Towards Accurate and Efficient Anharmonic Vibrational Frequencies with the Universal Interatomic Potential ANI-1ccx-gelu and Its Fine-Tuning

Seyedeh Fatemeh Alavi,[1] Yuxinxin Chen,[1] Yi-Fan Hou,[1] Fuchun Ge,[1] Peikun Zheng,[1,2] and Pavlo O. Dral[1,3*]

[1]*State Key Laboratory of Physical Chemistry of Solid Surfaces, Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, Department of Chemistry, and College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China.*

[2]*Present affiliation: Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States.*

[3]*Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, ul. Grudziądzka 5, 87-100 Toruń, Poland.*

Email: dral@xmu.edu.cn

Abstract

Calculating anharmonic vibrational modes of molecules for interpreting experimental spectra is one of the most interesting challenges of contemporary computational chemistry. However, the traditional QM methods are costly for this application. Machine learning techniques have emerged as a powerful tool for substituting the traditional QM methods. Universal interatomic potentials (UIPs) hold a particular promise to deliver accurate results at a fraction of the cost of the traditional QM methods but the performance of UIPs for calculating anharmonic vibrational frequencies remains hitherto unknown. Here we show that despite a known excellent performance of the representative UIP ANI-1ccx for thermochemical properties, it fails for the anharmonic frequencies due to the original unfortunate choice of the activation function. Hence, we recommend evaluating new UIPs on anharmonic frequencies as an additional important quality test. To remedy the shortcomings of ANI-1ccx, we introduce its reformulation ANI-1ccx-gelu with the GELU activation function which is capable of calculating IR anharmonic frequencies with reasonable accuracy (close to B3LYP/6-31G*). We also show that our new UIP can be fine-tuned to obtain very accurate anharmonic frequencies for some specific molecules but more effort is needed to improve the overall quality of UIP and its capability for fine-tuning. The new UIP will be included as part of our universal and updatable AI-enhanced QM methods (UAIQM) platform and is available together with tutorials in open-source MLatom at https://github.com/dralgroup/mlatom. The calculations can also be performed via web browser at https://XACScloud.com.

Introduction

Vibrational spectroscopy is a prominent area of research in contemporary molecular physics and theoretical chemistry that serves as a highly sensitive tool for molecule identification and the monitoring of conformational and structural transformations, both in the gas phase and in solution. It plays a vital role in detecting various species in the interstellar medium and the atmosphere[1-4] with applications in stellar classification, air quality monitoring, and identification of air pollutants. It is also widely used to study reactions at interfaces[5], including functional interfaces of actively developed technologies such as heterogeneous catalysis for chemical synthesis or decomposition[6, 7] or for energy conversion[8, 9], solar cells[10, 11], and electrochemical batteries[12, 13].

During the last 50 years, computational vibrational spectroscopy has gone through great development and has become a valuable tool for the interpretation and prediction of spectra[14]. A crucial task in the practical application of vibrational spectroscopy is the assignment of a measured frequency to its corresponding type of molecular motion. In the frequency range between 1200 and 1700 cm^{-1}, assignments are particularly challenging, because of couplings between the different degrees of freedom.

The most straightforward approach to evaluate the vibrational behavior of a molecule is through the use of the harmonic approximation, in which the potential energy function is taken as a truncated, second-order Taylor series. Calculating vibrational spectra within the harmonic approximation typically yields data that enable a semi-quantitative comparison with experimental results. However, this approximation does not consider anharmonic effects and mode coupling, which can significantly impact the accuracy of the results, especially when the molecule deviates from its equilibrium geometry. Consequently, interpreting experimental spectra may lead to notable errors. It is crucial to acknowledge that molecular potential energy surfaces (PESs) are inherently anharmonic, and this characteristic is manifested in the vibrational spectrum of the studied molecular system. Predictions of accurate anharmonic frequencies, which compare sufficiently well with experiment, remain a challenge to overcome and allow assignments of the vibrations and interpretation of spectroscopic features, especially for high-frequency modes[15-19].

Any of the above approaches for vibrational frequencies require the calculations using QM methods. Conventional spectroscopic computer techniques based on methods derived from quantum mechanics can be very time-consuming when using highly accurate methods such as

coupled cluster with single, double, and perturbative triple excitations, CCSD(T)[20-23], and beyond[24-28]. On the other side, there are density functional theory (DFT) methods that generally are far from spectroscopic accuracy but often provide a decent semi-quantitative agreement with the experiment[29].

Many efforts have been made to develop faster and more accurate QM methods and this research field is still active, but the traditional QM method development takes a lot of time and ultimately comes with little improvement in results. Machine Learning (ML) techniques have emerged as a powerful tool for supplementing the traditional QM method development allowing for rapid improvements[30, 31]. ML enables calculations with both high accuracy and very low computational cost. Although significant progress has been made in this field, the majority of ML applications in quantum chemistry are still rather limited to particular uses. Numerous studies using ML for calculating anharmonic vibrational frequencies have also been conducted[32-39]. These calculations require accurate, full-dimensional potential energy surfaces (PESs) for which ML methods have gained much attention[40]. ML potentials are used to generate statistical models that aim to identify an unbiased predicting function that optimally correlates a set of molecular structures with the given target energies and forces used as training data. These ML methods used to generate PESs can be classified into kernel-based approaches[41] (e.g., KREG[42], GAP[43]-SOAP[44], sGDML [45]) and neural network (NN) approaches (e.g., PhysNet[46], ANI[47], DPMD[48], BPNN[49], SchNet[50, 51], MEGNet[52], and MACE[53, 54]).

Numerous studies have been conducted on the calculation of anharmonic frequencies and spectra using various methods.[55-65] Recent studies illustrate a practical approach to calculate anharmonic frequencies from accurate, machine-learned potentials for small- (four atoms) to medium-sized (nine atoms) molecules, where the representation of the PESs is based on a neural network (NN) using the PhysNet architecture[59, 60]. In these studies, the second-order vibrational perturbation theory (VPT2) is a popular tool for calculating anharmonic frequencies which can be done with the ML-PESs used as an external energy function.[59, 60] Also, in these studies, transfer learning (TL) was used to achieve more accurate PESs by fine-tuning the PhysNet potential on additional data from higher levels of quantum chemical theory.[59]

The primary obstacle that persists is developing flexible ML methodologies that can successfully generalize to various scenarios. The above studies employed the ML interatomic potential specifically fitted for each molecule. It is, however, highly desirable to use universal interatomic potentials (UIP) which can generalize to any molecule. There is a surge in the number of UIPs with most of them targeting DFT-quality PESs.[66-71] One of the earliest

successful UIPs is ANI-1ccx[72] which is addressing the challenge of achieving PESs with CCSD(T)-level accuracy while maintaining transferability and extensibility across a wide range of chemical processes. ANI-1ccx was demonstrated to have excellent, near-chemical accuracy for predicting thermochemical properties which also involve the frequency calculations.[73] Hence, one would expect it must have a decent performance for anharmonic frequencies too. In this work, we show that, surprisingly, this is not the case.

To remedy this problem, here, we revise the ANI-1ccx and introduce the new UIP ANI-1ccx-gelu which has comparable performance to ANI-1ccx in terms of excellent speed and good accuracy for many applications. Particularly, for the anharmonic frequencies, ANI-1ccx-gelu produces decent, close-to-DFT results.

We also show that a particular advantage of ANI-1ccx-gelu is its cost-efficient fine-tuning for specific molecules which might achieve the desired level of accuracy in anharmonic frequencies, e.g., close to the CCSD(T) level. The remaining challenges with the use of ANI-1ccx-gelu and its fine-tuning are discussed too.

Results and discussion

Evaluation of the coupled-cluster-level UIP ANI-1ccx on anharmonic frequencies

ANI-1ccx is known to approach coupled-cluster-level accuracy for many properties. The reported scheme for the accurate calculation of the enthalpies of formation has employed the rigid-rotor, harmonic oscillator approximations for calculating zero-point vibrational frequencies and thermal corrections which were sufficient to nearly reach chemical accuracy. To put it in perspective, ANI-1ccx outperformed common DFT methods such as B3LYP/6-31G* and semi-empirical methods specifically fitted on enthalpies of formation. Hence, one might expect that such a high accuracy of ANI-1ccx would translate into at least decent performance for anharmonic frequencies.

Here we test this hypothesis and perform the VPT2 calculations with the ANI-1ccx for the HCOOH molecule (see *Methods*). Surprisingly, the errors in anharmonic frequencies compared to the experiment are huge: mean absolute error (MAE) is ca.1500 cm^{-1} with many negative frequencies (Table 1). This is far worse than DFT exemplified by B3LYP/6-31G* (here and in the following) or MP2 which have MAEs in the order of dozens of wavenumbers (Table 1).

We trace the underlying issue to the inappropriate choice of the activation function which leads to numerical instabilities for higher-order derivatives required for VPT2 anharmonic

frequencies. The possible solution is then to employ another activation function. Indeed, when we replace CELU with the GELU activation function and recalculate the anharmonic frequencies with the resulting ANI-1ccx (GELU) variant, the unphysical negative frequencies disappear and the MAE drops to ca. 680 cm^{-1} (Table 1). Similarly, for other molecules (CH_2O, HNO_2, CH_3OH, CH_3CHO, CH_3COOH, CH_3NO_2, and CH_3CONH_2) replacing CELU with the GELU led to the increase in accuracy and the elimination of negative frequencies (except CH_3COOH, see Table S3, Table **S6**, Table **S8**, Table **S10**, Table **S12**,Table **S14**).

Table 1. VPT2 anharmonic frequencies (in cm^{-1}) of HCOOH calculated using UIPs ANI-1ccx (with the default CELU activation function and after changing to the GELU activation function) and ANI-1ccx-gelu compared to the *ab initio* MP2 values[59] as well as with experiment[74] and a representative DFT method (B3LYP/6-31G*).[a]

mode	ANI-1ccx (CELU)	ANI-1ccx (GELU)	ANI-1ccx-gelu	DFT	MP2	Exp
1	572.95	451.205	601.02	620.51	619.54	626.16
2	-4292.22	394.215	639.48	676.56	641.78	640.72
3	1075.44	541.456	1022.52	1033.34	1036.38	1033.47
4	-1250.98	533.909	1137.08	1115.83	1098.18	1104.85
5	6955.88	641.05	1317.02	1265.34	1220.63	1306.20
6	1206.03	812.656	1407.21	1403.86	1380.94	1380.00
7	-2263.56	915.278	1797.65	1820.40	1760.53	1776.83
8	694.75	1799.38	3020.50	2924.46	2967.80	2942.00
9	2324.11	2089.36	3617.04	3476.54	3554.90	3570.50
MAE	1568.75	689.136	28.16	30.27	17.94	
RMSE	2535.00	793.4	35.61	40.43	30.90	

[a]Negative frequencies were removed when evaluating errors as is done in literature[59].

New UIP ANI-1ccx-gelu fixing the problems of ANI-1ccx

Despite the improvement in anharmonic frequencies, ANI-1ccx (GELU) still performs much worse than DFT or MP2. The possible reason is that in ANI-1ccx (GELU) we kept the ANI-1ccx (CELU) weights, which obviously might not be optimal for this activation function. Hence, here we refitted the weights using the procedure analogous to that employed for fitting ANI-1ccx. In brief, we first fitted the network weights to ca. 4.5 M DFT-level data and then fine-tuned some of them to 0.5 M coupled cluster-level data. We name the resulting universal interatomic potential ANI-1ccx-gelu. It consists of an ensemble of eight models as ANI-1ccx, which improves the accuracy of the model.

Our new UIP ANI-1ccx-gelu predicts the anharmonic frequencies of the HCOOH molecule with good accuracy: MAE is 28.16 cm^{-1} which is even somewhat better than B3LYP/6-31G* (30.27 cm^{-1}) but still worse than MP2 (17.94 cm^{-1}). We also tested the new UIP on a set of small molecules with available experimental anharmonic frequencies and it has a decent accuracy for many of them (Table 2). To put it in perspective, we also report the DFT errors for the same set of molecules and see that ANI-1ccx-gelu has an accuracy similar to DFT while both of the methods may have big errors for some molecules. Only one molecule ANI-1cc-gelu has catastrophically bad anharmonic frequencies (CH_3NO_2, MAE of 1373.65 cm^{-1}). It means that there is still room for improvement for the UIP as the deficiencies are likely to be due to insufficient training data.

Table 2. MAEs in VPT2 anharmonic frequencies (in cm^{-1}) of different molecules at ANI-1ccx-gelu and DFT (B3LYP/6-31G*) with respect to the experiment.[a]

Molecule	MAE (ANI-1ccx-gelu)	MAE (DFT)
HCOOH	28.16	30.27
H_2CO	317.24	45.75
HNO_2	18.52	50.45
CH_3OH	49.72	123.51
CH_3CHO	18.73	16.86
CH_3COOH	225.65	18.62
CH_3NO_2	1373.65	17.61
CH_3CONH_2	78.94	43.80

[a]Negative frequencies at both ANI-1ccx-gelu and DFT were removed when evaluating errors, as is done in literature[59]. See Supporting Information for the full lists of frequencies at each level in Table S3, Table **S6**Table S8Table S10Table S12, and Table **S14**.

In addition, ANI-1ccx-gelu retains a performance similar to that of ANI-1ccx and popular semi-empirical GFN2-xTB[75] and B3LYP[76]/6-31G* methods for the standard GMTKN55 benchmark[77] (its CHNO-containing subset with neutral and closed-shell species due to the limitations of the training set and model architecture, Figure 1).

Subsets	N	ANI-1ccx	ANI-1ccx-gelu	GFN2-xTB	B3LYP-D4/ 6-31G*
Basic properties and reaction energies for small systems					
G2RC	12	24.59	34.40	29.52	10.66
FH51	44	6.59	10.64	11.21	5.75
TAUT15	13	1.56	1.41	0.99	1.59
DC13	10	17.98	42.31	26.61	5.99
WTMAD-2		7.79	12.05	10.19	5.50
Reaction energies for large systems and isomerisation reactions					
DARC	14	1.20	1.50	17.44	2.08
BSR36	36	2.69	0.77	2.76	3.14
CDIE20	20	2.09	1.77	1.80	1.95
ISO34	34	1.24	1.37	6.90	2.89
ISOL24	16	2.83	5.55	14.59	3.65
C60ISO	9	91.96	49.34	5.80	2.55
WTMAD-2		5.41	4.00	8.64	4.83
Reaction barrier heights					
BHPERI	22	8.89	18.05	9.49	4.02
BHDIV10	5	13.15	13.33	10.83	4.22
INV24	13	10.81	15.37	2.63	1.68
BHROT27	19	1.27	1.11	0.91	0.75
PX13	8	7.97	11.00	3.65	8.34
WCPT18	12	6.01	5.66	3.83	6.49
WTMAD-2		7.27	10.87	5.25	3.53
Intermolecular noncovalent interactions					
ADIM6	6	0.64	0.17	1.15	0.47
S22	22	5.00	6.29	0.76	2.58
S66	66	2.52	3.56	0.73	2.10
WATER27	15	1.17	0.42	1.50	55.36
WTMAD-2		9.99	13.30	2.86	9.34
Intramolecular noncovalent interactions					
IDISP	6	12.39	16.15	6.78	1.82
ACONF	15	0.62	0.60	0.19	0.19
Amino20x4	72	0.91	1.13	0.98	0.71
PCONF21	18	3.18	4.03	1.76	1.88
MCONF	51	0.92	1.52	1.72	1.37
SCONF	17	1.58	2.25	1.64	3.66
BUT14DIOL	64	0.61	0.30	1.25	2.44
WTMAD-2		9.64	11.46	9.97	12.15
Total WTMAD-2		8.33	10.27	7.93	8.31

Figure 1. Performance of the GELU-improved universal ANI potential compared with ANI-1ccx, GFN2-xTB, and B3LYP/6-31G* on GMTKN55. Only neutral closed-shell CHNO-containing compounds are selected due to the limited coverage of training data. WTMAD-2 – weighted mean absolute deviation-2 in kcal/mol as defined in Ref. 77.

Efficient fine-tuning of UIP ANI-1ccx-gelu

The anharmonic frequencies computed with our new UIP ANI-1ccx-gelu show that they reach the DFT-level accuracy, much better than ANI-1ccx potential (Table 1). However, we often need higher accuracy, ideally closer to the spectroscopic accuracy (errors of ca. 1 cm^{-1}). Here

we explore whether it is possible to achieve this by fine-tuning our new UIP ANI-1ccx-gelu model to a new problem: a specific molecule and/or level of theory.

For this, we performed transfer learning by refitting part of the ANI-1ccx-gelu model parameters to CCSD(T)-F12, CCSD(T), and MP2 energies and forces for different conformers of HCOOH molecule (data taken from a previous study[77]). The generated VPT2 anharmonic frequencies by ANI-1ccx-gelu fine-tuned on MP2 data (ANI-1ccx-gelu/TL@MP2) show very good agreement with the reference MP2 results as the MAE is only 1.79 cm^{-1} (Table 3, see *Methods* for details). Importantly, this error is smaller than the error of the ANI model with GELU activation function trained on the same data from scratch (MAE of 3.03 cm^{-1}), i.e., fine-tuning is more accurate and a bit faster (see Table 3 for errors and Table 4 for timings). This demonstrates an excellent quality that may be achieved by fine-tuning to the desired level for a single molecule.

Table 3. VPT2 anharmonic frequencies (in cm^{-1}) of HCOOH calculated using TL UIP ANI-1ccx-gelu trained on MP2, CCSD(T), and CCSD(T)-F12 data. They are compared to their reference ab initio MP2 values[59] and experimental values[74].

mode	ANI-1ccx-gelu/ TL@ MP2	ANI	MP2	ANI-1ccx-gelu/ TL@ CCSD(T)	ANI-1ccx-gelu/ TL@ CCSD(T)-F12	Exp
1	619.78	621.88	619.54	619.11	620.18	626.16
2	642.25	642.10	641.78	632.32	637.63	640.72
3	1039.08	1038.64	1036.38	1029.55	1033.24	1033.47
4	1096.95	1085.76	1098.18	1099.21	1103.54	1104.85
5	1218.50	1221.13	1220.63	1293.70	1299.64	1306.20
6	1377.10	1382.86	1380.94	1375.13	1376.04	1380.00
7	1763.24	1764.45	1760.53	1767.72	1778.76	1776.83
8	2966.44	2967.09	2967.80	2930.51	2933.28	2942.00
9	3556.36	3557.77	3554.90	3562.76	3571.96	3570.50
MAE_{Exp}	18.24	18.67	17.94	7.86	3.69	
$RMSE_{Exp}$	31.30	30.92	13.90	8.31	4.57	
MAE_{MP2}	1.79	3.03				
$RMSE_{MP2}$	2.10	4.63				

Table 4. The computational cost for training or fine-tuning a single model on the HCOOH molecule PES using the RTX 2080Ti GPU. The range is shown for three repeats.

Models	Number of epochs	Time per epoch, s
ANI	565–800	8.5–9.0
ANI-1ccx-gelu (TL)	420–685	8.3–8.4
PhysNet	–[a]	186.2–196.1

[a]The original study[59] did not report the number of epochs needed to achieve the reported results, here we only evaluated the time per epoch on our hardware.

However, both MP2 and ANI-1ccx-gelu/TL@MP2 have MAE of ca. 18 cm^{-1} compared to the experiment. Hence, we look next to how the model fine-tuned on more accurate coupled cluster-level data performs. Indeed, ANI-1ccx-gelu/TL@CCSD(T)-F12 has an outstanding accuracy with MAE of only 3.69 cm^{-1} compared to experiment, while ANI-1ccx-gelu/TL@CCSD(T) trained on a slightly less accurate data shows MAE of 7.86 cm^{-1}. This demonstrates that the fine-tuning of ANI-1ccx-gelu can, in principle, successfully adjust to the target level for a specific molecule.

Nevertheless, an analysis of fine-tuning of ANI-1ccx-gelu to MP2 and coupled-cluster-level (where available) data for a range of molecules uncovers the remaining challenges (Table 5). Our results clearly show that fine-tuning works better for molecules that were already described well with the universal model (i.e., compare Table 2 and Table **5**). For such well-described molecules (particularly, HCOOH and HNO_2), the performance of fine-tuned ANI-1ccx-gelu is on par with that of the more complex and computationally costly (see Table 4 and discussion below) PhysNet model, i.e., MAEs are well below 10 cm^{-1}. However, for other molecules, errors can be significantly higher, although the comparison to experimental values indicates that better ML architecture does not always lead to better results. For example, the CH_3OH molecule is described with MAE of ca. 15 cm^{-1} with both ANI-1ccx-gelu and PhysNet trained on CCSD(T) data; similarly, for H_2CO both approaches produce MAE of the same order of magnitude (ca. 10 cm^{-1}). This might be due to several factors such as uncertainty of experimental data, inherent error of the VPT2 approach, insufficient accuracy of the reference coupled-cluster-level data, fortuitous error cancellation in some cases, insufficient coverage of the required PES regions in the training data, and the fact that ANI-1ccx-gelu was pre-trained on coupled-cluster data that might make it easier to be fine-tuned on the data at a similar level. In any case, the varying performance may reflect that the data used to train the universal model is of crucial importance for downstream tasks such as fine-tuning on a specific task. An

additional factor to consider is the distribution of the data used for fine-tuning: here we took the data that was specifically developed to train the PhysNet model, which generally shows excellent accuracy on these data, particularly when compared to the MP2 level. However, there is growing evidence that a different type of model may need a different set of data, particularly, if the model is not as accurate as PhysNet or state-of-the-art equivariant networks like MACE.[78, 79]

Table 5. MAEs in VPT2 anharmonic frequencies (in cm^{-1}) of different molecules at ANI-1ccx-gelu/TL@MP2, CCSD(T), and CCSD(T)-F12 with respect to MP2 and the experiment.[a]

Molecule	TL@ MP2 wrt MP2	TL@ CCSD(T) wrt exp	TL@ CCSD(T)-F12 wrt exp	PhysNet@ MP2 wrt MP2	PhysNet@ CCSD(T) wrt exp	PhysNet@ CCSD(T)-F12 wrt exp
HCOOH	1.79	7.86	3.69	1.55, 1.77	9.06, 9.47	3.59, 3.97
H_2CO	42.00	13.87	76.26	0.95, 2.03	9.04, 10.66	3.98, 5.28
HNO_2	2.41	5.04	6.29	1.05, 1.30	5.56, 6.14	6.74, 7.14
CH_3OH	14.25	15.53		0.92, 1.01	14.30, 15.07	
CH_3CHO	12.60			0.88, 0.72		
CH_3COOH	51.97			1.47, 1.49		
CH_3NO_2	62.36			2.12, 1.14		
CH_3CONH_2	28.91			2.71, 3.88		

[a]Negative frequencies at both ANI-1ccx-gelu and DFT were removed when evaluating errors, as is done in literature[59]. See Supporting Information for the full lists of frequencies at each level in Table S4, Table **S5**, Table **S7**, Table **S9**,Table **S11**, Table **S13**, Table **S15**, and Table **S17**.

However, when deciding which model to choose, an important factor to consider is the computational cost.[80] Table 4 provides timing for training models including ANI (i.e., training from scratch), TL ANI-1ccx-gelu (i.e., fine-tuning), and PhysNet on the HCOOH molecule PES using a single GPU. For ANI and ANI-1ccx-gelu (TL) models, the wall-clock times for each training epoch ranged from ca. 8 to 9 seconds, meanwhile, this value for PhysNet models is ca. 200 seconds which is at least 20 times slower than for ANI. Number of epochs needed to train the models also varies. Considering the number of epochs for each training, fine-tuning ANI-1ccx-gelu is also generally faster than training ANI from scratch. Although the original study[59] did not report the number of epochs needed to train PhysNet, our attempts to reproduce the reported PhysNet performance showed that the number of epochs must be very large, thousands of epochs, to achieve high accuracy. This amounts to one to two orders of magnitude more expensive training of PhysNet compared to ANI models. Based on this analysis of the computational cost, it is apparent that fine-tuning ANI-1ccx-gelu is the most cost-effective

while also rather effective, and, hence, it would be very beneficial to further improve this UIP's robustness for such downstream tasks as calculating accurate anharmonic frequencies.

Conclusions

We studied the calculation of the anharmonic frequencies with the VPT2 approach using the universal interatomic potentials. We found that ANI-1ccx despite its excellent accuracy for thermochemical properties is not suitable for predicting the accurate anharmonic frequencies which we traced down to the deficiencies in the activation function CELU. Hence, we introduced the reformulation of ANI-1ccx with a more suitable activation function GELU. The resulting universal model ANI-1ccx-gelu has good, DFT-level, accuracy for anharmonic frequencies while also providing decent performance for chemically relevant simulations. ANI-1ccx-gelu is, of course, more computationally efficient than quantum chemical approaches for calculating anharmonic frequencies and can be used for molecules with more atoms. Overall, our new UIP ANI-1ccx-gelu potential is an attractive alternative to DFT and semi-empirical quantum mechanical approaches. We will be included ANI-1ccx-gelu in our collection of universal and updatable AI-enhanced QM (UAIQM) models[81] giving the benefit that it can be selected automatically based on our reported uncertainty quantification procedure[73].

We also showed that ANI-1ccx-gelu can be fine-tuned to produce a high accuracy potential for special tasks but the performance strongly depends on the system. We attribute this dependence to the data distribution used to train the universal model. It indicates that more effort is needed to consistently achieve beyond-DFT quality for anharmonic frequencies. This is particularly pertinent considering the outstanding computational speed for making predictions with and fine-tuning ANI-1ccx-gelu compared to more complex and accurate machine learning interatomic potential architectures.

Methods

Computational details

All simulations were performed with the development versions of the MLatom program[82] interfaced to the TorchANI package[83] for training and predictions with the ANI-based models, the PhysNet package[46] for calculations with the PhysNet potential[46], and Gaussian 16 for geometry optimizations and anharmonic frequency calculations. The data set used is taken from Meuwly's group[59]. MP2 level in the paper denotes the MP2/aug-cc-pVDZ, CCSD(T) –

CCSD(T)/AVTZ and CCSD(T)-F12 – CCSD(T)-F12/AVTZ-F12 With the reference data generated using MOLPRO[59].

Anharmonic vibrational frequencies

The vibrational anharmonic frequencies are computed using VPT2 via MLatom's interface to the Gaussian 16 software,[84] utilizing ML PES as an external potential. VPT2 is employed to incorporate anharmonic effects and mode coupling into spectroscopic properties[85]. VPT2 assumes that the potential energy of a system can be expressed as a quartic force field, given by

$$V = \frac{1}{2}\sum \phi_{ijk}\hat{q}_i^2 + \frac{1}{3i}\sum \phi_{ijk}\,\hat{q}_i\hat{q}_j\hat{q}_k + \frac{1}{4!}\sum \phi_{ijk}\,\hat{q}_i\hat{q}_j\hat{q}_k\hat{q}_l. \quad (1)$$

Here, ω_i is a harmonic frequency, q_i are normal mode coordinates, and ϕ_{ijk} and ϕ_{ijkl} are third- and fourth-order derivatives of the potential V, with respect to normal mode coordinates, respectively. The first term in Eq. 1 is the harmonic part of the potential and the remaining terms describe anharmonic effects. The cubic and quartic force constants can be used to obtain anharmonic constants.[59] The anharmonic fundamental frequencies ν_i are obtained from the anharmonic constants χ according to

$$\nu_i = \omega_i + 2\chi_{ii} + \frac{1}{2}\sum_{i\neq j}\chi_{ij}\,. \quad (2)$$

Fine-tuning ANI-1ccx-gelu

ANI-1ccx-gelu was fine-tuned on CCSD(T)-F12, CCSD(T), and MP2 data across HCOOH, H_2CO, HNO_2, CH_3OH, CH_3CHO, CH_3COOH, CH_3NO_2, and CH_3CONH_2 molecules with 5401, 3601, 6404, 7201, 10073, 10910, 9001, and 12601 number of data points, respectively, taken from the data set reported earlier[59]. CCSD(T) data were available only for HCOOH, HNO_2, H_2CO, and CH_3OH. CCSD(T)-F12 data were available only for HCOOH, HNO_2, and H_2CO. Transfer learning is used with different weights fixed for different molecules and data based on the best performance of generated anharmonic frequencies (Table S18). Table S19 shows an example of the effects of different fixed layers on the performance of generated anharmonic frequencies for HCOOH. A sample of different models is trained until the limit of the learning rate (10^{-5}) is reached and the epoch's model with the best performance on the validation set is selected as the final model.

Training ANI from scratch

The dataset was randomly split according to 85%/10%/5% into training/validation/test sets for fitting the ANI-type ML potential (with the GELU activation function) on energies and forces. The ANI model hyperparameters were optimized too. These hyperparameters include batch size, early stopping, number of layers and their size, radial cutoff, and angular cutoff. We also optimized a hyperparameter for energy-weighting the loss function. The calculations and results of the weighting function parameters are explained in Supplementary Note 1, SI which showed that for this data distribution, no energy weighting is needed. We found that the optimized hyperparameters of the NN architecture are basically the same as the default of the TorchANI interface and, hence, we used the default parameters except for the angular cutoff R_{ca} which was set to 3.3 Å. Our tests showed that the generated anharmonic frequencies from different ANI models are a little different from each other. To improve the robustness, the ANI model's results reported in the main text (Table 3) are from the ensemble of nine of the best ANI neural networks trained on the same MP2 data but with different random splits. All of the ANI models in the ensemble were evaluated on separate test sets with the MAEs and root-mean-squared errors (RMSEs) in energies and forces reported in Table S1. MAEs and RMSEs show nine independently trained models can still differ appreciably, although nine models are of relatively high quality.

Supporting information

Out of sample energy and force errors for ANI models, VPT2 anharmonic frequencies obtained from ANI-1ccx (CELU), ANI-1ccx (GELU), ANI-1ccx-gelu (TL) models, and DFT.

Data availability

No new data was generated in this study (we re-used the existing data sets taken from the literature[59]).

Code availability

The open-source code and the ANI-1ccx-gelu model are available in open-source MLatom under the MIT license at https://github.com/dralgroup/mlatom.

Author contributions

P.O.D. designed and supervised the project, evaluated the results of the analysis, edited the first draft of the manuscript, and secured the funding.

S.F.A. performed all the evaluations on the anharmonic frequencies, training and fine-tuning the models for all molecules, and wrote the original manuscript.

Y.C. trained ANI-1ccx-gelu and benchmarked it on the GMTKN55 database.

Y.H. implemented an interface to Gaussian for the VPT2 calculations.

F.G. implemented the fine-tuning via layer fixing in the TorchANI interface.

P.Z. original implementations and tests, and the suggestion of using GELU activation function.

All authors discussed the results and the manuscript.

Acknowledgments

P.O.D. acknowledges funding from the National Natural Science Foundation of China (funding via the Outstanding Youth Scholars (Overseas, 2021) project) and via the Lab project of the State Key Laboratory of Physical Chemistry of Solid Surfaces.

References

(1) Heras, A.; Shipman, R.; Price, S.; de Graauw, T.; Walker, H.; Kessler, M.; Prusti, T.; Decin, L.; Vandenbussche, B.; Waters, L. Infrared spectral classification of normal stars. *Astron. Astrophys.* **2002**, *394* (2), 539-552.

(2) Hachmann, J.; Olivares-Amaya, R.; Jinich, A.; Appleton, A. L.; Blood-Forsythe, M. A.; Seress, L. R.; Román-Salgado, C.; Trepte, K.; Atahan-Evrenk, S.; Er, S. Lead candidates for high-performance organic photovoltaics from high-throughput quantum chemistry–the Harvard Clean Energy Project. *Energ. Environ. Sci.* **2014**, *7* (2), 698-704.

(3) Pyzer-Knapp, E. O.; Suh, C.; Gómez-Bombarelli, R.; Aguilera-Iparraguirre, J.; Aspuru-Guzik, A. What is high-throughput virtual screening? A perspective from organic materials discovery. *Annu. Rev. Mater. Res.* **2015**, *45*, 195-216.

(4) Lopez, S. A.; Pyzer-Knapp, E. O.; Simm, G. N.; Lutzow, T.; Li, K.; Seress, L. R.; Hachmann, J.; Aspuru-Guzik, A. The Harvard organic photovoltaic dataset. *Sci. Data.* **2016**, *3*, 160086. DOI: 10.1038/sdata.2016.86.

(5) Kanal, I. Y.; Hutchison, G. R. Rapid computational optimization of molecular properties using genetic algorithms: Searching across millions of compounds for organic photovoltaic materials. *arXiv preprint arXiv:1707.02949* **2017**.

(6) Vishwakarma, G. Machine Learning Model Selection for Predicting Properties of High Refractive Index Polymers. State University of New York at Buffalo, 2018.

(7) Pradhan, A. Tailor-made Materials: Inverse Engineering Compounds Using Feature Correlation. State University of New York at Buffalo, 2020.

(8) Yoshikawa, N.; Terayama, K.; Sumita, M.; Homma, T.; Oono, K.; Tsuda, K. Population-based de novo molecule generation, using grammatical evolution. *Chem. Phys. Lett.* **2018**, *47* (11), 1431-1434.

(9) Weymuth, T.; Reiher, M. Gradient-driven molecule construction: An inverse approach applied to the design of small-molecule fixating catalysts. *Int. J. Quantum Chem.* **2014**, *114* (13), 838-850.

(10) Kwon, Y.; Kang, S.; Choi, Y.-S.; Kim, I. Evolutionary design of molecules based on deep learning and a genetic algorithm. *Sci. Rep.* **2021**, *11* (1), 17304.

(11) Lee, Y.; Choi, G.; Yoon, M.; Kim, C. Genetic Algorithm for Constrained Molecular Inverse Design. *arXiv preprint arXiv:2112.03518* **2021**.

(12) Leguy, J.; Cauchy, T.; Glavatskikh, M.; Duval, B.; Da Mota, B. EvoMol: a flexible and interpretable evolutionary algorithm for unbiased de novo molecular generation. *J. Cheminformatics* **2020**, *12* (1), 1-19.

(13) Cai, W.; Zhang, M.; Zhang, Y. Batch mode active learning for regression with expected model change. *IEEE T. Neur. Net. Lear.* **2016**, *28* (7), 1668-1681.

(14) Grunenberg, J. *Computational spectroscopy: methods, experiments and applications*; John Wiley & Sons, 2011.

(15) Nejad, A.; Meyer, E.; Suhm, M. A. Glycolic acid as a vibrational anharmonicity benchmark. *J. Phys. Chem. Lett.* **2020**, *11* (13), 5228-5233.

(16) Mata, R. A.; Suhm, M. A. Benchmarking quantum chemical methods: Are we heading in the right direction? *Angewandte Chemie International Edition* **2017**, *56* (37), 11011-11018.

(17) Barone, V.; Biczysko, M.; Bloino, J. Fully anharmonic IR and Raman spectra of medium-size molecular systems: accuracy and interpretation. *Phys. Chem. Chem. Phys.* **2014**, *16* (5), 1759-1787.

(18) Qu, C.; Bowman, J. M. IR spectra of (HCOOH) 2 and (DCOOH) 2: Experiment, VSCF/VCI, and ab initio molecular dynamics calculations using full-dimensional potential and dipole moment surfaces. *Faraday Discuss.* **2018**, *9* (10), 2604-2610.

(19) Qu, C.; Bowman, J. M. Quantum and classical IR spectra of (HCOOH) 2,(DCOOH) 2 and (DCOOD) 2 using ab initio potential energy and dipole moment surfaces. *Faraday Discuss.* **2018**, *212*, 33-49.

(20) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157* (6), 479-483.

(21) Thomas, J. R.; DeLeeuw, B. J.; Vacek, G.; Crawford, T. D.; Yamaguchi, Y.; Schaefer III, H. F. The balance between theoretical method and basis set quality: A systematic study of equilibrium geometries, dipole moments, harmonic vibrational frequencies, and infrared intensities. *J. Chem. Phys.* **1993**, *99* (1), 403-416.

(22) Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. The prediction of molecular equilibrium structures by the standard electronic wave functions. *J. Chem. Phys.* **1997**, *106* (15), 6430-6440.

(23) Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. The accurate determination of molecular equilibrium structures. *J. Chem. Phys.* **2001**, *114* (15), 6548-6556.

(24) Koput, J. Ab initio potential energy surface and vibration-rotation energy levels of lithium monohydroxide. *J. Chem. Phys.* **2013**, *138* (23).

(25) Yachmenev, A.; Yurchenko, S. N.; Ribeyre, T.; Thiel, W. High-level ab initio potential energy surfaces and vibrational energies of H2CS. *J. Chem. Phys.* **2011**, *135* (7).

(26) Owens, A.; Yurchenko, S. N.; Yachmenev, A.; Tennyson, J.; Thiel, W. Accurate ab initio vibrational energies of methyl chloride. *J. Chem. Phys.* **2015**, *142* (24).

(27) Owens, A. A highly accurate potential energy surface for carbonyl sulphide (OCS): how important are the ab initio calculations? *Phys. Chem. Chem. Phys.* **2024**.

(28) Owens, A.; Yurchenko, S. N.; Yachmenev, A.; Tennyson, J.; Thiel, W. A highly accurate ab initio potential energy surface for methane. *J. Chem. Phys.* **2016**, *145* (10).

(29) Husch, T.; Vaucher, A. C.; Reiher, M. Semiempirical molecular orbital models based on the neglect of diatomic differential overlap approximation. *Int. J. Quantum Chem.* **2018**, *118* (24), e25799.

(30) Dral, P. O. Quantum Chemistry in the Age of Machine Learning. Elsevier: Amsterdam, Netherlands, 2023.

(31) von Lilienfeld, O. A.; Müller, K.-R.; Tkatchenko, A. Exploring chemical compound space with quantum-based machine learning. *Nat. Rev. Chem.* **2020**, *4* (7), 347–358. DOI: 10.1038/s41570-020-0189-9.

(32) Kwac, K.; Cho, M. Machine learning approach for describing vibrational solvatochromism. *J. Chem. Phys.* **2020**, *152* (17).

(33) Ye, S.; Zhang, G.; Jiang, J. AI-based spectroscopic monitoring of real-time interactions between SARS-CoV-2 and human ACE2. *P. Natl. Acad. Sci.* **2021**, *118* (26), e2025879118.

(34) Zhang, Y.; Ye, S.; Zhang, J.; Hu, C.; Jiang, J.; Jiang, B. Efficient and Accurate Simulations of Vibrational and Electronic Spectra with Symmetry-Preserving Neural Network Models for Tensorial Properties. *J. Phys. Chem. B* **2020**, *124* (33), 7284–7290. DOI: 10.1021/acs.jpcb.0c06926.

(35) Ye, S.; Zhong, K.; Zhang, J.; Hu, W.; Hirst, J. D.; Zhang, G.; Mukamel, S.; Jiang, J. A Machine Learning Protocol for Predicting Protein Infrared Spectra. *J. Am. Chem. Soc.* **2020**, *142* (45), 19071-19077. DOI: https://doi.org/10.1021/jacs.0c06530 From NLM Medline.

(36) Gastegger, M.; Schutt, K. T.; Muller, K. R. Machine learning of solvent effects on molecular spectra and reactions. *Chem. Sci.* **2021**, *12* (34), 11473–11483. DOI: 10.1039/d1sc02742e From NLM PubMed-not-MEDLINE.

(37) Kananenka, A. A.; Yao, K.; Corcelli, S. A.; Skinner, J. L. Machine Learning for Vibrational Spectroscopic Maps. *J. Chem. Theory Comput.* **2019**, *15* (12), 6850–6858. DOI: 10.1021/acs.jctc.9b00698.

(38) Kwac, K.; Freedman, H.; Cho, M. Machine Learning Approach for Describing Water OH Stretch Vibrations. *J. Chem. Theory Comput.* **2021**, *17* (10), 6353-6365.

(39) Ito, S.; Cui, Q. Multi-level free energy simulation with a staged transformation approach. *J. Chem. Phys.* **2020**, *153* (4).

(40) Unke, O. T.; Koner, D.; Patra, S.; Käser, S.; Meuwly, M. High-dimensional potential energy surfaces for molecular simulations: from empiricism to machine learning. *Mach. Learn.: Sci. Technol.* **2020**, *1* (1), 013001.

(41) Hofmann, T.; Schölkopf, B.; Smola, A. J. Kernel methods in machine learning. **2008**.

(42) Dral, P. O.; Owens, A.; Yurchenko, S. N.; Thiel, W. Structure-based sampling and self-correcting machine learning for accurate calculations of potential energy surfaces and vibrational levels. *J. Chem. Phys.* **2017**, *146* (24), 244108. DOI: 10.1063/1.4989536.

(43) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104* (13), 136403. DOI: 10.1103/Physrevlett.104.136403.

(44) Bartók, A. P.; Kondor, R.; Csányi, G. On representing chemical environments. *Phys. Rev. B* **2013**, *87* (18), 187115. DOI: 10.1103/physrevb.87.184115.

(45) Chmiela, S.; Sauceda, H. E.; Müller, K.-R.; Tkatchenko, A. Towards exact molecular dynamics simulations with machine-learned force fields. *Nat. Commun.* **2018**, *9* (1), 3887. DOI: 10.1038/s41467-018-06169-2.

(46) Unke, O. T.; Meuwly, M. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theory Comput.* **2019**, *15* (6), 3678–3693. DOI: 10.1021/acs.jctc.9b00181.

(47) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **2017**, *8* (4), 3192–3203. DOI: 10.1039/c6sc05720a.

(48) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys. Rev. Lett.* **2018**, *120* (14), 143001. DOI: 10.1103/PhysRevLett.120.143001.

(49) Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 146401. DOI: 10.1103/Physrevlett.98.146401.

(50) Schütt, K. T.; Kindermans, P.-J.; Sauceda, H. E.; Chmiela, S.; Tkatchenko, A.; Müller, K.-R. SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. *Adv. Neural. Inf. Process. Syst.* **2017**, *30*, 992–1002.

(51) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. Schnet–a deep learning architecture for molecules and materials. *J. Chem. Phys.* **2018**, *148* (24).

(52) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chemistry of materials : a publication of the American Chemical Society* **2019**, *31* (9), 3564–3572. DOI: 10.1021/acs.chemmater.9b01294.

(53) Batatia, I.; Kovács, D. P.; Simm, G. N. C.; Ortner, C.; Csányi, G. MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields. In *Advances in Neural Information Processing Systems, https://openreview.net/forum?id=YPpSngE-ZU*, 2022. DOI: https://doi.org/10.48550/arXiv.2206.07697.

(54) Batatia, I.; Batzner, S.; Kovács, D. P.; Musaelian, A.; Simm, G. N. C.; Ortner, C.; Kozinsky, B.; Csányi, G. The Design Space of E(3)-Equivariant Atom-Centered Interatomic Potentials. *arXiv:2205.06643* **2022**. DOI: 10.48550/arXiv.2205.06643.

(55) Manzhos, S.; Ihara, M.; Carrington, T. Machine learning for vibrational spectroscopy. In *Quantum Chemistry in the Age of Machine Learning*, Elsevier, 2023; pp 355-390.

(56) Manzhos, S.; Yamashita, K.; Carrington Jr., T. Using a neural network based method to solve the vibrational Schrodinger equation for H_2O. *Chem. Phys. Lett.* **2009**, *474* (1-3), 217–221. DOI: DOI 10.1016/j.cplett.2009.04.031.

(57) Chan, M.; Carrington, T.; Manzhos, S. Anharmonic vibrations of the carboxyl group in acetic acid on TiO 2: implications for adsorption mode assignment in dye-sensitized solar cells. *Phys. Chem. Chem. Phys.* **2013**, *15* (25), 10028-10034.

(58) Manzhos, S.; Carrington, T. Calculations of vibrational spectra with rectangular collocation: suitability for catalytic systems, tests of PES representations and selection of important regions of space. In *APS March Meeting Abstracts*, 2018; Vol. 2018, p Y02. 009.

(59) Käser, S.; Boittier, E. D.; Upadhyay, M.; Meuwly, M. Transfer Learning to CCSD(T): Accurate Anharmonic Frequencies from Machine Learning Models. *J. Chem. Theory Comput.* **2021**, *17* (6), 3687–3699. DOI: 10.1021/acs.jctc.1c00249 From NLM PubMed-not-MEDLINE.

(60) Kaser, S.; Vazquez-Salazar, L. I.; Meuwly, M.; Topfer, K. Neural network potentials for chemistry: concepts, applications and prospects. *Digit. Discov.* **2023**, *2* (1), 28–58. DOI: https://doi.org/10.1039/D2DD00102K From NLM PubMed-not-MEDLINE.

(61) Alavi, S. F.; Shayesteh, A. Einstein A coefficients for rovibronic lines of the $A^2\Pi \rightarrow X^2\Sigma^+$ and $B^2\Sigma^+ \rightarrow X^2\Sigma^+$ transitions of CaH and CaD. *MNRAS* **2018**, *474* (1), 2-11.

(62) Yousefi, M.; Bernath, P. F. Line lists for AlF and AlCl in the X1Σ+ ground state. *ApJS* **2018**, *237* (1), 8.

(63) Brooke, J. S.; Bernath, P. F.; Western, C. M.; Sneden, C.; Afşar, M.; Li, G.; Gordon, I. E. Line strengths of rovibrational and rotational transitions in the X2Π ground state of OH. *J. Quant. Spectrosc. Radiat. Transfer* **2016**, *168*, 142-157.

(64) Yurchenko, S. N.; Owens, A.; Kefala, K.; Tennyson, J. ExoMol line lists–LVII. High accuracy ro-vibrational line list for methane (CH4). *MNRAS* **2024**, *528* (2), 3719-3729.

(65) Owens, A.; Dooley, S.; McLaughlin, L.; Tan, B.; Zhang, G.; Yurchenko, S. N.; Tennyson, J. ExoMol line lists–XLV. Rovibronic molecular line lists of calcium monohydride (CaH) and magnesium monohydride (MgH). *MNRAS* **2022**, *511* (4), 5448-5461.

(66) Nováček, M.; Řezáč, J. PM6-ML: The Synergy of Semiempirical Quantum Chemistry and Machine Learning Transformed into a Practical Computational Method. *ChemRxiv* **2024**. DOI: https://doi.org/10.26434/chemrxiv-2024-3nwwv-v2.

(67) Chen, Y.; Yan, W.; Wang, Z.; Wu, J.; Xu, X. Constructing accurate and efficient general-purpose atomistic machine learning model with transferable accuracy for quantum chemistry. *arXiv:2408.05932 [physics.chem-ph]* **2024**. DOI: 10.48550/arXiv.2408.05932.

(68) Chen, C.; Ong, S. P. A universal graph deep learning interatomic potential for the periodic table. *Nat. Comput. Sci.* **2022**, *2* (11), 718–728. DOI: 10.1038/s43588-022-00349-3 From NLM Medline.

(69) Choudhary, K.; DeCost, B.; Major, L.; Butler, K.; Thiyagalingam, J.; Tavazza, F. Unified graph neural network force-field for the periodic table: solid state applications. *Digit. Discov.* **2023**, *2* (2), 346–355. DOI: 10.1039/d2dd00096b.

(70) Takamoto, S.; Shinagawa, C.; Motoki, D.; Nakago, K.; Li, W.; Kurata, I.; Watanabe, T.; Yayama, Y.; Iriguchi, H.; Asano, Y.; et al. Towards universal neural network potential for

material discovery applicable to arbitrary combination of 45 elements. *Nat. Commun.* **2022**, *13* (1), 2991. DOI: 10.1038/s41467-022-30687-9 From NLM Medline.

(71) Xie, F.; Lu, T.; Meng, S.; Liu, M. GPTFF: A high-accuracy out-of-the-box universal AI force field for arbitrary inorganic materials. *Sci. Bull.* **2024**. DOI: 10.1016/j.scib.2024.08.039 From NLM Publisher.

(72) Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nat. Commun.* **2019**, *10* (1), 2903. DOI: 10.1038/s41467-019-10827-4.

(73) Zheng, P.; Yang, W.; Wu, W.; Isayev, O.; Dral, P. O. Toward Chemical Accuracy in Predicting Enthalpies of Formation with General-Purpose Data-Driven Methods. *J. Phys. Chem. Lett.* **2022**, *13* (15), 3479–3491. DOI: 10.1021/acs.jpclett.2c00734 From NLM PubMed-not-MEDLINE.

(74) Tew, D. P.; Mizukami, W. Ab initio vibrational spectroscopy of cis-and trans-formic acid from a global potential energy surface. *J. Phys. Chem. A* **2016**, *120* (49), 9815-9828.

(75) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB-An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15* (3), 1652–1671. DOI: 10.1021/acs.jctc.8b01176.

(76) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem. C.* **1994**, *98*, 11623–11627.

(77) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19* (48), 32184–32215. DOI: 10.1039/c7cp04913g From NLM PubMed-not-MEDLINE.

(78) Poltavsky, I.; Charkin-Gorbulin, A.; Puleva, M.; Cordeiro Fonseca, G.; Batatia, I.; Browning, N. J.; Chmiela, S.; Cui, M.; Frank, J. T.; Heinen, S.; et al. Crash Testing Machine Learning Force Fields for Molecules, Materials, and Interfaces: Model Analysis in the TEA Challenge 2023. *ChemRxiv* **2024**. DOI: 10.26434/chemrxiv-2024-ctdm3.

(79) Niblett, S. P.; Kourtis, P.; Magdău, I.-B.; Grey, C. P.; Csányi, G. Transferability of datasets between Machine-Learning Interaction Potentials. *arXiv:2409.05590v1 [physics.chem-ph]* **2024**. DOI: https://doi.org/10.48550/arXiv.2409.05590.

(80) Pinheiro Jr, M.; Ge, F.; Ferré, N.; Dral, P. O.; Barbatti, M. Choosing the right molecular machine learning potential. *Chem. Sci.* **2021**, *12* (43), 14396–14413. DOI: 10.1039/D1SC03564A.

(81) Chen, Y.; Hou, Y.-F.; Isayev, O.; Dral, P. O. Universal and Updatable Artificial Intelligence-Enhanced Quantum Chemical Foundational Models. *ChemRxiv* **2024**. DOI: https://doi.org/10.26434/chemrxiv-2024-604wb.

(82) Dral, P. O.; Ge, F.; Hou, Y.-F.; Zheng, P.; Chen, Y.; Barbatti, M.; Isayev, O.; Wang, C.; Xue, B.-X.; Pinheiro Jr, M.; et al. MLatom 3: A Platform for Machine Learning-Enhanced Computational Chemistry Simulations and Workflows. *J. Chem. Theory Comput.* **2024**, *20* (3), 1193–1213. DOI: 10.1021/acs.jctc.3c01203 From NLM Publisher.

(83) Gao, X.; Ramezanghorbani, F.; Isayev, O.; Smith, J. S.; Roitberg, A. E. TorchANI: A Free and Open Source PyTorch-Based Deep Learning Implementation of the ANI Neural Network Potentials. *J. Chem. Inf. Model.* **2020**, *60* (7), 3408–3415. DOI: 10.1021/acs.jcim.0c00451.

(84) *Gaussian 16 Rev. A.01*; Wallingford, CT, 2016. (accessed.

(85) Nielsen, H. H. The vibration-rotation energies of molecules. *Rev. Mod. Phys.* **1951**, *23* (2), 90.

(86) Ge, F.; Wang, R.; Qu, C.; Zheng, P.; Nandi, A.; Conte, R.; Houston, P. L.; Bowman, J. M.; Dral, P. O. Tell machine learning potentials what they are needed for: Simulation-oriented training exemplified for glycine. *J. Phys. Chem. Lett.* **2024**, *15* (16), 4451-4460.

(87) Herndon, S. C.; Nelson Jr, D. D.; Li, Y.; Zahniser, M. S. Determination of line strengths for selected transitions in the v_2 band relative to the v_1 and v_5 bands of H_2CO. *J. Quant. Spectrosc. Radiat. Transfer* **2005**, *90* (2), 207-216.

(88) Guilmot, J.; Godefroid, M.; Herman, M. Rovibrational parameters for trans-nitrous acid. *J. Mol. Spectrosc.* **1993**, *160* (2), 387-400.

(89) Serrallach, A.; Meyer, R.; Günthard, H. H. Methanol and deuterated species: infrared data, valence force field, rotamers, and conformation. *J. Mol. Spectrosc.* **1974**, *52* (1), 94-129.

(90) Wiberg, K. B.; Thiel, Y.; Goodman, L.; Leszczynski, J. Acetaldehyde: Harmonic frequencies, force field, and infrared intensities. *J. Phys. Chem.* **1995**, *99* (38), 13850-13864.

(91) Goubet, M.; Soulard, P.; Pirali, O.; Asselin, P.; Réal, F.; Gruet, S.; Huet, T. R.; Roy, P.; Georges, R. Standard free energy of the equilibrium between the trans-monomer and the cyclic-dimer of acetic acid in the gas phase from infrared spectroscopy. *Phys. Chem. Chem. Phys.* **2015**, *17* (11), 7477-7488.

(92) Wells, A.; Wilson Jr, E. B. Infra‐Red and Raman Spectra of Polyatomic Molecules XIII. Nitromethane. *J. Chem. Phys.* **1941**, *9* (4), 314-318.

(93) Ganeshsrinivas, E.; Sathyanarayana, D.; Machida, K.; Miwa, Y. Simulation of the infrared spectra of acetamide by an extended molecular mechanics method. *J. Mol. Struct. THEOCHEM* **1996**, *361* (1-3), 217-227.

Supporting Information

for

Towards Accurate and Efficient Anharmonic Vibrational Frequencies with the Universal Interatomic Potential ANI-1ccx-gelu and Its Fine-Tuning

Seyedeh Fatemeh Alavi,[1] Yuxixin Chen,[1] Yi-Fan Hou,[1] Fuchun Ge, Peikun Zheng,[1,2] and Pavlo O. Dral[1,3*]

[1]*State Key Laboratory of Physical Chemistry of Solid Surfaces, Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, Department of Chemistry, and College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China.*

[2]*Present affiliation: Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States.*

[3]*Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, ul. Grudziądzka 5, 87-100 Toruń, Poland.*

Email: dral@xmu.edu.cn

Table of Contents

Supplementary Note 1. Energy weighting.

To generate accurate machine learning potentials to subsequently use them for simulation tasks with better accuracy, we applied the flexible energy weighing functions in training machine learning potentials which gives higher weights for the more important low-energy regions of PESs, as introduced in the previous work[86]. Here, we implemented training of ANI-type NNs with a weighting function to get the best results for anharmonic frequencies which are evaluated based on MAE and RMSE for frequencies.

We used the weighing function, $w(\Delta E)$ that provides larger weights w to points with lower energies as defined by energy ΔE (in Hartree) relative to the global minimum:[86]

$$w(\Delta E) = \max\{[-6(a\Delta E)^5 + 15(a\Delta E)^4 - 10(a\Delta E)^3] + 1, 0\} \qquad (S1)$$

Figure S1 shows the weighting function as the function of the energy ΔE relative to the global minimum. The dependence of the weighting function shape on the parameter a is shown by the colorful curves, and the purple histogram of the ΔE distribution of the data set is also shown.

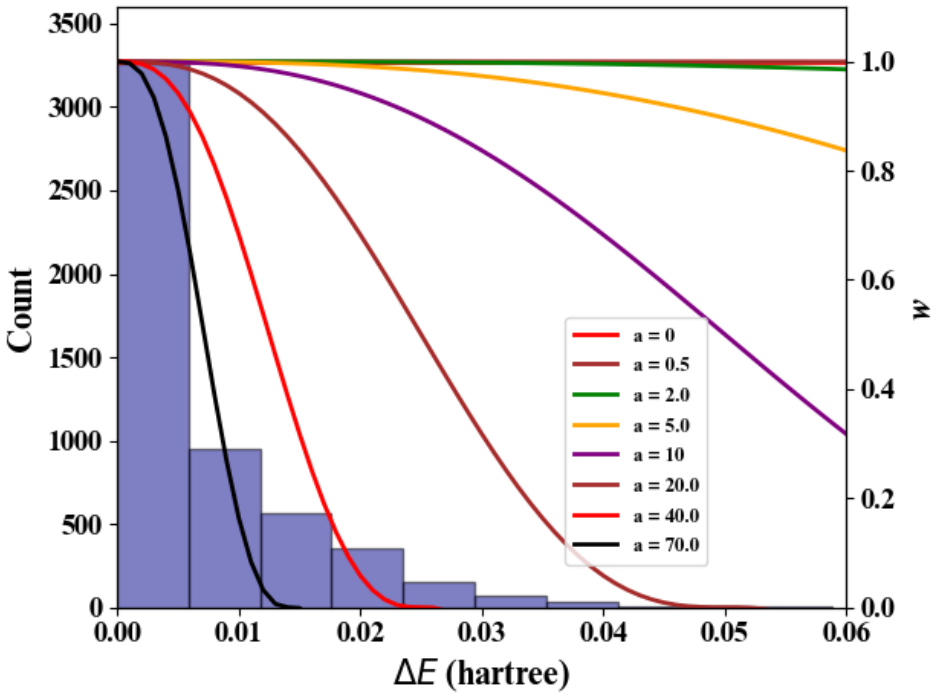

Figure S1. The weighting function as the function of the energy ΔE relative to the global minimum.

The anharmonic frequency calculations with various values of "*a*" parameter in the weighting function formula were run according to the above diagram. Table S1 shows the results of anharmonic frequencies and their corresponding accuracies, as well as the accuracy of the ANI

model (energies and forces as MAE and RMSE (test set)). According to these calculations, the best accuracy of anharmonic frequencies, MAE(v) = 3.73 is when there was no weighting function in calculations. Hence, we did not use the weighting function in any of the reported simulations.

Table S1. Errors in the anharmonic frequencies as a function of the parameter a in Eq. S1.

a	0	2.0	5.0	10	20	30	40	50	60	70	80
MAE	3.73	4.52	12.88	12.85	14.39	4.70	13.30	18.43	6.99	14.62	13.95
RMSE	4.90	7.20	26.43	26.89	28.10	5.94	27.31	33.90	7.69	16.45	16.25

Table S2. Performance measures of nine ANI models trained independently on the same MP2 data and evaluated on the test set. The MAEs and RMSEs are given in kcal/mol (/Å) and multiplied by a factor of 1000 for clarity.

Accuracy/Models	NN1	NN2	NN3	NN4	NN5	NN6	NN7	NN8	NN9
MAE(E)	7.29	7.27	7.35	8.30	7.56	6.93	8.19	8.19	7.56
RMSE(E)	4.35	5.00	5.07	4.58	4.41	3.78	5.04	5.04	5.67
MAE(F)	53.21	63.94	55.29	65.60	66.15	55.44	71.19	62.37	63.63
RMSE(F)	94.25	115.44	92.35	115.68	114.03	103.95	108.99	115.92	154.98

H_2CO

Table S3. VPT2 anharmonic frequencies (in cm^{-1}) of H_2CO calculated using UIPs ANI-1ccx (with the default CELU activation function and after changing to the GELU activation function) and ANI-1ccx-gelu compared to the *ab initio* MP2 values[59] as well as with experiment and a representative DFT method (B3LYP/6-31G*).

mode	ANI-1ccx (CELU)	ANI-1ccx (GELU)	ANI-1ccx-gelu	DFT	MP2	Exp
1	-1122.53	460.07	443.59	1179.50	1180.23	1167.00
2	2358.37	485.16	669.46	1258.43	1246.71	1249.00
3	11.967	528.25	1719.28	1529.47	1507.95	1500.00
4	981.928	812.17	1790.30	1824.10	1720.94	1746.00
5	4751.3	1681.30	2884.77	2760.37	2826.67	2782.00
6	3280.5	1609.58	3077.14	2719.62	2862.61	2843.00
MAE[a]	1153.66	951.75	317.24	45.75	18.80	
RMSE[a]	1272.66	968.88	403.04	61.78	23.28	

[a]Negative frequencies were removed from the ANI-1ccx (CELU) when evaluating errors.

Table S4. VPT2 anharmonic frequencies (in cm^{-1}) of H_2CO calculated using TL ANI-1ccx-gelu trained on MP2, CCSD(T), and CCSD(T)-F12 data (ANI-1ccx-gelu/TL@MP2, ANI-1ccx-gelu/TL@CCSD(T), and ANI-1ccx-gelu/TL@CCSD(T)-F12). They are compared to the reference ab initio MP2 values[59], experimental values[87], and two PhysNet models[59].

mode	ANI-1ccx-gelu/ TL@MP2	ANI-1ccx-gelu/ TL@CCSD(T)	ANI-1ccx-gelu/ TL@CCSD(T)-F12	PhysNet MP2	PhysNet CCSD(T)	PhysNet CCSD(T)-F12	MP2	Exp
1	1188.27	1146.39	1223.89	1180.04, 1179.94	1163.73, 1163.52	1166.36, 1166.73	1180.23	1167.00
2	1219.92	1265.18	1443.51	1246.72, 1246.67	1240.62, 1241.36	1246.01, 1245.61	1246.71	1249.00
3	1567.01	1496.96	1526.44	1507.55, 1507.47	1493.89, 1494.91	1498.43, 1498.36	1507.95	1500.00
4	1705.88	1727.49	1833.76	1719.02, 1716.10	1729.94, 1731.88	1745.32, 1744.66	1720.94	1746.00
5	2916.85	2778.09	2731.19	2823.82, 2823.11	2775.38, 2779.58	2783.01, 2778.06	2826.67	2782.00
6	2809.76	2822.06	2884.15	2862.29, 2859.67	2819.51, 2821.49	2825.99, 2821.91	2862.61	2843.00
MAE_{Exp}	54.26	13.87	76.26	18.49, 18.40	10.66, 9.04	3.98, 5.28	18.80	
MAE_{MP}	42.00			0.95, 2.03				

HCOOH

Table S5. VPT2 anharmonic frequencies (in cm^{-1}) of HCOOH calculated using TL ANI-1ccx-gelu trained on MP2, CCSD(T), and CCSD(T)-F12 data (ANI-1ccx-gelu/TL@MP2, ANI-1ccx-gelu/TL@CCSD(T), and ANI-1ccx-gelu/TL@CCSD(T)-F12). They are compared to the reference ab initio MP2 values[59], experimental values[74], and two PhysNet models[59].

mode	ANI-1ccx-gelu/ TL@MP2	ANI-1ccx-gelu/ TL@CCSD(T)	ANI-1ccx-gelu/ TL@CCSD(T)-F12	PhysNet MP2	PhysNet CCSD(T)	PhysNet CCSD(T)-F12	MP2	Exp
1	619.78	619.11	620.18	619.28, 618.68	620.37, 619.30	626.05, 625.78	619.54	626.16
2	642.25	632.32	637.63	642.69, 641.62	630.86, 630.52	637.65, 637.58	641.78	640.72
3	1039.08	1029.55	1033.24	1037.06, 1036.44	1028.34, 1028.38	1033.08, 1033.04	1036.38	1033.47
4	1096.95	1099.21	1103.54	1097.01, 1096.67	1099.54, 1098.65	1104.11, 1104.7	1098.18	1104.85
5	1218.5	1293.7	1299.64	1220.49, 1219.25	1294.92, 1293.36	1301.99, 1301.89	1220.63	1306.20
6	1377.1	1375.13	1376.04	1380.81, 1380.24	1375.30, 1374.86	1377.35, 1378.23	1380.94	1380.00
7	1763.24	1767.72	1778.76	1758.17, 1757.89	1765.98, 1768.89	1774.44, 1774.60	1760.53	1776.83
8	2966.44	2930.51	2933.28	2959.50, 2960.15	2923.05, 2922.24	2924.61, 2927.30	2967.80	2942.00
9	3556.36	3562.76	3571.96	3554.88, 3555.83	3557.15, 3562.98	3565.70, 3565.34	3554.90	3570.50
MAE_{Exp}	18.24	7.86	3.69	17.62, 17.61	9.47, 9.06	3.97, 3.59		
MAE_{MP}	1.79			1.55, 1.77				

HNO_2

Table S6. VPT2 anharmonic frequencies (in cm^{-1}) of HNO_2 calculated using UIPs ANI-1ccx (with the default CELU activation function and after changing to the GELU activation function) and ANI-1ccx-gelu compared to the *ab initio* MP2 values[59] as well as with experiment and a representative DFT method (B3LYP/6-31G*).

mode	ANI-1ccx (CELU)	ANI-1ccx (GELU)	ANI-1ccx-gelu	DFT	MP2	Exp
1	1172.40	297.61	543.23	468.66	551.35	543.88
2	547.02	391.09	616.23	611.20	565.97	595.62
3	-6145.50	576.89	822.50	837.37	765.83	790.12
4	-26.90	652.00	1268.59	1273.05	1233.45	1263.21
5	-3229.15	1148.83	1734.42	1780.37	1637.31	1699.76
6	4269.76	2032.99	3608.21	3516.59	3577.17	3590.77
MAE[a]		563.99	18.52	50.45	27.87	
RMSE[a]		736.18	22.39	58.07	32.90	

Table S7. VPT2 anharmonic frequencies (in cm^{-1}) of HNO_2 calculated using TL ANI-1ccx-gelu trained on MP2, CCSD(T), and CCSD(T)-F12 data (ANI-1ccx-gelu/TL@MP2, ANI-1ccx-gelu/ TL@CCSD(T), and ANI-1ccx-gelu/TL@CCSD(T)-F12). They are compared to their reference ab initio MP2 values[59], experimental values[88], and two PhysNet models[59].

mode	ANI-1ccx-gelu/ TL@MP2	ANI-1ccx-gelu/ TL@CCSD(T)	ANI-1ccx-gelu/ TL@CCSD(T)-F12	PhysNet MP2	PhysNet CCSD(T)	PhysNet CCSD(T)-F12	MP2	Exp
1	553.46	533.95	544.76	551.59, 551.72	533.67, 533.93	541.10, 542.37	551.35	543.88
2	561.89	595.69	606.09	566.02, 566.12	595.98, 596.32	608.55, 608.92	565.97	595.62
3	763.09	789.92	797.48	766.25, 765.86	785.91, 786.93	799.70, 799.04	765.83	790.12
4	1237.23	1266.17	1262.30	1235.12, 1230.53	1260.08, 1260.91	1267.74, 1267.88	1233.45	1263.21
5	1638.30	1691.36	1712.11	1637.07, 1636.84	1691.60, 1691.48	1709.88, 1710.24	1637.31	1699.76
6	3576.41	3582.09	3585.02	3580.86, 3581.00	3580.01, 3581.84	3590.28, 3586.80	3577.17	3590.77
MAE_{Exp}	28.69	5.04	6.29	26.98, 27.83	6.14, 5.56	6.74, 7.14	27.87	
MAE_{MP2}	2.41			1.05, 1.30				

CH_3OH

Table S8. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3OH calculated using UIPs ANI-1ccx (with the default CELU activation function and after changing to the GELU activation function) and ANI-1ccx-gelu compared to the *ab initio* MP2 values[59] as well as with experiment and a representative DFT method (B3LYP/6-31G*).

mode	ANI-1ccx (CELU)	ANI-1ccx (GELU)	ANI-1ccx-gelu	DFT	MP2	Exp
1	7852.77	-186.92	230.48	1174.09	240.64	271.50
2	1257.71	336.74	1038.17	1036.97	1030.29	1033.50
3	1597.55	856.19	1070.95	1149.66	1069.25	1074.50
4	195.92	1149.17	1134.79	1293.10	1155.02	1145.00
5	5559.27	621.93	1185.31	1428.81	1321.98	1332.00
6	1186.87	732.20	1441.71	1474.75	1457.30	1454.50
7	730.12	790.91	1489.90	1564.51	1483.22	1465.00
8	4333.03	1024.04	1500.86	1518.61	1489.48	1479.50
9	-627.34	1761.84	2811.30	2842.28	2996.01	2844.20
10	-1805.08	1827.01	2721.53	2947.86	2996.40	2970.00
11	-2211.29	1823.85	2957.73	2939.25	3046.42	2999.00
12	5448.49	1908.93	3690.35	3668.19	3687.39	3681.50
MAE[a]	2125.32	786.75	49.72	123.51	26.82	
RMSE[a]	3134.87	915.61	86.21	268.91	48.01	

[a]Negative frequencies were removed from the ANI-1ccx (CELU) and (GELU) when evaluating errors.

Table S9. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3OH calculated using TL ANI-1ccx-gelu trained on MP2 and CCSD(T) data (ANI-1ccx-gelu/TL@MP2 and ANI-1ccx-gelu/TL@CCSD(T)). They are compared to their reference ab initio MP2 values[59], experimental values[89], and two PhysNet models[59].

Mode	ANI-1ccx-gelu/ TL@MP2	ANI-1ccx-gelu/ TL@CCSD(T)	PhysNet MP2	PhysNet CCSD(T)	MP2	Exp
1	234.12	226.99	240.84, 241.49	236.97, 238.08	240.64	271.50
2	1031.82	1029.47	1031.16, 1030.51	1026.41, 1027.04	1030.29	1033.50
3	1067.12	1065.81	1069.23, 1069.65	1063.32, 1063.78	1069.25	1074.50
4	1156.04	1143.68	1154.53, 1154.90	1145.68, 1145.95	1155.02	1145.00
5	1325.82	1318.68	1321.89, 1321.99	1321.31, 1322.05	1321.99	1332.00
6	1457.58	1450.93	1457.42, 1457.78	1447.89, 1448.52	1457.30	1454.50
7	1487.78	1480.39	1483.39, 1483.11	1469.21, 1468.55	1483.22	1465.00
8	1494.63	1492.60	1490.11, 1489.77	1475.75, 1475.62	1489.48	1479.50
9	2861.79	2833.13	2998.34, 2997.26	2827.05, 2829.22	2996.01	2844.20
10	2999.13	2916.02	3000.01, 3001.00	2911.42, 2912.12	2996.40	2970.00
11	3047.11	2982.86	3048.63, 3048.07	2987.57, 2988.73	3046.42	2999.00
12	3679.09	3682.74	3688.76, 3686.33	3666.53, 3667.99	3687.39	3681.50
MAE_{Exp}	16.83	15.53	27.57, 27.28	15.07, 14.30	26.82	
MAE_{MP2}	14.25		1.01, 0.92			

CH_3CHO

Table S10. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3CHO calculated using UIPs ANI-1ccx (with the default CELU activation function and after changing to the GELU activation function) and ANI-1ccx-gelu compared to the *ab initio* MP2 values[59] as well as with experiment and a representative DFT method (B3LYP/6-31G*).

mode	ANI-1ccx (CELU)	ANI-1ccx (GELU)	ANI-1ccx-gelu	DFT	MP2	Exp
1	664.36	-197.14	157.22	135.10	149.94	143.80
2	2057.13	332.66	510.05	507.56	505.96	508.80
3	7629.99	225.85	800.59	768.40	765.11	764.10
4	5724.89	499.69	881.33	858.16	867.05	865.90
5	7231.06	507.10	1050.39	1112.48	1109.98	1097.80
6	1679.99	1160.83	1109.99	1120.89	1116.41	1113.80
7	7363.04	559.22	1360.35	1368.29	1351.80	1352.60
8	-890.87	684.54	1400.79	1423.09	1403.23	1394.90
9	844.01	747.26	1467.55	1450.15	1437.76	1433.50
10	1227.45	784.49	1463.63	1457.71	1445.08	1436.30
11	2786.33	952.99	1777.98	1817.76	1726.61	1746.00
12	2201.96	1892.56	2727.45	2691.74	2735.51	2715.40
13	1292.05	1777.12	2902.95	2943.00	2964.93	2923.20
14	1452.91	1694.83	2986.43	2952.47	3012.12	2964.30
15	2584.41	1850.82	3012.64	3014.43	3059.86	3014.30
MAE[a]	2316.30	696.79	18.73	16.86	14.85	
RMSE[a]	3320.37	775.54	23.22	23.72	21.99	

[a]Negative frequencies were removed from the ANI-1ccx (CELU) and (GELU) when evaluating errors.

Table S11. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3CHO calculated using TL ANI-1ccx-gelu trained on MP2 data (ANI-1ccx-gelu/TL@MP2). They are compared to their reference ab initio MP2 values[59], experimental values[90], and two PhysNet models[59].

mode	ANI-1ccx-gelu/ TL@MP2	PhysNet MP2	MP2	Exp
1	197.63	150.37, 151.46	149.94	143.80
2	507.33	505.60, 505.45	505.96	508.80
3	742.84	766.03, 766.42	765.11	764.10
4	870.50	867.46, 868.57	867.05	865.90
5	1146.28	1110.82, 1110.70	1109.98	1097.80
6	1122.81	1115.91, 1116.88	1116.41	1113.80
7	1351.77	1352.20, 1351.93	1351.80	1352.60
8	1410.81	1403.31, 1403.48	1403.23	1394.90
9	1448.23	1437.73, 1438.49	1437.76	1433.50
10	1450.22	1445.30, 1444.24	1445.08	1436.30
11	1735.74	1728.74, 1726.97	1726.61	1746.00
12	2745.56	2735.11, 2735.32	2735.51	2715.40
13	2969.15	2966.08, 2965.81	2964.93	2923.20
14	2994.34	3012.46, 3011.77	3012.12	2964.30
15	3052.68	3064.91, 3060.87	3059.86	3014.30
MAE_{Exp}	22.59	15.27, 15.32	14.85	
MAE_{MP2}	12.60	0.88, 0.72		

CH_3COOH

Table S12. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3COOH calculated using UIPs ANI-1ccx (with the default CELU activation function and after changing to the GELU activation function) and ANI-1ccx-gelu compared to the *ab initio* MP2 values[59] as well as with experiment and a representative DFT method (B3LYP/6-31G*).

mode	ANI-1ccx (CELU)	ANI-1ccx (GELU)	ANI-1ccx-gelu	DFT	MP2	Exp
1	37635.00	-754.29	2510.34	39.50	77.93	-
2	13913.00	3798.67	1067.99	419.84	422.94	424.00
3	15296.10	-7649.07	390.23	537.85	536.39	534.50
4	2833.06	-5183.72	919.66	574.98	575.61	581.50
5	-1532.99	1516.88	1889.33	670.37	639.50	642.00
6	7906.26	-2657.60	1224.31	849.11	855.10	847.00
7	6152.98	-5942.66	947.59	990.86	987.46	991.00
8	8520.33	-1839.45	1217.86	1054.76	1047.31	1049.00
9	2889.26	-7954.63	1338.90	1186.61	1159.17	1184.00
10	1817.96	-3620.34	1217.84	1295.59	1320.83	1266.00
11	8330.82	-2502.88	1452.60	1400.46	1378.53	1384.50
12	5488.43	-4909.94	1309.11	1462.07	1439.69	1430.00
13	3438.94	-1328.35	1578.28	1457.54	1450.29	1430.00
14	5093.70	-2157.59	1969.75	1824.30	1780.90	1792.00
15	6560.85	20.05	2947.07	2969.35	2989.40	2944.00
16	4596.76	320.83	3112.48	2995.54	3040.41	2996.00
17	4107.41	1174.76	3064.82	3034.36	3080.19	3051.00
18	15155.30	251.70	3564.16	3501.89	3568.62	3585.50
MAE[a]	5413.20		225.65	18.62	16.90	
RMSE[a]	6940.46		375.10	27.43	23.74	

[a]Negative frequencies were removed from the ANI-1ccx (GELU) when evaluating errors.

Table S13. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3COOH calculated using TL ANI-1ccx-gelu trained on MP2 data (ANI-1ccx-gelu/TL@MP2). They are compared to their reference ab initio MP2 values[59], experimental values[91], and two PhysNet models[59].

mode	ANI-1ccx-gelu/ TL@MP2	PhysNet MP2	MP2	Exp
1	73.42	74.53, 73.11	77.93	-
2	422.99	422.52, 423.21	422.94	424.00
3	546.19	536.36, 538.21	536.39	534.50
4	583.82	574.91, 575.17	575.61	581.50
5	661.42	638.93, 641.25	639.50	642.00
6	873.80	855.20, 855.37	855.10	847.00
7	1008.91	987.08, 987.15	987.46	991.00
8	1076.54	1047.38, 1045.69	1047.31	1049.00
9	1205.64	1159.13, 1161.15	1159.17	1184.00
10	1343.16	1320.53, 1322.89	1320.83	1266.00
11	1422.48	1377.70, 1378.64	1378.53	1384.50
12	1493.19	1438.15, 1437.89	1439.69	1430.00
13	1498.45	1449.95, 1450.01	1450.29	1430.00
14	1810.13	1780.63, 1787.68	1780.90	1792.00
15	3098.08	2984.12, 2989.62	2989.40	2944.00
16	3182.44	3037.04, 3040.77	3040.41	2996.00
17	3224.29	3082.90, 3080.65	3080.19	3051.00
18	3751.78	3574.65, 3567.07	3568.62	3585.50
MAE_{Exp}	63.14	16.25, 16.67	16.90	
MAE_{MP2}	51.97	1.47, 1.49		

CH_3NO_2

Table S14. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3NO_2 calculated using UIPs ANI-1ccx (with the default CELU activation function and after changing to the GELU activation function) and ANI-1ccx-gelu compared to the *ab initio* MP2 values[59] as well as with experiment and a representative DFT method (B3LYP/6-31G*).

mode	ANI-1ccx (CELU)	ANI-1ccx (GELU)	ANI-1ccx-gelu	DFT	MP2	Exp
1	-5219.93	1007.07	128671.00	5.30	-47.12	-
2	-4719.78	369.56	2787.77	474.93	477.77	479.00
3	345.48	299.85	3071.22	605.43	593.54	599.00
4	-2140.72	365.95	2873.98	652.74	663.10	647.00
5	-7937.01	462.19	1436.55	907.65	923.87	921.00
6	-5536.19	462.89	2326.74	1097.49	1105.09	1097.00
7	-1164.60	684.14	2642.67	1120.45	1119.84	1153.00
8	-941.58	653.21	2745.92	1372.72	1385.96	1384.00
9	100.09	1229.59	3632.49	1423.36	1395.09	1413.00
10	2584.80	611.66	3469.29	1452.77	1448.14	1449.00
11	-1.25	1058.44	2401.08	1458.07	1449.79	1488.00
12	176.87	1018.99	2165.75	1644.35	1727.56	1582.00
13	-5650.30	1800.83	3340.32	2992.39	3006.22	2965.00
14	-502.78	1834.95	3769.87	3056.19	3086.01	3048.00
15	-1109.20	1847.94	3843.19	3078.68	3106.89	3048.00
MAE[a]		627.76	1373.85	17.61	29.25	
RMSE[a]		733.09	1552.19	24.08	47.09	

[a]Negative frequencies were removed from MP2 when evaluating errors.

Table S15. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3NO_2 calculated using TL ANI-1ccx-gelu trained on MP2 data (ANI-1ccx-gelu/TL@MP2). They are compared to their reference ab initio MP2 values[59], experimental value[92], and two PhysNet models[59].

mode	ANI-1ccx-gelu/ TL@MP2	PhysNet MP2	MP2	Exp
1	47.28	-84.94, -55.88	-47.12	-
2	478.25	479.22, 478.01	477.77	479.00
3	611.13	593.12, 594.60	593.54	599.00
4	669.90	661.89, 662.60	663.10	647.00
5	939.61	922.30, 925.33	923.87	921.00
6	1128.32	1104.10, 1106.03	1105.09	1097.00
7	1149.50	1120.33, 1119.07	1119.84	1153.00
8	1411.33	1385.86, 1385.15	1385.96	1384.00
9	1430.27	1394.18, 1394.87	1395.09	1413.00
10	1491.13	1447.35, 1448.15	1448.14	1449.00
11	1501.61	1448.21, 1450.30	1449.79	1488.00
12	1745.44	1713.45, 1726.72	1727.56	1582.00
13	3115.51	3011.56, 3008.23	3006.22	2965.00
14	3219.69	3086.54, 3087.68	3086.01	3048.00
15	3245.90	3107.14, 3101.98	3106.89	3048.00
MAE_{Exp}[a]	62.36	28.56, 29.12	29.25	
MAE_{MP2}[a]	46.34	2.12, 1.14		

[a]Negative frequencies were removed from PhysNet and MP2 when evaluating errors as is done in literature[59].

CH_3CONH_2

Table S16. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3CONH_2 calculated using UIPs ANI-1ccx (with the default CELU activation function and after changing to the GELU activation function) and ANI-1ccx-gelu compared to the *ab initio* MP2 values[59] as well as with experiment and a representative DFT method (B3LYP/6-31G*).

mode	ANI-1ccx (CELU)	ANI-1ccx (GELU)	ANI-1ccx-gelu	DFT	MP2	Exp
1	3718.78	14.49	171.81	-79.40	-2425.72	-
2	362.96	192.23	1276.11	-2106.28	-213.48	269.00
3	7292.51	42.76	365.80	445.79	451.29	427.00
4	617.61	303.37	516.67	581.05	477.92	507.00
5	3065.86	409.73	557.87	481.69	574.20	548.00
6	949.13	432.37	632.99	709.77	700.92	625.00
7	3903.98	349.42	873.02	814.81	840.39	858.00
8	866.73	517.56	989.53	976.95	979.28	965.00
9	-3855.97	556.04	1018.14	1041.90	1033.17	1040.00
10	1524.49	720.39	1108.52	1093.84	1054.59	1134.00
11	7935.90	551.06	1328.52	1325.81	1340.62	1319.00
12	7984.97	622.18	1386.10	1414.77	1371.24	1385.00
13	1853.51	767.87	1448.44	1456.84	1443.50	1432.00
14	4107.44	784.72	1488.94	1489.66	1463.58	1433.00
15	2460.15	899.70	1610.71	1589.37	1552.95	1600.00
16	2339.88	1225.83	1763.59	1766.00	1725.92	1733.00
17	734.90	1795.35	2933.32	2865.75	2994.98	2860.00
18	5010.43	1744.47	3004.59	3080.65	3008.41	2900.00
19	1190.29	1766.06	3050.05	2914.63	3070.81	2967.00
20	10793.20	1921.66	3444.96	3482.86	3501.75	3450.00
21	5833.81	2054.44	3544.31	3607.80	3657.46	3550.00
MAE[a]	2466.63	667.24	78.94	43.80	47.98	
RMSE[a]	3479.92	786.75	228.91	59.31	62.29	

[a]Negative frequencies were removed from the ANI-1ccx (CELU), DFT, and MP2 when evaluating errors.

Table S17. VPT2 anharmonic frequencies (in cm^{-1}) of CH_3CONH_2 calculated using TL ANI-1ccx-gelu trained on MP2 data (ANI-1ccx-gelu/TL@MP2). They are compared to their reference ab initio MP2 values[59], experimental value[93], and two PhysNet models[59].

mode	ANI-1ccx-gelu/ TL@MP2	PhysNet MP2	DFT	MP2	Exp
1	86.76	-2337.85, -2548.69	-79.40	-2425.72	-
2	-1961.15	-206.36, -204.12	-2106.28	-213.48	269.00
3	439.27	453.25, 457.15	445.79	451.29	427.00
4	624.98	479.72, 473.40	581.05	477.92	507.00
5	475.70	574.42, 576.91	481.69	574.20	548.00
6	679.93	701.19, 706.62	709.77	700.92	625.00
7	847.06	839.33, 840.76	814.81	840.39	858.00
8	977.93	981.95, 984.00	976.95	979.28	965.00
9	1027.64	1037.24, 1038.03	1041.90	1033.17	1040.00
10	1092.39	1056.99, 1054.48	1093.84	1054.59	1134.00
11	1324.71	1342.26, 1344.85	1325.81	1340.62	1319.00
12	1372.05	1370.49, 1370.80	1414.77	1371.24	1385.00
13	1437.07	1444.48, 1443.41	1456.84	1443.50	1432.00
14	1472.07	1464.30, 1459.85	1489.66	1463.58	1433.00
15	1566.25	1556.28, 1556.57	1589.37	1552.95	1600.00
16	1726.40	1718.97, 1726.61	1766.00	1725.92	1733.00
17	2968.73	2996.35, 2990.34	2865.75	2994.98	2860.00
18	3108.67	3005.78, 2993.60	3080.65	3008.41	2900.00
19	3034.73	3068.72, 3072.59	2914.63	3070.81	2967.00
20	3503.60	3496.12, 3504.90	3482.86	3501.75	3450.00
21	3666.99	3646.50, 3665.08	3607.80	3657.46	3550.00
MAE_{Exp}[a]	52.33	47.23, 48.40	43.80	47.98	
MAE_{MP}[a]	28.91	2.71, 3.88	46.65		

[a]Negative frequencies were removed when evaluating errors as is done in literature[59].

Fixed weights in fine-tuning

Table S18. Weights fixed in transfer learning for different molecules and data based on the best performance of generated anharmonic frequencies. The weights are fixed between the layers in the feed-forward neural network used in the ANI-type potentials. The layers are numbered from 1 to 5, with layer 1 – input layer, layer 5 – output layer, and layers 2–4 – the hidden layers.

Molecules/data	MP2	CCSD(T)	CCSD(T)-F12
HCOOH	(1, 2) & (4, 5)	(3, 4) & (4, 5)	(3, 4) & (4, 5)
H_2CO	(2, 3) & (3, 4)	(1, 2) & (4, 5)	(1, 2) & (4, 5)
HNO_2	(1, 2) & (2, 3)	(1, 2) & (2, 3)	(1, 2) & (2, 3)
CH_3OH	(2, 3) & (4, 5)	(1, 2) & (2, 3)	-
CH_3CHO	(1, 2) & (4, 5)	-	-
CH_3COOH	(1, 2) & (4, 5)	-	-
CH_3NO_2	(1, 2) & (2, 3)	-	-
CH_3CONH_2	(1, 2) & (2, 3)	-	-

Table S19. VPT2 anharmonic frequencies (in cm^{-1}) of HCOOH calculated using ANI-1ccx-gelu/TL@CCSD(T)-F12 fine-tuned with different NN weights fixed. They are compared to the experimental values[74] and determine the effect of different layers fixed on the performance of generated anharmonic frequencies. The weights are fixed between the layers in the feed-forward neural network used in the ANI-type potentials. The layers are numbered from 1 to 5, with layer 1 – input layer, layer 5 – output layer, and layers 2–4 – the hidden layers. Fixing weights between layers 3, 4 and 4, 5 shows the best accuracy in anharmonic frequencies.

mode	(3, 4) & (4, 5)	(2, 3) & (4, 5)	(1, 2) & (3, 4)	(1, 2) & (4, 5)	(2, 3) & (3, 4)	(1, 2) & (2, 3)	(1, 2) & (2, 3) & (3, 4)	(1, 2) & (2, 3) & (4, 5)	(1, 2) & (3, 4) & (4, 5)	No layers
1	620.18	618.68	623.67	625.84	618.44	636.08	632.25	624.58	622.83	616.03
2	637.63	638.43	637.94	637.87	637.90	637.02	633.36	637.78	637.54	638.99
3	1033.24	1033.98	1035.23	1035.53	1033.74	1035.70	1032.81	1034.51	1034.81	1034.54
4	1103.54	1098.74	1101.56	1104.22	1104.42	1112.56	1097.96	1100.37	1101.89	1102.59
5	1299.64	1297.76	1297.20	1297.21	1300.82	1309.40	1293.22	1298.36	1295.34	1299.27
6	1376.04	1374.86	1375.72	1374.47	1377.71	1381.28	1367.00	1376.72	1376.46	1377.41
7	1778.76	1779.91	1775.09	1779.86	1781.35	1780.65	1773.06	1778.54	1776.83	1780.47
8	2933.28	2931.38	2928.04	2930.04	2932.17	2941.04	2923.60	2931.92	2930.23	2931.98
9	3571.96	3562.08	3565.18	3565.01	3568.17	3573.67	3564.17	3563.57	3565.70	3566.73
MAE	3.69	5.79	4.96	4.54	3.95	4.00	8.39	4.43	4.64	4.68
RMSE	4.57	6.58	6.26	5.84	4.99	4.88	9.83	5.34	6.00	5.72