Supplementary Information

Global Neural Network Potential with Explicit Many-body Functions for Improved Descriptions of Complex Potential Energy Surface

Pei-Lin Kang[1], Zheng-Xin Yang[1], Cheng Shang[1,2], Zhi-Pan Liu[1,2,3]*

[1]Collaborative Innovation Center of Chemistry for Energy Material, Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433, China

[2]Key Laboratory of Synthetic and Self-Assembly Chemistry for Organic Functional Molecules, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China

[3]Shanghai Qi Zhi Institution, Shanghai 200030, China

Table of Contents

1. Training information of Li-Co-O

Table S1 Structure information of the LiCoO dataset for NN training. Listed data are the number of the structures in the global dataset, as distinguished by the chemical formula, the number of atoms (Natoms), the type of structures (cluster, bulk, layer) and its total number (Ntotal).

Chemical Formula	Natoms	cluster	layer	bulk	total
O1-Co16	17	0	1	0	1
O2-Co16	18	0	0	10	10
O2-Co32	34	0	1	0	1
O4	4	0	32	0	32
O5-Co30	35	0	3	0	3
O6-Co8	14	0	0	2	2
O7-Co8	15	0	0	14	14
O8-Co6	14	80	0	0	80
O8-Co8	16	12	1	237	250
O10-Co8	18	0	6	273	279
O11	11	0	169	84	253
O12-Co8	20	11	4	553	568
O12-Co12	24	0	0	1	1
O12-Co16	28	0	1	1	2
O13-Co12	25	0	0	2	2
O14-Co12	26	0	0	3	3
O14-Co16	30	0	4	1	5
O15-Co12	27	0	0	2	2
O15-Co16	31	0	0	1	1
O16-Co8	24	0	54	116	170
O16-Co12	28	0	8	308	316
O16-Co16	32	0	74	303	377
O18-Co12	30	0	0	25	25
O20-Co12	32	0	0	2	2
O20-Co16	36	0	21	68	89
O24-Co12	36	0	29	51	80
O24-Co16	40	0	87	147	234
O24-Co24	48	0	1	0	1
O26-Co20	46	0	0	2	2
O26-Co24	50	0	1	0	1
O28-Co14	42	0	0	1	1
O32-Co16	48	0	13	1	14
O32-Co24	56	0	16	22	38
O32-Co32	64	0	7	4	11
O36-Co24	60	0	3	1	4
O40-Co32	72	0	3	0	3
O48-Co24	72	0	2	5	7
O48-Co32	80	0	5	4	9
O54-Co27	81	0	1	20	21

O64-Co48	112	0	0	1	1
O96-Co48	144	0	1	9	10
Li1-O4-Co3	8	0	0	1	1
Li1-O5-Co3	9	0	0	2	2
Li1-O6-Co3	10	0	0	3	3
Li1-O7-Co3	11	0	0	9	9
Li1-O8-Co3	12	0	0	4	4
Li1-O8-Co4	13	0	0	287	287
Li1-O12-Co7	20	0	0	1	1
Li1-O16-Co8	25	0	27	35	62
Li2-O4-Co2	8	0	0	3	3
Li2-O5-Co2	9	0	0	2	2
Li2-O7-Co2	11	0	0	2	2
Li2-O8-Co2	12	0	0	1	1
Li2-O8-Co4	14	0	0	282	282
Li2-O10-Co6	18	0	0	1	1
Li2-O14-Co6	22	0	100	295	395
Li2-O14-Co7	23	0	8	13	21
Li2-O16-Co6	24	0	30	68	98
Li2-O16-Co8	26	0	346	542	888
Li2-O24-Co14	40	0	2	2	4
Li2-O26-Co14	42	0	0	5	5
Li2-O28-Co14	44	0	0	6	6
Li2-O30-Co14	46	0	1	10	11
Li2-O32-Co14	48	0	0	1	1
Li2-O32-Co16	50	0	9	12	21
Li3-O4-Co2	9	0	2	2	4
Li3-O8-Co4	15	0	8	257	265
Li3-O24-Co12	39	0	2	147	149
Li3-O54-Co27	84	0	0	18	18
Li4-O4-Co2	10	0	1	3	4
Li4-O8-Co4	16	0	0	198	198
Li4-O12-Co10	26	0	0	1	1
Li4-O14-Co4	22	0	6	59	65
Li4-O16-Co4	24	0	11	40	51
Li4-O16-Co8	28	0	329	518	847
Li4-O16-Co12	32	0	0	11	11
Li4-O20-Co12	36	0	7	21	28
Li4-O22-Co12	38	0	6	9	15
Li4-O24-Co12	40	0	8	95	103
Li4-O26-Co12	42	0	0	6	6
Li4-O28-Co12	44	0	63	84	147
Li4-O28-Co14	46	0	9	13	22
Li4-O30-Co12	46	0	0	4	4

Li4-O32-Co12	48	0	0	2	2
Li4-O32-Co16	52	0	31	40	71
Li4-O52-Co28	84	0	3	3	6
Li4-O56-Co28	88	0	1	0	1
Li4-O60-Co28	92	0	0	9	9
Li4-O64-Co32	100	0	1	0	1
Li4-O96-Co48	148	0	0	40	40
Li5-O4-Co2	11	0	2	7	9
Li5-O14-Co7	26	0	12	11	23
Li5-O16-Co8	29	0	42	117	159
Li6-O4-Co2	12	0	2	5	7
Li6-O8-Co4	18	0	3	2	5
Li6-O14-Co2	22	0	6	3	9
Li6-O16-Co2	24	0	12	4	16
Li6-O16-Co8	30	0	374	678	1052
Li6-O16-Co10	32	0	0	1	1
Li6-O22-Co10	38	0	2	9	11
Li6-O24-Co10	40	0	0	4	4
Li6-O24-Co12	42	0	0	107	107
Li6-O26-Co10	42	0	1	2	3
Li6-O30-Co10	46	0	0	2	2
Li6-O48-Co24	78	0	16	124	140
Li6-O54-Co27	87	0	0	15	15
Li7-O14-Co7	28	0	5	21	26
Li7-O16-Co8	31	0	9	13	22
Li8-O8-Co4	20	0	4	2	6
Li8-O12-Co6	26	0	0	8	8
Li8-O16-Co8	32	0	259	537	796
Li8-O20-Co8	36	0	5	22	27
Li8-O24-Co8	40	0	0	9	9
Li8-O28-Co8	44	0	0	2	2
Li8-O30-Co8	46	0	1	4	5
Li8-O32-Co16	56	0	28	24	52
Li8-O40-Co24	72	0	0	1	1
Li8-O44-Co24	76	0	1	6	7
Li8-O48-Co24	80	0	9	20	29
Li8-O52-Co24	84	0	5	4	9
Li8-O56-Co24	88	0	2	6	8
Li8-O56-Co28	92	0	0	1	1
Li8-O96-Co48	152	0	0	49	49
Li8-O104-Co56	168	0	1	1	2
Li9-O16-Co8	33	0	19	38	57
Li9-O24-Co12	45	0	1	99	100
Li9-O54-Co27	90	0	0	23	23

Li10-O8-Co4	22	0	15	50	65
Li10-O20-Co6	36	0	0	1	1
Li10-O28-Co14	52	0	5	6	11
Li10-O30-Co6	46	0	1	1	2
Li10-O32-Co16	58	0	52	34	86
Li11-O8-Co4	23	0	10	47	57
Li12-O8-Co4	24	0	39	96	135
Li12-O16-Co4	32	0	0	3	3
Li12-O16-Co8	36	0	13	30	43
Li12-O22-Co4	38	0	1	0	1
Li12-O24-Co4	40	0	1	5	6
Li12-O24-Co12	48	0	20	123	143
Li12-O32-Co16	60	0	31	27	58
Li12-O44-Co20	76	0	0	2	2
Li12-O48-Co24	84	0	0	25	25
Li12-O54-Co27	93	0	0	23	23
Li14-O28-Co14	56	0	3	0	3
Li14-O32-Co16	62	0	5	4	9
Li15-O54-Co27	96	0	0	11	11
Li16-O16-Co8	40	0	16	36	52
Li16-O32-Co16	64	0	50	47	97
Li16-O64-Co32	112	0	1	2	3
Li16-O96-Co48	160	0	0	3	3
Li18-O32-Co16	66	0	17	12	29
Li18-O48-Co24	90	0	0	10	10
Li18-O54-Co27	99	0	0	11	11
Li20-O16-Co8	44	0	1	1	2
Li20-O64-Co32	116	0	1	1	2
Li21-O54-Co27	102	0	0	8	8
Li22-O16-Co8	46	0	24	6	30
Li24-O16-Co8	48	0	6	7	13
Li24-O32-Co16	72	0	6	1	7
Li24-O48-Co24	96	0	0	6	6
Li24-O54-Co27	105	0	0	15	15
Li24-O64-Co32	120	0	2	0	2
Li27-O54-Co27	108	0	0	23	23
Li32-O64-Co32	128	0	2	1	3
Li36-O64-Co32	132	0	1	0	1
Li44-O32-Co16	92	0	0	4	4
Li48	48	0	61	0	61
Li48-O32-Co16	96	0	1	0	1
Li64-O32	96	0	36	0	36
total	--	103	2801	8104	11008

2. Training information of Ti-O

Table S2 Structure information of the TiO dataset for NN training. Listed data are the number of the structures in the global dataset, as distinguished by the chemical formula, the number of atoms (Natoms), the type of structures (cluster, bulk, layer) and its total number (Ntotal).

Chemical Formula	Natoms	cluster	layer	bulk	total
Ti8	8	0	13	104	117
Ti13	13	0	8	7	15
Ti16	16	1103	3	3804	4910
Ti21	21	0	4	1421	1425
Ti32	32	0	5	337	342
O1-Ti12	13	0	1583	0	1583
O2-Ti8	10	0	104	323	427
O2-Ti10	12	0	871	0	871
O3-Ti18	21	0	151	0	151
O4	4	0	94	0	94
O4-Ti4	8	0	180	589	769
O4-Ti8	12	0	93	356	449
O4-Ti22	26	0	12	28	40
O5-Ti27	32	0	0	34	34
O6-Ti4	10	0	0	47	47
O6-Ti8	14	0	104	423	527
O6-Ti15	21	0	133	0	133
O6-Ti18	24	0	9	55	64
O8-Ti4	12	2120	190	11379	13689
O8-Ti8	16	0	95	361	456
O8-Ti15	23	0	22	41	63
O10-Ti8	18	0	135	437	572
O10-Ti22	32	0	0	34	34
O11	11	0	478	146	624
O12-Ti8	20	0	0	378	378
O13-Ti6	19	0	145	478	623
O13-Ti7	20	0	44	19	63
O14-Ti7	21	0	0	1	1
O14-Ti8	22	0	0	198	198
O15-Ti9	24	0	2	283	285
O16-Ti8	24	0	127	541	668
O18-Ti7	25	0	0	1	1
O18-Ti8	26	0	0	3	3
O18-Ti12	30	0	349	17	366
O120-Ti59	179	0	1	9	10
O120-Ti60	180	0	1	0	1
O126-Ti63	189	0	0	15	15
O126-Ti64	190	0	0	12	12
O127-Ti63	190	0	0	1	1

O127-Ti64	191	0	0	1	1
O128-Ti63	191	0	0	4	4
O128-Ti64	192	0	0	10	10
O239-Ti119	358	0	1	0	1
O240-Ti120	360	0	2	0	2
total	--	3223	4959	21897	30079

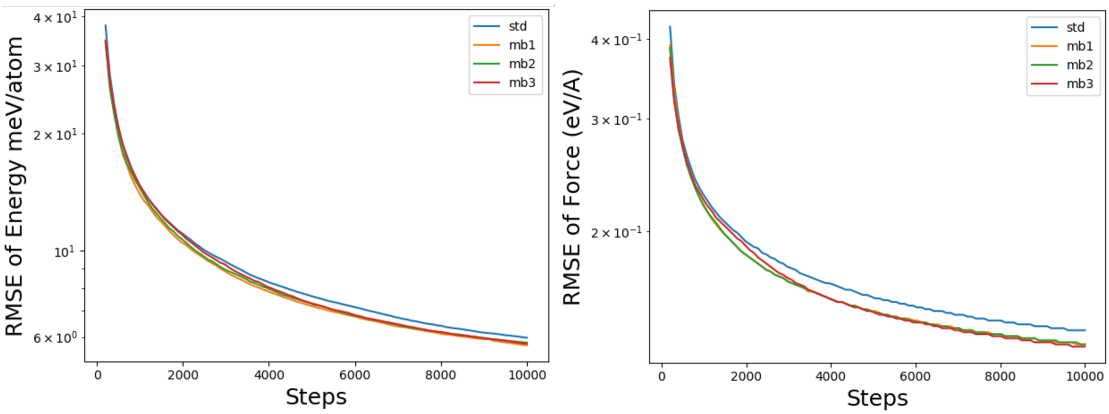


Figure S1 Comparison of learning curve of 4 Ti-O neural networks with different architecture.

3. MB functions in fitting Claisen rearrangement

Table S3. MB function Parameters in MBNN used for fitting Claisen rearrangement (in Figure 4).

MB terms	r_{cut} (Å)	m_d/m_t	l_t/l_q
f_{d1}	2	6	
f_{d1}	2.5	6	
f_{d1}	3	3	
f_{d1}	3	1	
f_{d1}	3.5	3	
f_{d1}	4.5	4	
f_{d1}	5.5	3	
f_{d1}	6.5	1	
f_{d1}	8.5	4	
f_{d1}	10	3	
f_{d1}	10	1	
f_{d1}	15	3	
f_{d1}	3	6	
f_{t2}	3	0	1
f_{t1}	3.5	2	1
f_{t2}	3.5	6	2
f_{t1}	4.5	2	2
f_{t1}	5.5	2	1
f_{t2}	5.5	2	1
f_q	2.5		1
f_q	2.5		1
f_q	2.5		1

4. Test results of Claisen rearrangement

The benchmark of 31 structures along the reaction channel using the MBNN and standard NN trained on the SSW dataset is shown in Table S4.

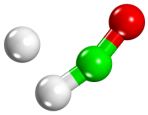
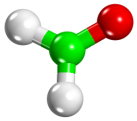
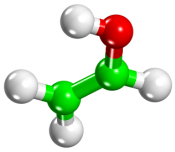
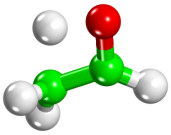
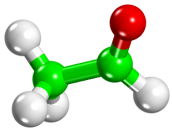
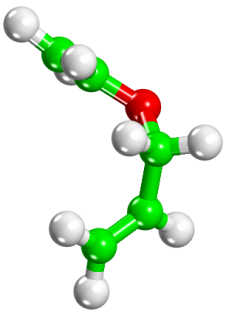
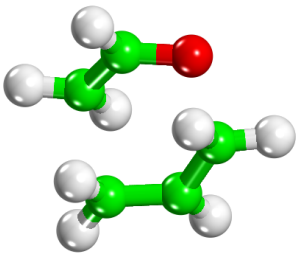
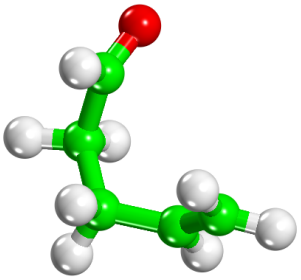
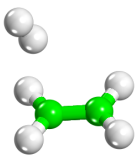
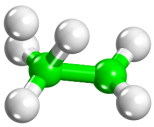
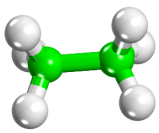
Table S4 Absolute error statistics for 31 structures along the reaction channel. The energy in the first column is the relative energy from DFT calculations.

Energy/eV	Error (meV/atom)	
	std-NN	**MBNN**
0.723	0.281	0.866
0.726	0.284	0.875
0.723	0.274	0.869
0.724	0.314	0.782
0.732	0.396	0.704
0.760	0.285	0.450
0.777	1.128	0.263
0.806	0.766	0.141
0.837	1.276	1.002
0.899	0.394	2.887
0.979	2.394	0.981
1.164	0.594	0.133
1.401	0.317	0.593
1.651	4.221	1.081
1.701	10.568	5.151
1.702	12.069	5.189
1.682	11.720	2.077
1.504	13.036	1.493
0.574	8.008	5.719
0.236	0.742	3.155
0.088	0.648	1.217
0.053	0.289	4.149
0.023	0.959	0.258
0.013	1.894	0.346
0.010	1.754	0.474
0.008	1.787	0.428
0.006	1.734	0.364
0.002	0.642	0.351
0.0002	0.381	0.715
0.0001	0.432	0.701
0.000	0.440	0.696
MAE	**2.58**	**1.42**
RMSE	**4.65**	**2.11**

5. Benchmark results for 11 distinct organic reactions

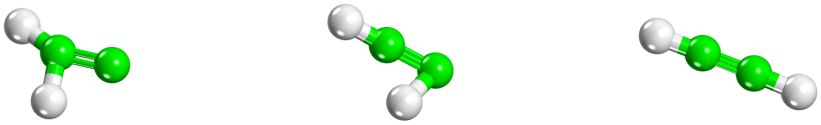
The benchmark results of the initial state (IS), transition state(TS) and final state (FS) of 11 distinct organic reactions are shown in Table S3.

Table S5 Potential energies of IS, TS and FS of 11 distinct organic reactions corresponding to Figure 5 in the main text. The energy of IS of each reaction from DFT calculations is set to be zero. The energy is in the unit of eV. mb: MBNN; std: standard NN

	IS	**TS**	**FS**
		$H_2CO \rightarrow H_2 + CO$	
			
DFT	0.00	2.82	-0.60
mb	-0.05	2.81	-0.61
std	0.00	2.88	-0.59
		$H_2CCHOH \rightarrow H_3CCHO$	
			
DFT	0.00	2.25	-0.35
mb	0.03	2.19	-0.37
std	0.08	2.21	-0.36
		Claisen rearrangement	
			
DFT	0.00	0.90	-0.78
mb	0.00	0.99	-0.75
std	0.00	0.97	-0.73
		$H_3CCH_3 \rightarrow H_2CCH_2 + H_2$	
			
DFT	0.00	3.07	-1.64

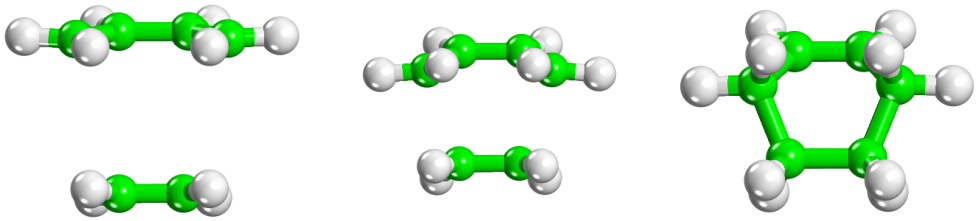
mb	0.05	3.23	-1.63
std	0.09	3.50	-1.60

HCCH → CCH_2



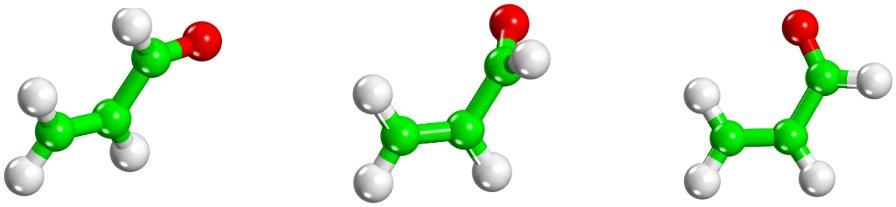
DFT	0.00	-0.02	-1.97
mb	0.02	-0.04	-1.99
std	0.09	-0.01	-1.98

parent Diels-Alder



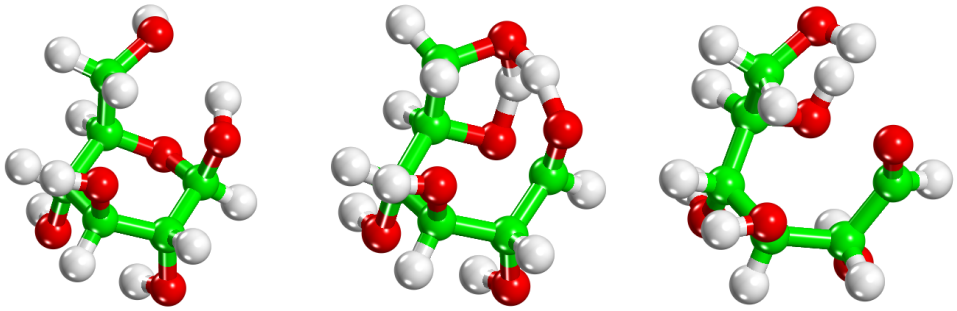
DFT	0.00	0.52	-1.80
mb	0.09	0.57	-1.75
std	0.12	0.56	-1.76

Acrolein rotation

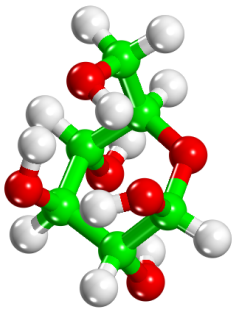
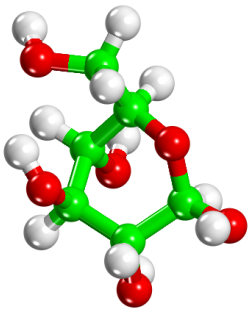
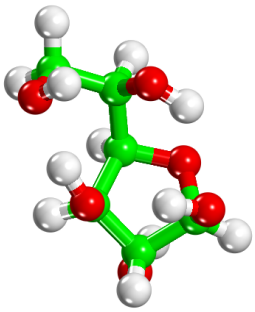
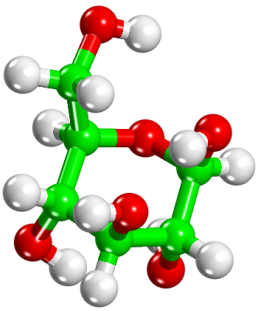
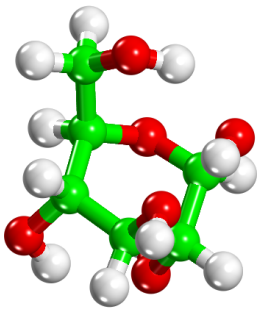
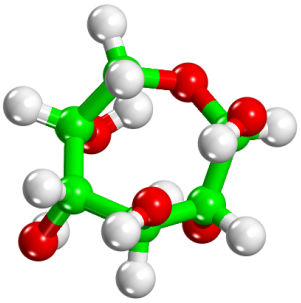
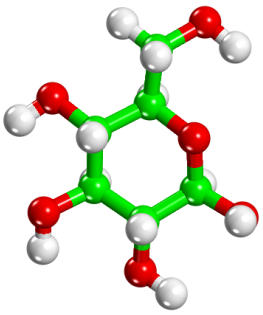
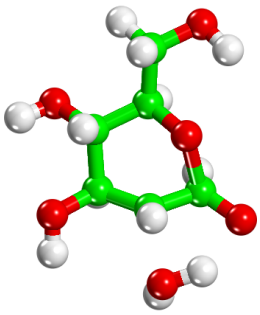
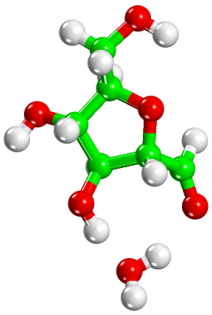


DFT	0.00	0.38	0.11
mb	0.13	0.45	0.14
std	0.11	0.49	0.09

glucopyranose → glucose



DFT	0.00	1.09	0.40
mb	-0.01	1.09	0.43
std	-0.05	1.08	0.43

glucopyranose → glucofuranose			
			
DFT	0.00	0.54	-0.11
mb	0.05	0.51	-0.11
std	0.06	0.52	-0.12
glucopyranose → glucoseptanose			
			
DFT	0.00	0.08	0.05
mb	0.01	0.11	0.09
std	0.01	0.12	0.03
glucopyranose → anhydromannose			
			
DFT	0.00	2.39	-0.06
mb	0.10	2.44	-0.09
std	0.05	2.54	-0.08