Supplementary Information For

Are Energy and Forces Really Enough? Using Structure to Evaluate the Accuracy and Transferability of Machine Learning Potentials of Biomolecules

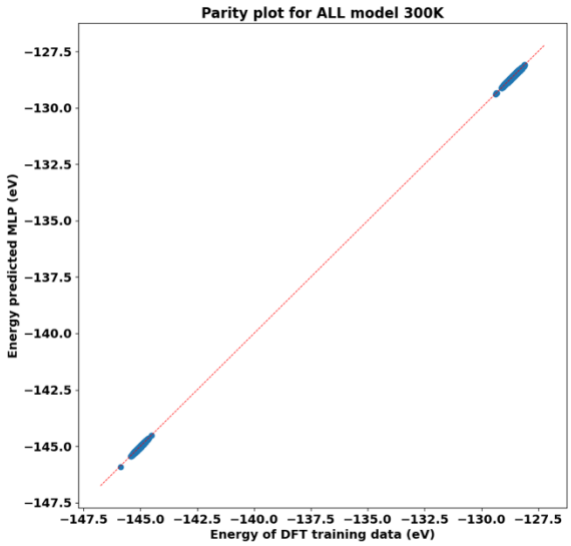
Lejla Š Biberić[1], Nisarg Joshi[1], Jim Pfaendtner[1,2]

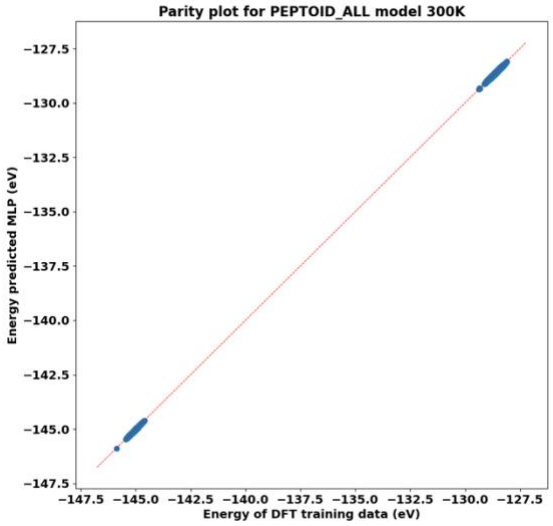
[1]Department of Biomolecular and Chemical Engineering, North Carolina State University, Raleigh, North Carolina, 27607, United States of America

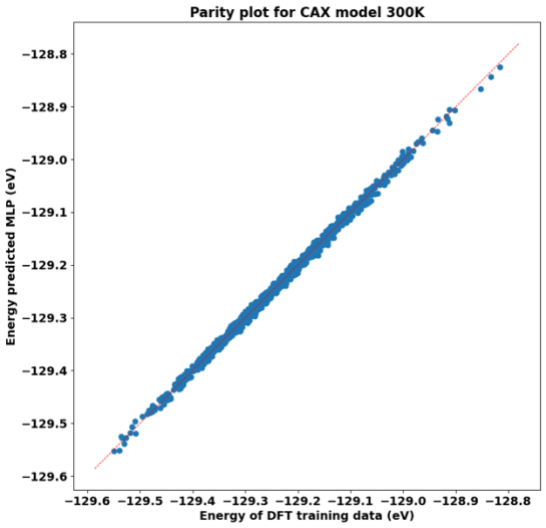
[2]Department of Chemistry, North Carolina State University, Raleigh, North Carolina, 27607, United States of America

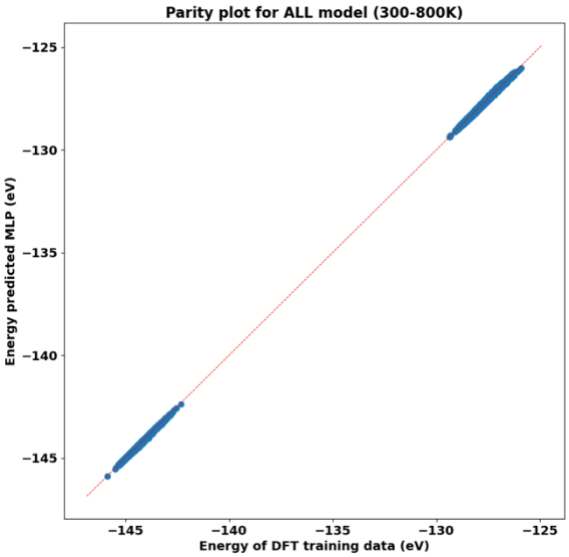
*Correspondence can be sent to pfaendtner@ncsu.edu

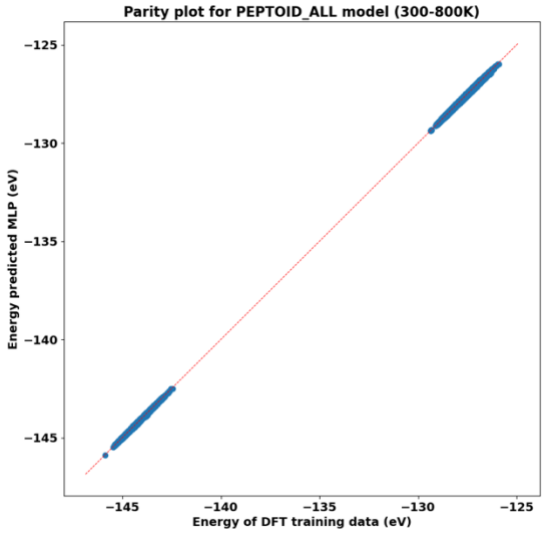
Supplementary Figure 1: Parity plots of the best and worst models. The All model refers to the isomeric model. The worst models are the $C_{\alpha x}$ and sarc. The best models are the All and peptoid.






Parity plot for CAX model 300K
Energy predicted MLP (eV)
−128.8
−128.9
−129.0
−129.1
−129.2
−129.3
−129.4
−129.5
−129.6
−129.6 −129.5 −129.4 −129.3 −129.2 −129.1 −129.0 −128.9 −128.8
Energy of DFT training data (eV)


Parity plot for ALL model (300-800K)
Energy predicted MLP (eV)
−125
−130
−135
−140
−145
−145 −140 −135 −130 −125
Energy of DFT training data (eV)



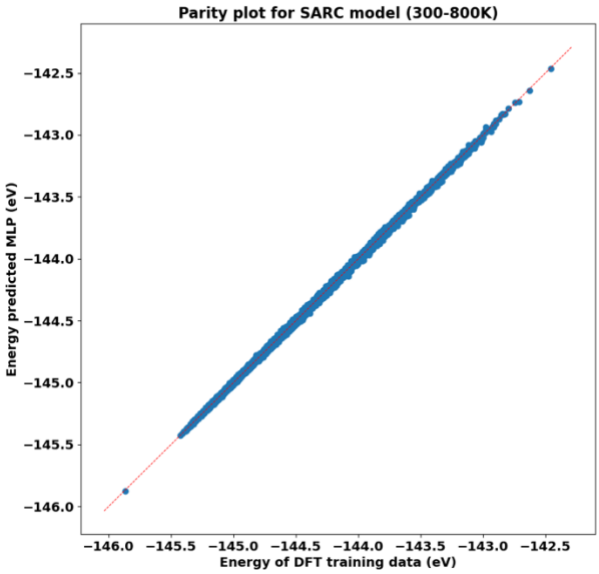


Table S1: All KL divergence values for the 300K models.

	SYSTEM													
Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine
c7eq	1.706354853	7.579490648	13.35279247	12.1218	10.84396	10.845546	8.38468	14.9611	15.034	9.69891	10.4528634	4.64292275	12.2205544	10.5366362
cax	13.35870269	0.974515488	13.79243378	10.6844	13.49913	13.118078	12.520888	14.9837	15.0416	13.7495	12.1723018	7.16660909	13.3771848	11.9970569
cisA	11.39186251	12.64513608	7.692894835	9.09124	11.65483	11.650554	5.1548595	14.777	14.9733	11.6659	11.0697492	12.0184993	10.7135184	11.0035142
cisAd	13.54454428	8.477989171	12.27762411	6.82059	5.46347	10.221548	10.412287	13.3809	13.1649	13.6607	10.7424593	11.0112667	10.2487697	10.4182135
cisc7B	13.75222625	6.975918465	4.738463038	9.1888	6.453686	11.007446	8.4017164	8.67636	14.6315	12.9009	9.67270817	10.3640724	9.0140027	9.31401818
sarc	12.64528845	6.493255437	6.575264497	9.18341	6.684412	4.9930138	11.246279	11.2945	11.3328	8.07185	8.85200587	9.56927194	8.75852406	8.93869025
transA	3.198366039	14.23836183	13.56803721	13.009609	7.508038	8.9642502	1.5800911	1.68712	1.72673	13.7495	7.92301079	8.71836393	6.86341049	7.27562237
transAd	7.937552644	14.05319007	13.20555278	6.10932	10.31644	9.1466582	2.035022	3.69211	1.58738	10.7284	7.88116147	10.9953714	6.58463997	7.5648025
transc7B	8.175072804	13.36693621	9.814831686	10.8759	8.534079	8.0012421	1.8395589	1.71674	6.59097	9.48237	7.83976824	10.7710045	6.76761525	7.65725731
mAmine	1.46733314	1.192734421	14.04191735	10.8737	12.60922	8.1567761	2.4967153	14.9563	14.3612	6.89929	8.70552372	1.33003378	11.0708404	8.9062167
tmAmine	1.609425663	2.095939778	14.02854961	10.4066	13.30111	12.564563	3.7317628	14.844	5.92375	8.51243	8.70181306	1.85268272	10.6857615	8.72285512
amide	1.571213157	2.25345344	12.22899869	11.2225	13.04723	9.7787161	8.7793085	4.12046	5.99092	10.7605	7.97533777	1.9123333	9.30973819	7.66587043
Peptide Only	1.85346861	1.680370351	13.57378458	9.49648	13.30529	12.455845	11.714644	14.7075	14.1754	13.1582	10.6121054	1.76691948	12.7755682	10.3292018
Peptoid Only	3.719608147	1.121598961	12.65540065	9.46052	8.471057	3.7326973	1.2117673	2.87928	1.5457	6.8544	5.16520294	2.42060355	5.70806037	4.97751441
all	1.68743936	0.874243874	5.003525013	8.93443	6.19902	6.0385106	1.4480528	2.01023	2.60989	7.11616	4.19215046	1.28084162	4.60623684	3.86726012
MACE	1.607299386	1.123853926	8.162587374	10.0343	8.348426	7.1299682	0.390543	2.16955	1.29871	5.82628	4.60914673	1.36557666	5.36200516	4.47390993
	6.20160987	5.94668676	10.91954111	9.84460081	9.764963405	9.237838264	5.709260984	8.80354879	8.74930996	10.1772077				

Table S2: all KL divergence values for the 300-800K models.

	SYSTEM														
	Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine
ALL MODEL	c7eq	2.122442925	0.991076423	12.94947678	12.5259	12.733	13.3307	8.04194	14.5126	6.90679	7.70491	9.18187726	1.55675967	11.5714774	9.34598456
	cax	1.61111226	0.69641259	13.78258143	12.5968	12.9437	12.1096	1.68199	8.08002	2.26783	4.29815	7.00682178	1.15376243	9.06607741	7.30778519
	cisA	12.39514586	4.266873553	12.10053058	9.13756	5.93494	6.14901	10.482	10.8919	5.28711	8.27592	8.4920937	8.33100971	8.56899938	8.51611279
	cisAd	13.5204804	0.811564279	9.853771428	7.97924	5.86598	6.3262	8.70945	9.65454	14.9556	12.5018	9.01786003	7.16602234	9.04924993	8.63075491
	cisc7B	13.91133047	5.287963411	12.73787573	9.1924	6.09997	8.59367	5.50883	9.16955	11.492	13.5274	9.55209982	9.59964694	8.97061339	9.11039862
	sarc	14.1462924	14.75286757	3.482344122	9.21618	10.7628	3.87801	13.8659	11.7022	13.705	5.34274	10.0854371	14.44958	9.51606723	10.6124034
	transA	3.042160886	13.40350156	5.506377263	8.90454	9.54805619	8.48773	3.52349	2.47848	1.57849	11.0187	6.74915331	8.22283122	5.71816556	6.27475793
	transAd	2.407194948	14.11264144	4.460661239	10.1054	9.1519	7.48207	2.13197	3.2231	2.29499	13.0796	6.84495127	8.25991819	5.550017	6.15221727
	transc7B	5.375238296	11.58664144	6.880810664	8.3372	10.701	7.96049	1.34954	1.38406	1.1023	12.8312	6.75084659	8.48093987	5.38791575	6.07525444
	mAmine	1.714110397	0.732823725	13.16902597	11.7424	12.4563	11.1424	9.22389	2.04326	3.14366	5.16524	7.0533122	1.22346706	8.98870627	7.26309756
	tmAmine	2.770356844	0.542679352	8.88803107	13.4842	12.5124	9.91012	3.14501	13.4878	4.53274	6.96828	7.62416001	1.6565181	9.42289773	7.69703559
	Peptide Only	2.169339362	1.364516216	13.32552577	13.5609	13.5141	12.8632	1.67514	14.8985	15.0398	6.5351	9.49462266	1.76692779	12.1253245	9.8234586
	Peptoid Only	7.640479062	1.182227192	3.530883172	9.2544	5.90681	4.03155	1.26765	1.93884	1.25251	7.32726	4.33326065	4.41135313	3.88323377	4.00059363
	all	1.745138728	0.958189189	11.2898487	9.03933	5.97506	5.1873	1.56814	1.37059	1.5914	3.71789	4.24428913	1.35166396	5.14595313	4.30277776
	MACE	1.607299386	1.123853926	8.162587374	10.0343	8.34843	7.12997	0.39054	2.16955	1.29871	5.82628	4.60914673	1.36557666	5.36200516	4.47390993
		5.745208148	4.787588791	9.341355419	10.3407187	9.49696758	8.30546162	4.837698	7.13366382	5.76326322	8.27469617				

Table S3: change in KL divergence going from 300K to 300-800K models.

		SYSTEM														
	Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine	Model Change
ALL MODEL	c7eq	−0.4161	6.58841	0.40332	−0.4041	−1.889	−2.4851	0.34274	0.44847	8.12726	1.994	1.27099	3.08616	0.64908	1.19065	1.27098612
	cax	11.7476	0.2781	0.00985	−1.9124	0.5554	1.00848	10.8389	6.90368	12.7738	9.45135	5.16548	6.01285	4.31111	4.68927	5.16548
	cisA	−1.0033	8.37826	−4.4076	−0.0463	5.71989	5.50154	−5.3271	3.8851	9.68616	3.38994	2.57766	3.68749	2.14452	2.4874	2.57765551
	cisAd	0.02406	7.66642	2.42385	−1.1587	−0.4025	3.89535	1.70284	3.7264	−1.7906	1.15887	1.7246	3.84524	1.19952	1.78746	1.72459927
	cisc7B	−0.1591	1.68796	−7.9994	−0.0036	0.35372	2.41378	2.89289	−0.4932	3.13955	−0.6265	0.12061	0.76443	0.04339	0.20362	0.12060834
	sarc	−1.501	−8.2596	3.09292	−0.0328	−4.0784	1.115	−2.6196	−0.4077	−2.3722	2.72911	−1.2334	−4.8803	−0.7575	−1.6737	-1.2334312
	transA	0.15621	0.83486	8.06166	4.10507	−2.04	0.47652	−1.9434	−0.7914	0.14824	2.73079	1.17386	0.49553	1.14524	1.00086	1.17385748
	transAd	5.53036	−0.0595	8.74489	−3.9961	1.16454	1.66459	−0.0969	0.46901	−0.7076	−2.3512	1.03621	2.73545	1.03462	1.41259	1.0362102
	transc7B	2.79983	1.78029	2.93402	2.53869	−2.1669	0.04075	0.49002	0.33268	5.48867	−3.3488	1.08892	2.29006	1.3797	1.582	1.08892165
	mAmine	−0.2468	0.45991	0.87289	−0.8687	0.15294	−2.9856	−6.7272	12.913	11.2176	1.73404	1.65221	0.10657	2.08213	1.64312	1.65221152
	tmAmine	−1.1609	1.55326	5.14052	−3.0776	0.78873	2.65445	0.58675	1.35621	1.39101	1.54415	1.07765	0.19616	1.26286	1.02582	1.07765305
	Peptide Only	−0.3159	0.31585	0.24826	−4.0645	−0.2088	−0.4073	10.0395	−0.191	−0.8644	6.62314	1.11748	−8E−06	0.65024	0.50574	1.11748274
	Peptoid Only	−3.9209	−0.0606	9.12452	0.20613	2.56424	−0.2988	−0.0559	0.94044	0.29318	−0.4729	0.83194	−1.9907	1.82483	0.97692	0.83194229
	all	−0.0577	−0.0839	−6.2863	−0.1049	0.22396	0.85121	−0.1201	0.63964	1.01849	3.39827	−0.0521	−0.0708	−0.5397	−0.4355	-0.0521387
	MACE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Avg System		0.81974447	1.50569306	1.59738055	-0.6299784	0.05269044	0.96034084	0.71452838	2.12366909	3.3963634	1.99673839	1.25371702	1.16271877	1.17357062	1.17115909	

Table S4: ω angles at 300K.

	SYSTEM														
	Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine
300 MODEL	c7eq	1.249534129	0.86436949	14.46606276	13.6877	13.7865	3.04415	15.5271	1.99266	2.93656	13.5996	8.11542773	1.05695181	9.34868489	7.50607754
	cax	0.584514643	0.729750268	14.54449853	14.0135	13.334	13.8681	15.5271	8.4603	13.3504	13.7328	10.8145047	0.65713246	13.2997062	10.4902454
	cisA	0.680435737	1.240382753	1.640491925	1.43852	1.98791	1.16926	15.5271	1.27581	0.72093	2.28691	2.79677793	0.96040925	3.39429344	2.85343028
	cisAd	0.470095894	0.95030173	1.694551696	1.5004	1.27555	0.51451	15.5271	0.8773	0.94136	0.66768	2.44188829	0.71019881	3.19011568	2.63902304
	cisc7B	1.187807169	1.684649938	0.960410564	1.29826	1.45596	0.79403	15.5271	0.73168	0.77782	2.03629	2.6454056	1.43622855	3.07790107	2.71308495
	sarc	1.246615479	0.676322945	0.908071593	1.51378	2.49751	1.15507	15.5271	1.28235	0.69828	1.73042	2.72355548	0.96146921	3.36888572	2.83390427
	transA	1.113214919	10.37650078	5.440190628	2.19504	1.86291	0.58944	15.5271	0.46435	0.44451	11.2713	4.92846165	5.74485785	3.78908316	4.22369976
	transAd	0.674304895	0.83417579	2.446980688	3.87523	1.78236	0.68066	15.5271	0.52254	0.72647	4.00294	3.10728048	0.75424034	3.65162662	3.007763
	transc7B	6.20854188	11.97497693	2.874706398	2.76341	1.58062	0.54875	15.5271	0.4852	0.77493	2.00307	4.47413671	9.0917594	3.5078248	4.74869915
	mAmine	0.775144218	1.137051242	13.19766051	13.723	12.9216	11.9644	15.5271	12.2419	12.7068	1.07451	9.52692665	0.95609773	13.1832231	10.4660841
	tmAmine	0.936756201	1.344740648	12.41614493	10.8325	10.6338	12.9331	15.5271	13.4269	11.902	1.13943	9.10925122	1.14074842	12.5245128	9.99478736
	Peptide Only	0.916223817	0.857861191	14.49726155	13.9725	13.75	12.8325	15.5271	7.43632	8.35262	13.2886	10.1431015	0.8870425	12.3383272	9.79359723
	Peptoid Only	0.538222393	0.909813666	2.796989007	1.68192	1.56746	0.53604	15.5271	0.8004	0.48886	1.00871	2.58555547	0.72401803	3.34268738	2.76076086
	all	1.374749727	0.81061088	1.197220804	2.38471	1.49843	0.62028	15.5271	0.52895	0.5313	0.81232	2.52857132	1.0926803	3.18400516	2.7192663
	MACE	1.079331096	0.893244162	1.19295989	1.76822	2.32713	0.9218	15.5271	0.96802	0.63358	1.37089	2.66823166	0.98628763	3.33412144	2.81238059
		1.269032813	2.352316827	6.018280098	5.77657696	5.48411731	4.14480714	15.5271427	3.43297786	3.73243065	4.66836858				
												cisAd			
	Peptides on Peptoids	Peptoids on Peptides	Amines on Peptides	Amines on Peptoids								All			
	11.32419556	2.808451917	1.048423077	12.85386794								Peptoid			
	Peptide on Peptide	Peptoids on Peptoids										cisc7B			
	0.857042132	3.425675783										MACE			

Table S5: ω angles 300-800K.

	SYSTEM														
	Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine
ALL MODEL	c7eq	0.920965391	0.600109308	14.53157481	14.2827	13.9581	13.2557	15.5271	10.8877	8.76893	14.2215	10.695434	0.7605373	13.030258	10.303653
	cax	0.690003894	0.69401703	14.47720576	13.903	14.0307	13.2384	15.5271	11.5259	6.46483	14.0882	10.463942	0.6920105	12.738171	10.061246
	cisA	0.84312485	0.744052537	1.918175581	1.85472	1.41673	0.89438	15.5271	0.74918	0.60072	0.91496	2.5463196	0.7935887	3.2801512	2.7275818
	cisAd	1.425790717	1.770827887	1.386293942	2.02514	1.91832	0.52976	15.5271	0.46443	0.48946	2.40914	2.7946289	1.5983093	3.1915048	2.8374614
	cisc7B	1.055376041	1.714278989	2.119030074	1.89627	2.0694	0.5131	15.5271	0.94389	0.7959	2.35245	2.8986836	1.3848275	3.4092478	2.9593766
	sarc	6.885943057	7.353866758	1.825175431	2.45365	1.55928	0.65205	15.5271	1.38965	0.68477	1.78478	4.0116313	7.1199049	3.4416742	4.2590588
	transA	1.407876282	9.027799191	1.503682127	1.65576	1.3801	1.4047	15.5271	0.7437	0.59338	1.9719	3.5216052	5.2178377	3.2583535	3.6937944
	transAd	1.242097574	9.82920486	1.241295883	7.69012	2.55548	1.27581	15.5271	0.61382	0.40458	5.6947	4.6074242	5.5356512	4.186892	4.4866162
	transc7B	0.960011366	8.920186065	8.466485974	2.70067	5.38911	2.27237	15.5271	0.69738	0.44204	1.8531	4.7228499	4.9400987	5.0707433	5.0417112
	mAmine	0.89134573	0.885478701	14.47720194	12.529	13.6582	7.82201	15.5271	3.44148	2.11357	14.0882	8.543361	0.8884122	9.9383695	7.9272679
	tmAmine	1.047983143	0.88256867	14.39965689	14.2675	14.0307	13.6789	15.5271	7.51229	3.75345	14.2659	9.9366166	0.9652759	11.881391	9.4555874
	Peptide Only	0.594549309	0.851153213	14.54004041	14.2657	13.7758	13.8765	15.5271	12.8218	10.3167	14.0882	11.065773	0.7228513	13.589118	10.729948
	Peptoid Only	0.887517322	0.790137336	1.586248228	2.38322	1.57665	0.41872	15.5271	0.62468	0.44766	1.29354	2.5535513	0.8388273	3.2234741	2.6935526
	all	1.814667757	1.409855311	3.047395064	1.53833	1.37339	0.77439	15.5271	0.87241	0.39078	1.52527	2.8273641	1.6122615	3.3605498	2.9720413
	MACE	1.079331096	0.893244162	1.19295989	1.76822	2.32713	0.9218	15.5271	0.96802	0.63358	1.37089	2.6682317	0.9862876	3.3341214	2.8123806
		1.449772235	3.091118668	6.4474948	6.3476017	6.067943	4.7685752	15.527143	3.6170912	2.4600274	6.1281773				
												cisA			
												Peptoid			
	Peptides on Peptoids	Peptoids on Peptides	Amines on Peptides	Amines on Peptoids								MACE			
	12.88421446	3.798602584	0.926844061	10.90988012								cisAd			
	Peptide on Peptide	Peptoids on Peptoids										All			
	0.726273906	3.691223827													

Table S6: W2 results for all models trained at 300K.

	Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine
300 MODEL	c7eq	1.1402	165.9702	29.1436	35.1143	46.0932	48.1983	97.8208	103.17	112.621	59.9194	69.91912	83.5552	67.45163	71.0302
	cax	173.1627	1.4725	107.8371	40.1175	186.588	148.021	132.979	119.368	237.034	252.318	139.8897	87.3176	138.8491	127.3976
	cisA	200.3516	110.0153	134.3146	9.4869	157.523	220.002	93.0394	85.223	119.138	221.46	135.0554	155.1835	116.961	125.4549
	cisAd	198.645	108.8822	14.0174	19.8716	13.9602	201.559	52.9752	25.2367	89.3715	245.197	96.97165	153.7636	59.57027	80.50212
	cisc7B	188.8419	32.5669	4.9917	12.2968	2.9994	167.425	33.437	18.2963	124.842	194	77.96962	110.7044	52.04113	65.07741
	sarc	142.0214	129.6757	26.8425	10.8086	65.4944	10.4354	89.2322	88.516	120.907	134.825	81.87581	135.8486	58.8908	75.99252
	transA	93.7985	18.2321	32.962	44.1852	46.0291	66.6328	1.0572	4.3785	1.6037	231.115	53.99939	56.0153	28.12121	34.3199
	transAd	188.2938	159.5671	35.6317	6.213	147.166	103.721	0.8157	50.8378	1.4353	69.5316	76.32132	173.9305	49.40296	77.07573
	transc7B	57.0262	17.4661	14.5467	26.1058	111.856	26.0268	34.9968	7.7138	81.3233	117.566	49.46275	37.24615	43.22413	41.89569
	mAmine	2.3108	63.6962	137.3383	29.6007	118.258	55.1489	5.4408	120.039	84.9498	35.446	65.22281	33.0035	78.68216	68.53134
	tmAmine	1.8714	1.1438	131.4158	81.3741	143.4308	94.6896	13.1743	142.934	131.394	43.5586	78.49863	1.5076	105.4875	82.38086
	Peptide Only	1.078	67.666	32.8178	61.529	26.1094	64.2693	156.653	117.107	163.585	72.2268	76.30404	34.372	88.86709	76.75707
	Peptoid Only	57.0949	1.5678	29.8352	13.8913	10.8555	6.6516	2.117	27.3529	11.3622	42.1623	20.28907	29.33135	14.58081	17.85871
	all	1.5122	13.4695	5.7604	8.278	30.5087	2.8553	4.1236	14.1975	39.1962	34.2188	15.41202	7.49085	14.98853	13.32238
	MACE	2.4808	4.462	13.6049	21.6764	11.008	5.2965	20.1269	15.2915	3.1995	33.8009	13.09474	3.4714	12.88624	10.79406
		87.30862667	59.72356	50.07064667	28.03661	74.52533	81.3955	49.19919	62.64411	94.19726	119.1564				
	Peptide on Peptoids	Peptoids on Peptides	Amines on Peptides	Amines on Peptoids											
	103.15035	117.5274143	17.25555	92.08482857											

Table S7: W2 results for all models trained from 300-800K.

					SYSTEM										
	Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine
ALL MODEL	c7eq	5.3968	2.0011	26.2812	38.349	48.0181	132.636	17.0877	99.6287	60.2985	58.3438	48.80411	3.69895	60.32848571	47.7441444
	cax	2.144	12.2795	146.6956	30.3927	29.951	132.406	2.2096	29.1943	53.8462	29.505	46.86235	7.21175	60.67071429	48.7909444
	cisA	178.6382	39.2109	7.8792	6.5814	18.8757	14.5637	141.294	102.057	68.5621	85.9319	66.35943	108.92455	51.4019	64.1847111
	cisAd	173.3951	2.9842	11.5634	10.7068	1.8102	30.015	148.179	38.9822	149.548	194.62	76.1804	88.18965	55.82925714	63.0204556
	cisc7B	185.1961	6.3142	92.9231	6.4747	13.5458	83.8272	53.8451	120.871	139.862	143.847	84.6706	95.75515	73.04985714	78.0954778
	sarc	179.1203	148.8337	8.0144	12.6383	152.241	4.1956	142.897	106.659	132.967	67.3617	95.49277	163.977	79.94457143	98.6184444
	transA	74.4803	19.0036	3.7627	12.7015	11.6144	5.7161	0.7162	11.1284	2.3644	194.964	33.64516	46.74195	6.857671429	15.7208444
	transAd	60.4503	29.2525	6.7143	23.1513	43.3195	4.608	0.8932	20.1411	31.4348	269.344	48.93087	44.8514	18.60888571	24.4405556
	transc7B	137.5451	11.2641	12.5824	10.3903	125.839	3.3252	1.8107	9.3882	4.8717	264.249	58.12658	74.4046	24.02968571	35.2241111
	mAmine	1.8366	8.7477	42.2698	150.176	121.611	131.931	67.5537	12.0039	54.7937	52.0044	64.29282	5.29215	82.90564286	65.6582
	tmAmine	0.83	2.6272	29.1131	140.551	32.3122	93.8582	62.7445	165.748	96.4749	64.431	68.86895	1.7286	88.6859	69.3620556
	Peptide Only	29.4806	1.3484	60.1063	126.591	109.605	166.994	8.6781	152.468	183.736	54.3814	89.33897	15.4145	115.4541857	93.2231444
	Peptoid Only	95.0883	1.5959	6.3622	6.9672	14.0285	3.1069	1.8771	15.1807	42.5005	134.695	32.14024	48.3421	12.86044286	20.7452556
	all	0.9093	1.7449	34.8202	5.0387	6.4303	2.2323	5.6249	10.941	11.2924	44.6396	12.36736	1.3271	10.9114	8.78155556
	MACE	2.4808	4.462	13.6049	21.6764	11.008	5.2965	20.1269	15.2915	3.1995	33.8009	13.09474	3.4714	12.88624286	10.7940556
		75.13278667	19.44466	33.51285333	40.1590733	49.34736	54.3141267	45.0358933	60.6455467	69.0500733	112.80786				
	Peptide on Peptoids	Peptoids on Peptides	Amines on Peptides	Amines on Peptoids											
	60.4996	88.97775714	3.510375	85.79577143											

Table S8: Change in W2 going from 300K to 300-800K.

						SYSTEM										
	Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine	Model Change
ALL MODEL	c7eq	-4.2566	163.969	2.8624	-3.2347	-1.9249	-84.438	80.7331	3.5413	52.3227	1.5756	21.115	79.8563	7.12314	23.2861	21.11501
	cax	171.019	-10.807	-38.859	9.7248	156.637	15.6153	130.769	90.1735	183.188	222.813	93.0273	80.1059	78.1784	78.6067	93.02731
	cisA	21.7134	70.8044	126.435	2.9055	138.648	205.439	-48.255	-16.834	50.5755	135.528	68.696	46.2589	65.5591	61.2702	68.69601
	cisAd	25.2499	105.898	2.454	9.1648	12.15	171.544	-95.204	-13.746	-60.176	50.5775	20.7913	65.574	3.74101	17.4817	20.79125
	cisc7B	3.6458	26.2527	-87.931	5.8221	-10.546	83.5973	-20.408	-102.58	-15.02	50.1528	-6.701	14.9493	-21.009	-13.018	-6.70098
	sarc	-37.099	-19.158	18.8281	-1.8297	-86.746	6.2398	-53.665	-18.143	-12.06	67.4637	-13.617	-28.128	-21.054	-22.626	-13.61696
	transA	19.3182	-0.7715	29.1993	31.4837	34.4147	60.9167	0.341	-6.7499	-0.7607	36.1508	20.3542	9.27335	21.2635	18.5991	20.35423
	transAd	127.844	130.315	28.9174	-16.938	103.847	99.113	-0.0775	30.6967	-30	-199.81	27.3905	129.079	30.7941	52.6352	27.39045
	transc7B	-80.519	6.202	1.9643	15.7155	-13.984	22.7016	33.1861	-1.6744	76.4516	-146.68	-8.6638	-37.158	19.1944	6.67158	-8.66383
	mAmine	0.4742	54.9485	95.0685	-120.58	-3.3532	-76.782	-62.113	108.035	30.1561	-16.558	0.92999	27.7114	-4.2235	2.87314	0.92999
	tmAmine	1.0414	-1.4834	102.303	-59.176	111.119	0.8314	-49.57	-22.814	34.9187	-20.872	9.62968	-0.221	16.8016	13.0188	9.62968
	Peptide Only	-28.403	66.3176	-27.289	-65.062	-83.496	-102.73	147.974	-35.361	-20.151	17.8454	-13.035	18.9575	-26.587	-16.466	-13.03493
	Peptoid Only	-37.993	-0.0281	23.473	6.9241	-3.173	3.5447	0.2399	12.1722	-31.138	-92.533	-11.851	-19.011	1.72037	-2.8865	-11.85117
	all	0.6029	11.7246	-29.06	3.2393	24.0784	0.623	-1.5013	3.2565	27.9038	-10.421	3.04466	6.16375	4.07713	4.54082	3.04466
	MACE															
Avg System		13.0455429	43.1559643	17.7404929	-12.98835	26.9763929	29.0157571	4.46067857	2.14131429	20.4435714	6.80200714	15.0793371	28.1007536	12.5414082	15.9990405	

Table S9: MAE Energy and Forces for Universal Models at 300K.

Models 300K	E (MAE) eV (10^{-3})	F (MAE) eV/Å (10^{-3})
Isomer	9.9	12.0
Peptide PES	5.5	8.7
Peptoid PES	7.1	11.0
Amide	5.6	9.1
Methylamine	3.4	8.8
Trimethylamine	4.2	8.7

Table S10: MAE Energy and Forces for Universal Models from 300-800K.

Models All K	E (MAE) eV (10^{-3})	F (MAE) eV/Å (10^{-3})
Isomer	15.3	19.2
Peptide PES	8.7	13.0
Peptoid PES	11.2	16.1
Amide	8.7	14.0
Methylamine	7.8	16.0
Trimethylamine	9.0	15.2

Table S11: KL divergence of the energies for models trained at 300K.

						SYSTEM									
	Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine
300 MODEL	c7eq	13.65948276	1.403572888	8.210202264	0.80183	1.00258	2.0726	23.8087	1.97796	18.8272	26.9077	9.86719103	7.53152782	8.10016141	7.97379839
	cax	2.237556389	18.51077918	0.16199822	0.53959	0.00174	0.38831	6.25452	0.19055	4.55631	26.907	5.97483701	10.3741678	1.72757518	3.6490402
	cisA	28.28458402	28.30232936	11.12182446	20.7888	11.0938	28.2259	9.6443	0.10695	4.10099	26.9045	16.8574	28.2934567	12.1546542	15.7410548
	cisAd	28.28355563	28.25202583	4.219927255	14.956	12.1038	27.318	7.45589	0.0864	5.03163	26.9046	15.4611724	28.2677907	10.1673695	14.1896853
	cisc7B	28.28614377	28.30232574	2.898909433	16.9932	7.94649	28.2226	9.20337	0.10182	4.68696	26.9047	15.3546543	28.2942348	10.0076225	14.0713141
	sarc	28.27537136	28.29953979	26.3383014	26.6718	28.1168	6.08765	28.1992	28.0107	28.2554	26.8969	25.5151738	28.2874556	24.525706	25.3616504
	transA	16.0127822	5.098339394	2.132343295	0.33772	2.19478	28.2278	4.18399	27.4693	12.9728	26.9053	12.5535208	10.5555608	11.0741059	10.9588737
	transAd	10.59526195	6.121672018	0.089616917	0.26478	0.28969	28.2276	12.041	21.2931	20.4243	26.9044	12.6251417	8.35846699	11.8042908	11.0385522
	transc7B	0.741972066	7.110284789	0.084180583	0.66011	0.19938	28.2271	13.5426	26.473	13.9097	26.9047	11.7853042	3.92612843	11.8708665	10.1053691
	mAmine	4.953205717	2.294273634	0.864345488	0.20446	0.11305	12.586	1.72673	16.5086	2.0278	26.9051	6.81835406	3.62373968	4.86156208	4.58649043
	tmAmine	11.55457073	21.97028223	28.4006633	28.3444	28.1175	28.2274	27.1281	16.2447	28.0602	26.9025	24.4950268	16.7624265	26.3604201	24.2275326
	Peptide Only	3.974563255	8.794181008	0.377693219	1.42566	1.24325	4.29363	0.70198	2.26423	1.40697	26.9085	5.13906927	6.38437213	1.67334406	2.72023919
	Peptoid Only	1.881106353	25.77581058	1.860567683	23.9775	17.1235	10.6175	4.3417	25.5085	16.9541	26.9065	15.4946739	13.8284585	14.3404679	14.226688
	all	10.15729095	22.0628681	6.975289083	20.9138	14.2987	21.1205	7.88522	19.3237	5.8484	26.9063	15.5492012	16.1100795	13.766511	14.287304
	MACE	19.07919418	16.0685038	27.74006745	25.7957	24.9086	22.5059	23.3493	3.43529	19.3918		20.25272	17.573849	21.0181117	20.25272
		13.86510942	16.55778589	8.098395337	12.1783529	9.91691186	18.4232268	11.964444	12.5996503	12.4303109	26.9049194				
	Peptide on Peptoids	Peptoids on Peptides	Amines on Peptides	Amines on Peptoids											
	4.913868294	19.42615628	10.19308308	15.6109911											

Table S12: KL divergence of the energies for models trained from 300-800K.

		SYSTEM													
	Models	c7eq	cax	cisA	cisAd	cisc7B	sarc	transA	transAd	transc7B	Valine	Average	Peptide Only	Peptoid Only	No Valine
ALL MODEL	c7eq	6.74987201	4.000719763	27.0583557	17.1398459	15.571	28.0602904	28.0997214	20.8691111	28.227	26.906	20.26826	5.375296	23.57508	19.53068
	cax	6.39236405	9.254070888	2.81470546	0.33634983	0.5904	3.6357165	17.1574165	1.83009715	22.079	26.907	9.099719	7.823217	6.92047	7.121081
	cisA	28.2856408	28.30237045	4.83027596	8.33654549	15.429	28.2251706	12.2273235	0.44972048	8.6852	26.905	16.16764	28.29401	11.16902	14.97457
	cisAd	22.9667203	15.66810478	2.2080942	17.0678665	10.294	28.2274925	0.42008174	8.15492981	1.6035	26.904	13.35157	19.31741	9.710916	11.84569
	cisc7B	28.2856257	28.30239662	5.44600106	25.3040837	15.649	28.2259945	11.731961	0.66463784	11.133	26.904	18.16465	28.29401	14.02205	17.1936
	sarc	28.2793685	28.29810999	28.3992911	28.3430929	28.115	11.1910586	28.1982002	28.009728	28.255	26.893	26.39825	28.28874	25.78738	26.34323
	transA	0.76662508	2.445817025	23.0171285	14.0766244	18.923	28.2276904	6.60956682	23.1609363	17.361	26.904	16.14932	1.606221	18.76804	14.9543
	transAd	3.06872669	6.413434563	18.0043287	2.953673	2.6014	28.2275835	4.77447343	25.7528804	8.2599	26.903	12.69599	4.741081	12.93917	11.11738
	transc7B	4.76366328	3.975527557	2.37913572	3.22796288	2.6574	28.227694	10.1307624	26.5375191	7.2099	26.903	11.60127	4.369595	11.48148	9.901061
	mAmine	14.8158415	9.6728889	6.16769119	2.18754789	5.9288	28.2215723	18.1003332	1.52398758	17.87	26.904	13.13926	12.24437	11.42852	11.60982
	tmAmine	7.2816806	8.028334889	28.3940191	27.7433496	27.567	28.2256888	28.1998597	26.8689426	28.257	26.901	23.74666	7.655008	27.89364	23.39616
	Peptide Only	15.485206	14.72876841	21.8782242	5.34683487	8.748	28.0715824	27.0945407	12.5680021	27.69	26.907	18.85178	15.10699	18.77101	17.95679
	Peptoid Only	4.69214777	27.07811353	3.33203164	11.1147364	15.584	13.9837453	5.68701402	23.2108581	13.461	26.907	14.50507	15.88513	12.3391	13.12711
	all	14.6513352	8.994686939	3.556562	18.760802	12.22	21.0164734	2.72812496	27.1621117	16.792	26.906	15.27888	11.82301	14.60525	13.98698
	MACE	19.0791942	16.0685038	27.7400675	25.7957441	24.909	22.5059061	23.3493472	3.43528645	19.392	26.902	20.91766	17.57385	21.01811	20.25272
		13.70426744	14.08212321	13.68172747	13.84900396	13.65254	23.61824397	14.96724845	15.34658325	17.08496	26.90394				
		Peptide on Peptoids	Peptoids on Peptides	Amines on Peptides	Amines on Peptoids										
		15.24777287	16.41586652	9.949686483	19.66108014										

Supplementary Figure 2: Input script of the model training used for NequIP.

```
root: results/model
run_name: cax_300
seed: 123
dataset_seed: 456
append: true
default_dtype: float32
allow_tf32: false
# device:  cuda

model_builders:
 - SimpleIrrepsConfig
 - EnergyModel
 - PerSpeciesRescale
 - ForceOutput
 - RescaleEnergyEtc

# network
r_max: 5.0
num_layers: 6

l_max: 1
parity: true
num_features: 32

nonlinearity_type: gate
resnet: false

nonlinearity_scalars:
  e: silu
  o: tanh
```

```
nonlinearity_gates:
  e: silu
  o: tanh

# radial network basis
num_basis: 8
BesselBasis_trainable: true
    1.  PolynomialCutoff_p: 6


# radial network
invariant_layers: 2
invariant_neurons: 64
avg_num_neighbors: auto
use_sc: true

dataset: npz
dataset_file_name: 300cax_vasp.npz
key_mapping:
  z: atomic_numbers
  E: total_energy
  F: forces
  R: pos
  pbc: pbc
  cell: cell
npz_fixed_field_keys:
  - atomic_numbers
  - cell
  - pbc

chemical_symbols:
 - H
 - C
 - N
 - O

wandb: true
wandb_project: cax

verbose: info
log_batch_freq: 10
log_epoch_freq: 1
save_checkpoint_freq: -1
save_ema_checkpoint_freq: -1
```

```
# training
n_train: 7000
n_val: 2000
learning_rate: 0.005 #0.001
batch_size: 2
max_epochs: 1000
train_val_split: random
shuffle: true
metrics_key: validation_loss
use_ema: true
ema_decay: 0.99
ema_use_num_updates: true
report_init_validation: true


# Early Stopping
early_stopping_patiences:
  validation_loss: 50

early_stopping_delta:
  validation_loss: 0.005

early_stopping_cumulative_delta: false

early_stopping_lower_bounds:
  LR: 1.0e-5

early_stopping_upper_bounds:
  wall: 1.0e+100

# loss function
loss_coeffs:
  forces: 1
  total_energy:
    - 1
    - PerAtomMSELoss

# output metrics
metrics_components:
  - - forces
    - mae
  - - forces
    - rmse
```

```
    - PerSpecies: True
      report_per_component: False
  - - forces
    - mae
    - PerSpecies: True
      report_per_component: False
  - - total_energy
    - mae
  - - total_energy
    - mae
    - PerAtom: True

optimizer_name: Adam
optimizer_amsgrad: true
optimizer_betas: !!python/tuple
  - 0.9
  - 0.999
optimizer_eps: 1.0e-08
optimizer_weight_decay: 0

max_gradient_norm: null


lr_scheduler_name: ReduceLROnPlateau
lr_scheduler_patience: 100
lr_scheduler_factor: 0.5


per_species_rescale_scales_trainable: false

per_species_rescale_shifts_trainable: false

per_species_rescale_shifts: dataset_per_atom_total_energy_mean
per_species_rescale_scales: dataset_forces_rms
```