offered Candidates. *Phys. Rev. Mater.* **2018**, *2*, 083802.