

Artificial Intelligence-Enhanced Quantum Chemical Method with Broad Applicability

Peikun Zheng,^a Roman Zubatyuk,^b Wei Wu,^a Olexandr Isayev,^{b*} Pavlo O. Dral^{a*}

^aState 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

^bDepartment of Chemistry, Carnegie Mellon University, Pittsburgh PA, 15213, USA

E-Mail: olexandr@olexandrisayev.com
dral@xmu.edu.cn

Abstract

High-level quantum mechanical (QM) calculations are indispensable for accurate explanation of natural phenomena on the atomistic level. Their staggering computational cost, however, poses great limitations, which luckily can be lifted to a great extent by exploiting advances in artificial intelligence (AI). Here we introduce the general-purpose, highly transferable artificial intelligence–quantum mechanical method 1 (AIQM1). It approaches the accuracy of the ‘gold-standard’ coupled cluster QM method with low computational speed of the approximate low-level semiempirical QM methods. AIQM1 can provide accurate ground-state energies for diverse organic compounds as well as geometries for even challenging systems such as large conjugated compounds (fullerene C₆₀) close to experiment. Noteworthy, our method’s accuracy is also good for ions and excited-state properties, although the neural network part of AIQM1 was never fitted to these properties.

Introduction

Quantum mechanical (QM) methods used in chemistry are invaluable for today’s modern science as they allow insight into electronic structure on an atomistic level that is otherwise unattainable experimentally. This in turn helps to find answers to fundamental scientific questions in chemistry and related fields, such as chemical physics and biology, and assists applied science in designing better materials and discover new medicines.

The usefulness of QM methods in practical applications is determined by their accuracy and computational cost. The trade-off between these two factors guides the choice of the QM

method. On the one side, we have very accurate, but slow high-level *ab initio* QM methods such as coupled cluster with single, double, and perturbative triple excitations, CCSD(T),¹ which has established itself as the “gold standard” in most applications, particularly, for closed-shell molecules.²⁻⁴ On the other side, we have very fast semiempirical QM (SQM) methods that have rather limited accuracy.⁵ The sweet spot of moderate computational cost and often sufficient accuracy is occupied by density functional theory (DFT) that has become a workhorse in the investigation of medium-sized systems (Figure 1a).⁶ The efforts for developing faster and more accurate QM methods is an active research field, but it is clear that traditional approaches to QM method development require years of hard human work and typically yield only relatively modest improvements.

Advances in artificial intelligence (AI) bring chemistry research to a radically new level and provide a much-needed alternative to the traditional QM method development.^{7,8} AI allows to perform calculations with both high accuracy and very low computational cost that was previously unattainable with the traditional QM methods. Nevertheless, most of the applications of AI to quantum chemistry are either proof-of-principle or limited to specific applications. Developing general-purpose AI approaches with transferability of QM methods remains a big challenge. A significant step towards transferable accurate AI approaches is the family of ANI potentials⁹⁻¹³ that can describe energies and forces of compounds of different size and composition in equilibrium and non-equilibrium configurations with accuracy approaching DFT⁹⁻¹² or even coupled cluster QM level¹³ (Figure 1a). They can be also applied to much larger systems than those included in the training dataset, because the total energy is calculated within the local approximation by the sum of the atomic contributions with each atom feeling the environment only within some cutoff.

While impressive, ANI potentials are however less transferable than general-purpose QM methods, because they are limited to closed-shell, neutral organic compounds and the use of the local approximation imposes farther limitations on their transferability, e.g., to large, highly conjugated systems (Figure 1b,c). A rational approach is to exploit synergies of AI and QM methods by merging them.⁷ This approach has already given rise to an increasing number of hybrid AI/QM methods,^{7,8,14-16} although most of them are either proof-of-principle or based on relatively slow DFT or trained on data of limited quantity and quality potentially restricting their transferability and accuracy.

Here we describe the general-purpose artificial intelligence–quantum mechanical method 1 (AIQM1) that approaches the coupled cluster accuracy with transferability of the QM methods and computational speed of the SQM methods (Figure 1d).

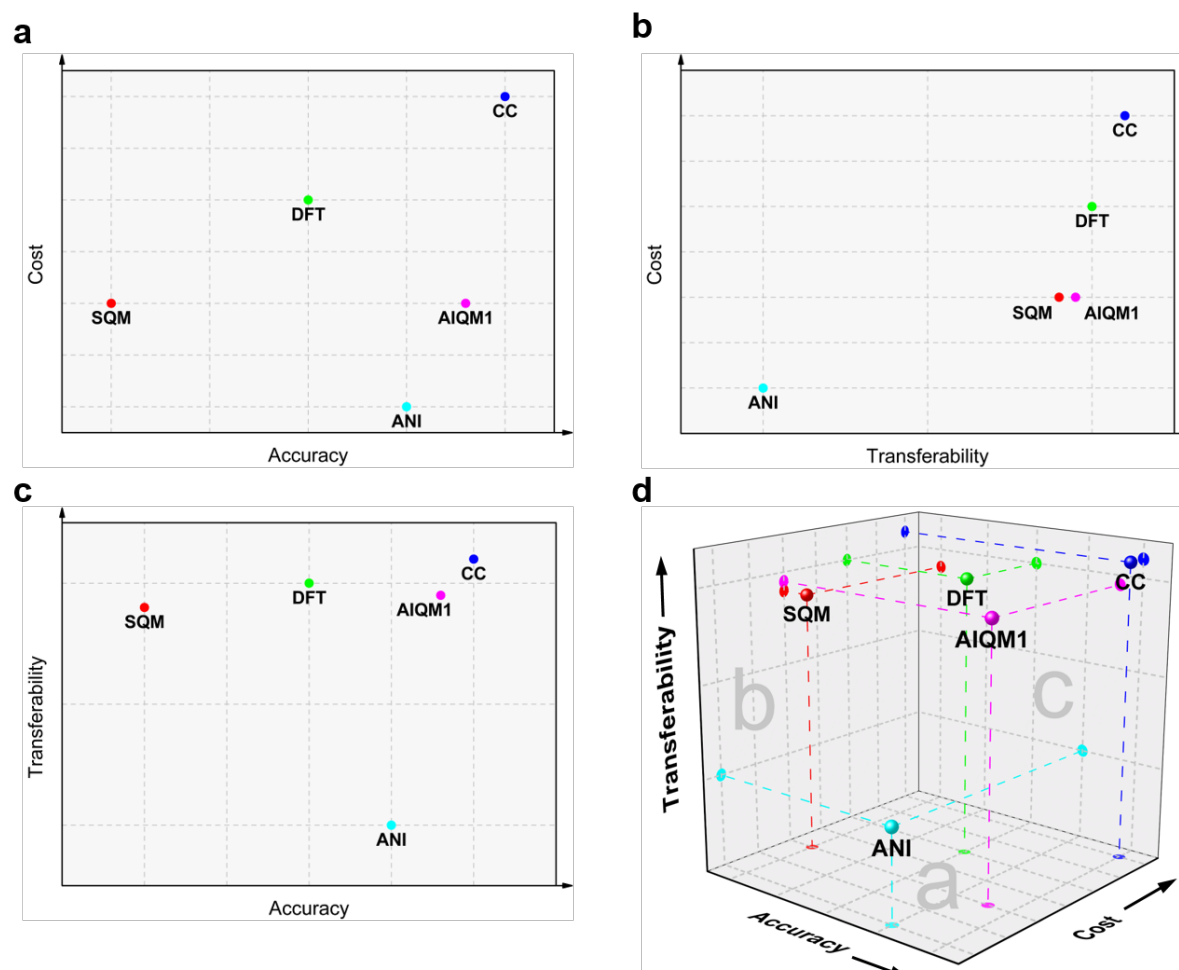


Figure 1. **Simplified scheme of quantum chemistry approximations.** Traditional quantum mechanical approaches such as the “gold-standard” coupled cluster (CC), “work-horse” density functional theory (DFT), “fast and approximate” semiempirical quantum mechanical (SQM) methods, artificial intelligence-based ANI, and the new artificial intelligence–quantum mechanical method 1 (AIQM1). They are compared with respect to a) cost and accuracy, b) cost and transferability, c) accuracy and transferability, d) cost, accuracy, and transferability. The accuracy of ANI is with respect to what it is applicable to.

Results

Method structure

The AIQM1 method consists of three main parts (Figure 2): 1) SQM Hamiltonian, 2) neural network (NN) correction to the potential, 3) dispersion corrections. The AIQM1 total energy E_{AIQM1} is the sum of the contributions from these three parts, E_{SQM} , E_{NN} , E_{disp} , respectively:

$$E_{\text{AIQM1}} = E_{\text{SQM}} + E_{\text{NN}} + E_{\text{disp}}. \quad (1)$$

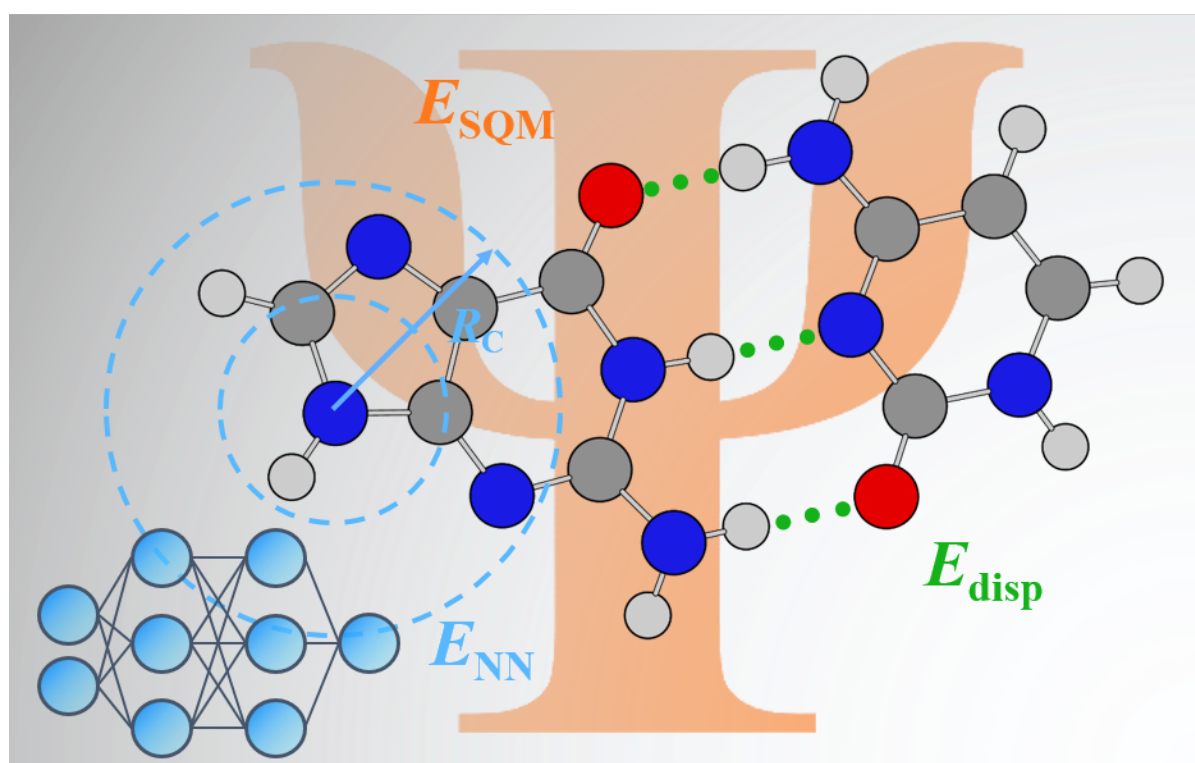


Figure 2. The structure of the AIQM1 method.

For the first part, we have chosen the orthogonalization- and dispersion-corrected method 2 (ODM2) Hamiltonian,¹⁷ which provides the most consistent and accurate predictions across different properties (from ground-state to excited-state and noncovalent interactions) among other SQM methods, particularly those based on neglect of diatomic differential overlap (NDDO) approximation. We remove the original D3-based dispersion corrections from the ODM2 approach and denote the modified approach as ODM2*. Instead, we add the state-of-the-art D4 dispersion corrections^{18,19} including with Axilrod–Teller–Muto three-body

contributions^{20,21} — the third part of AIQM1 method. These corrections are essential to describe properly dispersion terms in noncovalent interactions as they are described poorly by both SQM⁵ and local NN approaches such as ANI-1ccx²². For the second part, we took the ANI-type of NN potentials. We preserved the NN-architecture of ANI-1x that predicts E_{NN} by summing over N_{atoms} atomic contributions E_A :¹²

$$E_{\text{NN}} = \sum_A^{N_{\text{atoms}}} E_A. \quad (2)$$

We made only two minor modifications to NN model based on ANI-1x. First, we changed the activation function to GELU instead of CELU, because GELU is infinitely differentiable. This is important for applications where higher derivatives are required, e.g., geometry optimization and frequency calculations. Second, we increased the angular cutoff to 4 Å to assist with a better description of long-range interactions. Note that within ANI framework, atomic contributions are centered before fitting NN, i.e., the atomic contributions also include element-dependent terms obtained by linear fitting to the reference scalar values.

Method training and validation

The NN weights were obtained in two steps. In the first step, we fitted NN weights on the differences between the ground-state potentials calculated at DFT ω B97X/def2-TZVPP and ODM2* for the 4.6M geometries of the ANI-1x data set¹⁰. These DFT energies were used to train another successful general-purpose NN potential AIMNet.²³ This step is based on the Δ -learning²⁴ approach introduced by one of us and used here to correct the low-level SQM method to the target accuracy of the higher-level DFT method with comparatively small additional computational cost. (Calculations for the entire ANI-1x data set on a single CPU are ca. 10 times faster with a single ANI-type network NN compared to SQM calculations, but the difference should become larger for bigger systems and parallel computing.) The loss function L in this step is the geometric mean of the loss functions for energy differences between DFT and ODM2* (L_E , scalar values) and differences in forces (L_F , energy gradients $\frac{\partial E_{\text{NN}}}{\partial R}$ taken with opposite sign, vector values) contributions:

$$L = \sqrt{L_E L_F}, \quad (3)$$

with L_E and L_F defined analogously to the loss functions for energies and forces used in ANI-2x⁹.

In this way we trained an ensemble of eight NN models, which provides better accuracy than a single NN¹³ (see Methods). The method obtained in this first step is denoted by AIQM1@DFT* and it approaches DFT accuracy at the SQM cost for the hold-out test set as its mean absolute deviation (MAD) is only 0.7 kcal/mol for energies and 1.6 kcal/mol/Å for forces (Figure 3a).

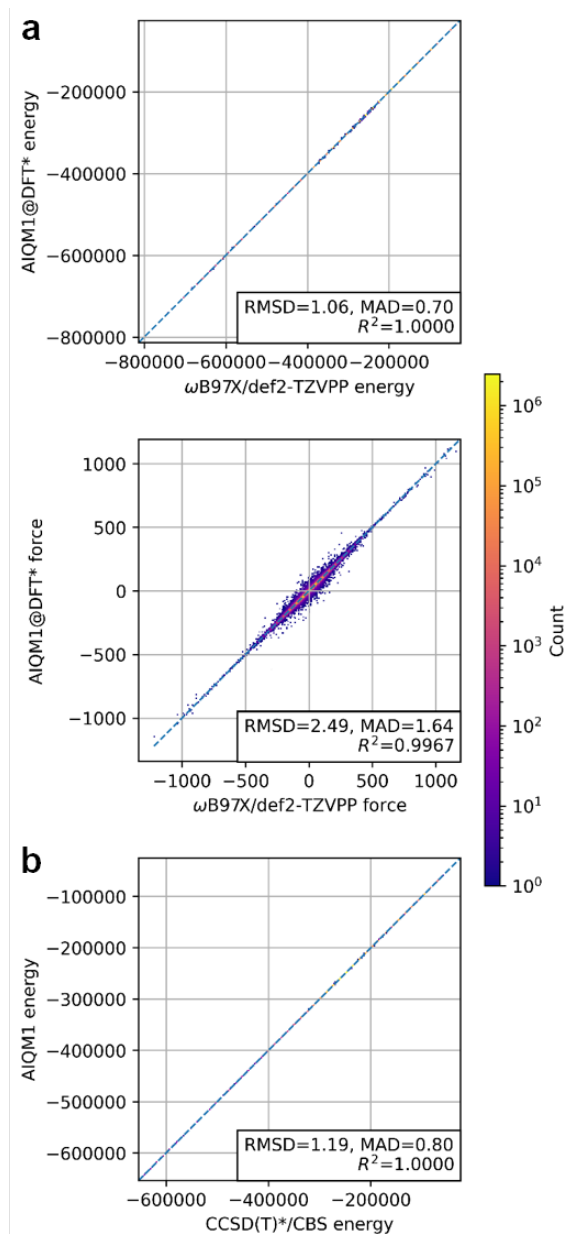


Figure 3. **Correlation between AIQM1 variants and reference methods for the hold-out test set.** a) Correlation between AIQM1@DFT* and ωB97X/def2-TZVPP energies and forces. b) Correlation between AIQM1 and CCSD(T)* / CBS. Root-mean-squared errors (RMSEs), mean absolute deviations (MADs), and squared correlation coefficients R² are also shown on the plots.

Since AIQM1@DFT* has no explicit dispersion corrections, we add the D4 dispersion corrections fitted¹⁹ for the DFT functional ω B97X and denote the resulting method as AIQM1@DFT.

In the second step of NN fitting, we used transfer learning²⁵ to reach coupled cluster accuracy using the 0.5M data points as was done for creating ANI-1ccx method¹³. Transfer learning is a powerful technique allowing to leverage more abundant training data for a related task to obtain the model for the target task using much fewer training points. For developing the AIQM1 method, we adjusted the weights of the first and third hidden layers of NN from the first step to minimize the loss function L_E for differences between the ground-state energies at CCSD(T)*/CBS and ODM2* with D4 corrections. The resulting approach is our final AIQM1 method and it approaches closely coupled cluster level for the hold-out test set as its MAD for energies is 0.8 kcal/mol (Figure 3b).

Performance for energies

AIQM1 has an excellent accuracy in energies for a broad range of data sets not used for fitting its NN part. A very important energy-based property is heat (enthalpy) of formation – a fundamental thermochemical quantity, which is notoriously difficult to accurately predict with quantum chemistry. Typically, only very computationally expensive QM methods are able to achieve the desired “chemical accuracy” for heats of formation (errors below 1 kcal/mol). Thus, AI was suggested as a potent approach to specifically target accurate and cost-efficient predictions of heats of formation by improving upon predictions made by the low-cost QM methods (DFT²⁶⁻²⁸ and SQM²⁹ methods). In contrast, in our approach we did not fit NN-part to better reproduce the heats of formation, we merely had to offset the bias in AIQM1 heats of formations at 298 K with respect to the experimental reference data in the CHNO data set³⁰ (popular data set to develop SQM methods) by just fitting four parameters — atomic energies of H, C, N, and O elements (see Methods).

AIQM1 performance is remarkable for heats of formation as it easily reaches chemical accuracy for the CHNO data set (MAD of 0.9 kcal/mol), even though this property was not included in the training set of its NN part. It is the first time that a QM method with semiempirical speed has broken this threshold as, e.g., ODM2 method with the best reported accuracy among semiempirical methods to date has three times higher MAD of 2.6 kcal/mol. Similarly, AIQM1 has MAD of 0.9 kcal/mol in heats of formation for the CHNO subset³¹ of

the independent G3/99 test set³², a set that formed a backbone for developing and testing many QM methods such as popular, but very expensive composite approaches G4³³ and G4MP2³⁴ targeting the coveted chemical accuracy. Thus, AIQM1 can be used as a computationally-efficient alternative to such composite methods.

Heats of formation can be considered as “absolute energies”. In chemistry, we often have to deal with relative energies such as isomerization energies, reaction energies and enthalpies as well as relative energies between conformers, because relative energies determine the outcome of reactions and 3D structures of molecules in thermal equilibrium. AIQM1 not only has good accuracy for absolute energies, but also faithfully reproduces relative energies. One example is the heats of formation and isomerization enthalpies at 298 K of organic compounds in the ISOMERS44 data set,^{31,35} for which AIQM1 has MAD of 0.4 and 0.5 kcal/mol, respectively.

Other types of relative energies, such as zero-point energy-excluded reaction energies at 0 K are also reproduced with AIQM1 very well. For example, isomerization energies in the IsoL6/11 data set³⁶ are reproduced by AIQM1 with chemical accuracy (MAD 0.6 kcal/mol, Figure 4a), while errors of ODM2 and ANI-1ccx are much larger (both with MAD of 1.5 kcal/mol)^{13,17}. AIQM1 accuracy is very close to CCSD(T)*/CBS (MAD 0.5 kcal/mol)¹³. Similarly, for another data set, reaction energies in the HC7/11 set,³⁷ AIQM1 accuracy is also very close to that¹³ of CCSD(T)*/CBS (MADs of 1.4 and 1.6 kcal/mol, respectively) and clearly outperforms both ODM2 and ANI-1ccx with MADs of 5.37 and 2.53 kcal/mol,^{13,17} respectively (Figure 4b). Curiously, for both IsoL6/11 and HC7/11 data sets, even AIQM1@DFT (MADs 1.5 and 9.2 kcal/mol, respectively) is much better than more expensive ω B97X/6-31G* tested previously¹³ (MADs 3.8 and 16.4 kcal/mol, respectively).

Relative energies of the configurations of the same molecule are also important as they determine, e.g., what rotational conformers are more stable, which is crucial for determining 3D structures of flexible molecules. AIQM1 confidently handles this task as its median MAD for the popular torsion benchmark set³⁸ is just 0.19 kcal/mol, which is much lower than that of ODM2 and ANI-1ccx¹³ (0.74 and 0.23 kcal/mol, respectively) and much more expensive ω B97X/6-31G* (0.33 kcal/mol)¹³ and B3LYP-D3/6-311+G** (0.24 kcal/mol)³⁸. AIQM1 is only slightly inferior to MP2/CBS (median MAD 0.11 kcal/mol),³⁸ which is in turn much slower than DFT. Now we can turn into investigating the performance of AIQM1 for predicting geometries themselves.

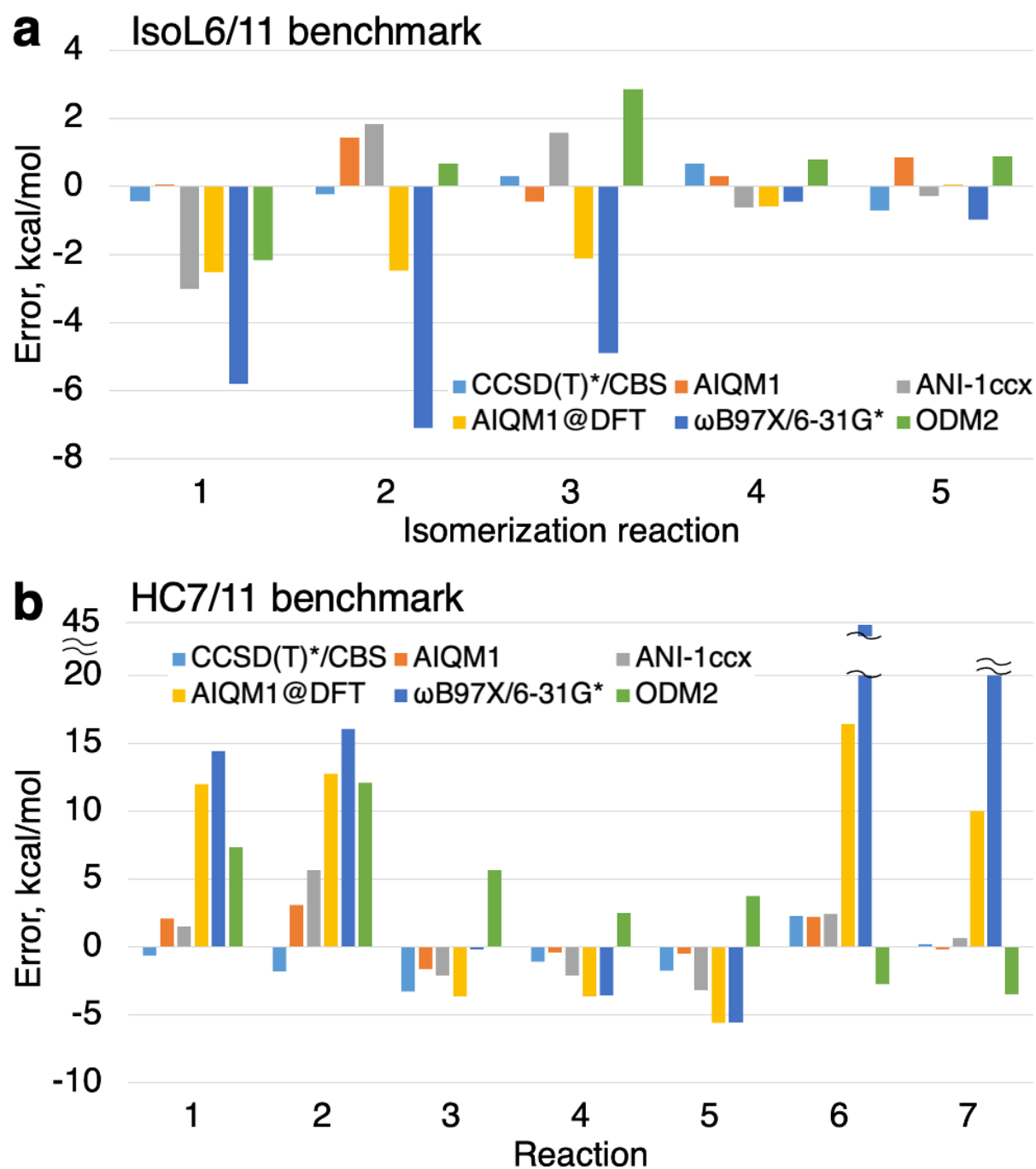


Figure 4. **Performance of AIQM1 for ground-state energies.** Comparison between errors of CCSD(T)*/CBS, AIQM1, ANI-1ccx, AIQM1@DFT, ω B97X/6-31G*, and ODM2 for the reaction energies in the a) IsoL6/11 and b) HC7/11 benchmark sets. Values for CCSD(T)*/CBS, ANI-1ccx, and ω B97X/6-31G* are taken from Ref. 13, values for ODM2 – from Ref. 17.

Performance for geometries

Theoretical prediction of molecular geometries is one of the most common applications of quantum chemistry, which is essential for chemical research as conclusive geometries are not always available from experiment. Geometry optimization is an iterative procedure requiring forces (and often Hessians), which makes it much more computationally expensive than energy calculations for a single geometry. SQM methods are much less accurate for geometries than common DFT methods and general-purpose NN potentials fail to deal with subtle conjugation effects, e.g., ANI-1ccx predicts that all bond lengths in C_{60} are equal to 1.451 Å, while it is known from experiment³⁹⁻⁴² that bond length between two adjacent hexagon rings is shorter than bond length between pentagon and hexagon rings (Figure 5a).

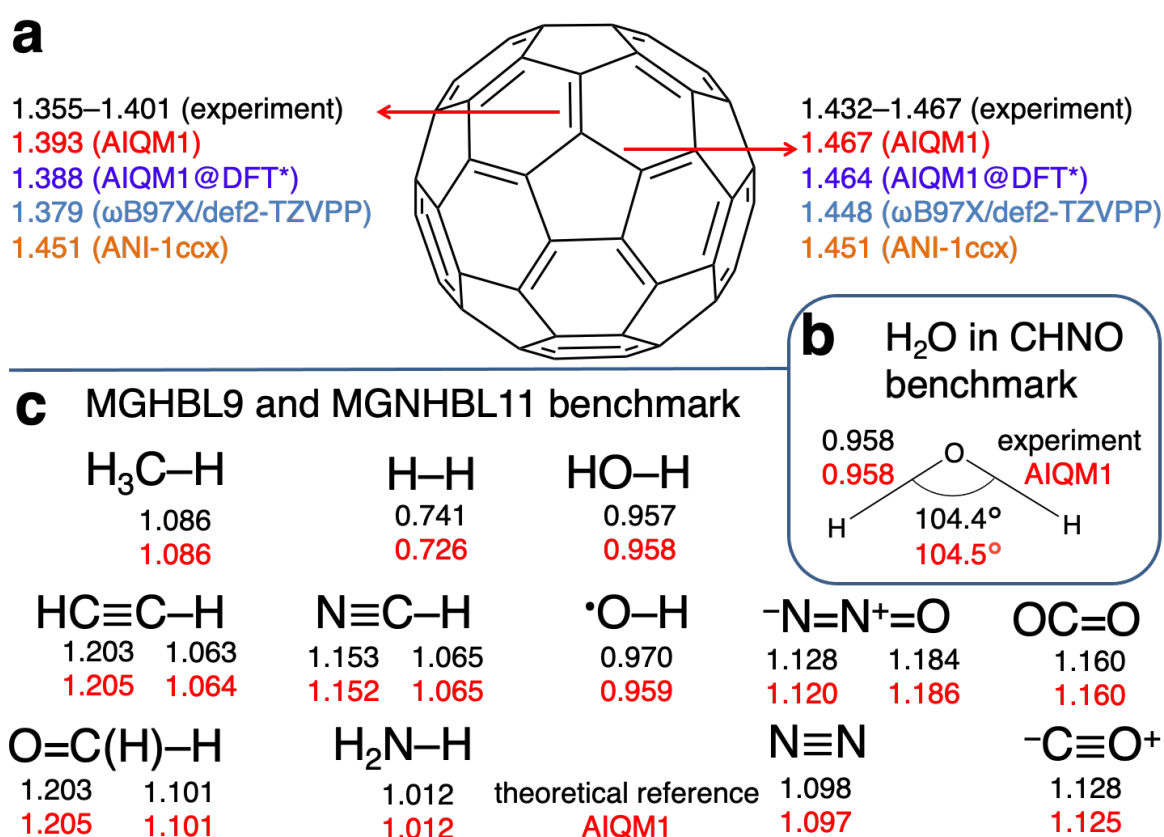


Figure 5. **Performance of AIQM1 for finding ground-state minimum geometries.** a) Short and long bond lengths in C_{60} as calculated at different levels of theory and compared to experimental values³⁹⁻⁴². b) Hydrogenic bond lengths (MGHBL9 benchmark)⁴³ and nonhydrogenic bond lengths (MGNHBL11 benchmark)^{43,44}. c) Geometry of a water molecule (CHNO benchmark)³⁰. Bond lengths are in Å.

AIQM1 successfully distinguishes these two bond types in C_{60} and predicts short and long bond lengths to be 1.393 and 1.467 Å, respectively (Figure 5a). For this molecule, we cannot compare AIQM1 predictions with CCSD(T)*/CBS due to the staggering cost of this coupled cluster approach, while experimental data are not conclusive as they range from 1.355 to 1.401 Å for short bond length and from 1.432 to 1.467 Å for long bond length depending on measurement conditions³⁹⁻⁴². Instead, we compare AIQM1@DFT* predicting 1.388 and 1.464 Å to ω B97X/def2-TZVPP predictions of 1.379 and 1.449 Å, which are in acceptable agreement (Figure 5a), while the cost of geometry optimization with AIQM1@DFT* is 14 s on a single CPU core vs 31 min on 32 CPU cores at DFT.

For smaller molecules, where reliable experimental and theoretical data is available, AIQM1 has very good accuracy, much better than, e.g., the accuracy of ODM2 or ANI-1ccx. For the CHNO data set³⁰ with experimental reference data, the MADs of AIQM1, ODM2, and ANI-1ccx are 0.007, 0.015, and 0.011 Å in bond lengths, 0.72°, 2.04°, and 1.00° in bond angles, and 2.38°, 4.07°, and 5.86° in dihedral angles, respectively (see, e.g., excellent prediction of water geometry, Figure 5b). Similarly, for hydrogenic (MGHBL9)⁴³ and nonhydrogenic bond lengths (MGNHBL11)^{43,44} data sets with accurate theoretical data used to test DFT methods, MAD of AIQM1 in bond lengths is 0.004 and 0.002 Å (Figure 5c), respectively, which is again much better than ODM2 (0.023 and 0.026 Å) or ANI-1ccx (0.093 and 0.004 Å).

Performance for noncovalent interactions

AIQM1 is transferable to noncovalent interactions too, which are very challenging even for the state-of-the-art QM methods and NN potentials. For the standard benchmark set S66x8⁴⁵ with CCSD(T)/CBS reference noncovalent interaction energies, AIQM1 has rather good accuracy as its MAD is 0.6 kcal/mol (Figure 6), which is comparable to ODM2 (0.8 kcal/mol) and DFT, e.g., ω B97X-D/6-31G* (1.2 kcal/mol) and ω B97X-D4/def2-TZVPP (0.5 kcal/mol). Hence, AIQM1 is a good cost-efficient alternative to such DFT methods.

The method performance is particularly good for hydrogen-bonded complexes. For clusters of neutral water molecules $(H_2O)_n$, and charged clusters $H^+(H_2O)_n$ and $OH^-(H_2O)_n$ (WATER27 data set⁴⁶ with revised values⁴⁷ for $(H_2O)_{20}$ clusters), AIQM1 has MAD of only 2.1 kcal/mol (Figure 6) compared to 4.5 of ODM2. This makes the method competitive in terms of accuracy with popular dispersion corrected DFT approaches, which have similar errors,⁴⁷ but are much slower. AIQM1 is therefore promising method for simulating chemical processes in water solutions, essential for biological processes. It is noteworthy that this data

set contains charged species, which cannot be adequately described by ANI-1ccx, which brings us to the next topic.

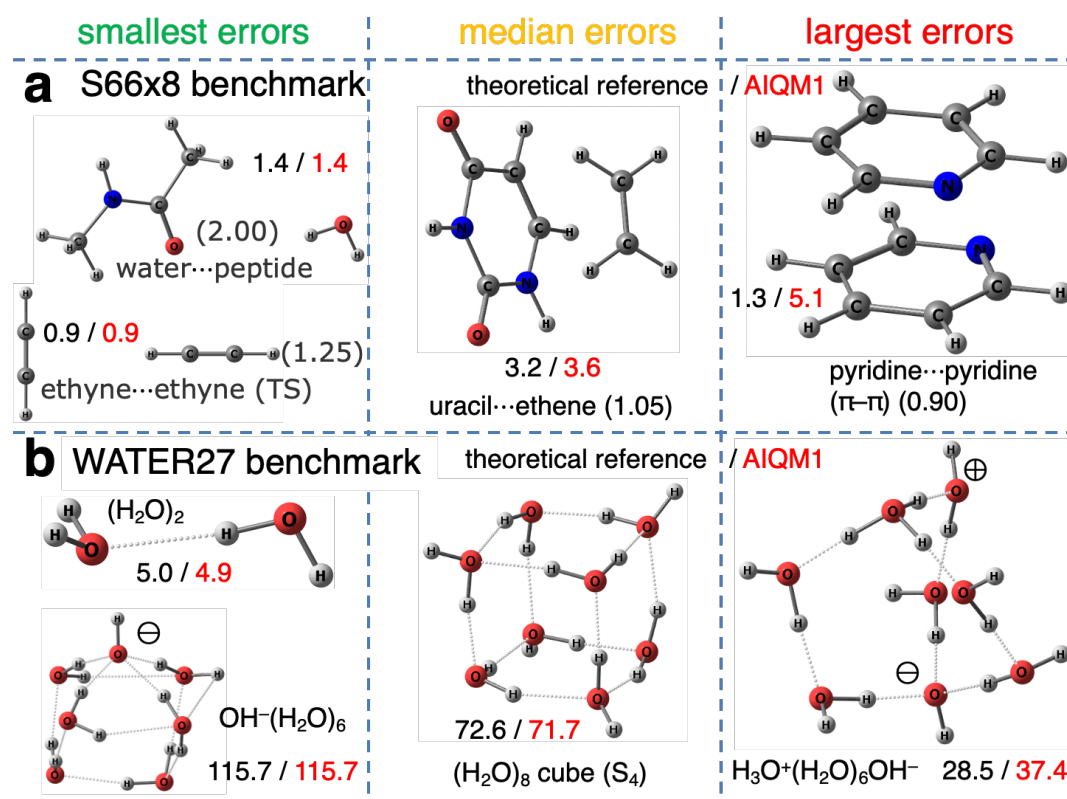


Figure 6. **Performance of AIQM1 for noncovalent interactions.** Selection of complexes with errors (in kcal/mol) ranging from smallest to median to largest values for a) the S66x8 benchmark and b) the WATER27 benchmark.

Beyond closed-shell, neutral molecules

AIQM1 is transferable beyond closed-shell, neutral species used for fitting its NN part and even improves upon the ODM2 method (ANI potentials cannot be used at all for such simulations). We saw before that AIQM1 performs well for charged protonated and deprotonated water clusters. Other examples are proton affinities, where MAD is improved from 16.6 (ODM2) to 10.5 (AIQM1) kcal/mol for the PA data set,⁴⁶ MAD in adiabatic ionization potentials (IP21 set)⁴⁶ from 15.6 to 2.2 kcal/mol, and MAD in adiabatic electron affinities (EA13 set)⁴⁶ from 12.7 to 7.7 kcal/mol.

Interestingly, geometries are also improved for charged species as for the CATIONS41 data set,^{31,48} the MADs of AIQM1 and ODM2 are 0.018 and 0.030 Å in bond lengths, 1.36° and

2.41° in bond angles, and 1.27° and 2.96° in dihedral angles, respectively.

Beyond ground-state properties

Finally, AIQM1 method is also transferable to electronically excited states and, e.g., it can be used for multi-reference configuration interaction (MRCI) calculations to predict excitation energies, oscillator strengths and nonadiabatic couplings for simulating spectra and performing nonadiabatic excited-state dynamics. AIQM1/MRCI is three orders of magnitude faster than popular linear-response time-dependent (TD) DFT approaches such as TD-B3LYP, while its accuracy is similar for vertical excitation energies (MAD of AIQM1/MRCI is 0.35 eV, which is close to TD-B3LYP/TZVP with MAD of 0.33 eV⁴⁹ for the Thiel's data set⁴⁹, Figure 7a).

It makes AIQM1 attractive for performing more computationally expensive tasks such as optimization of excited-state geometries (required, e.g., for simulating fluorescence spectra). We tested its performance on the ExGeom set^{17,49} with excited-state geometries and AIQM1 MAD for bond lengths is 0.018 Å vs CC2 reference (with TZVP basis set) and 0.019 Å vs TDDFT reference (specifically, TD-B3LYP/TZVP). This is rather good result given that uncertainties of the reference calculations are in the same order of magnitude (MAD of TDDFT reference vs CC2 reference is 0.014 Å, Figure 7b).⁴⁹ Accurate experimental values are very hard to obtain. However, for the available experimental bond lengths in the ExGeom data set, AIQM1 gives better or similar predictions compared to TDDFT and CC2 for C–O bond in 1nπ* and 3nπ* excited states, while its error is much bigger for the 3ππ* excited state of formaldehyde (Figure 7c).

Overall, AIQM1 seems to be a better choice than current routinely used QM methods in terms of performance/cost ratio at least for some types of excitations, which holds a great promise for using this method for exploration of dynamical properties arising from the manifold of electronic states, e.g., by performing nonadiabatic excited-state dynamics, which should be an interesting topic for future explorations. In any case, the AIQM1 method is only the first step in the direction of creating a general-purpose AI-based method for excited-state simulations — an important, but open topic in chemistry⁵⁰ — as obviously training models on excited-state properties will be crucial for future improvements.

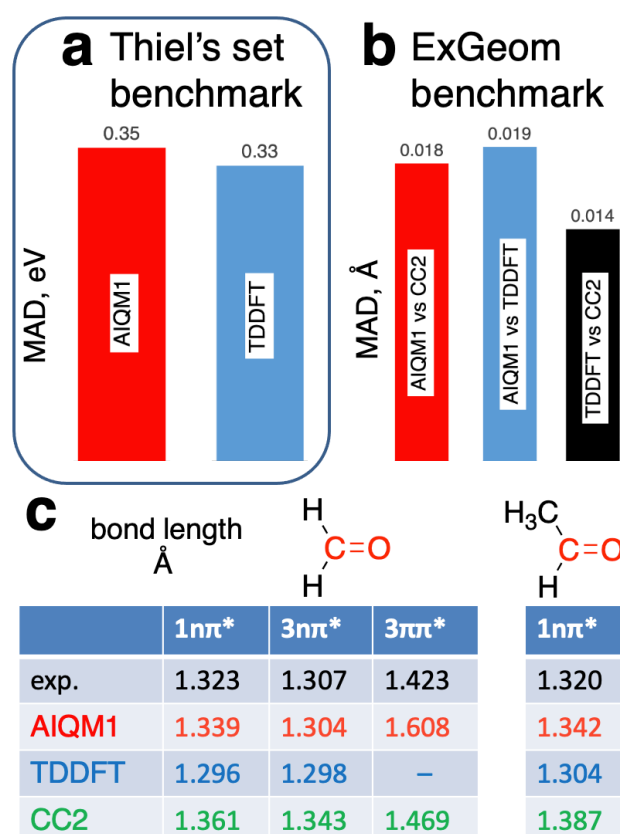


Figure 7. **Performance of AIQM1 for excited states.** a) Mean absolute error (MAD) in vertical excitation energies for Thiel's benchmark set⁴⁹. b) MAD in bond lengths for the ExGeom benchmark set^{17,49}. c) Bond length for C–O bond length for formaldehyde and acetaldehyde as compared to experiment (one value for TDDFT is missing due to the failed geometry optimization at this level⁴⁹). Values for TDDFT, CC2, TDDFT vs CC2, and experiment are taken from Ref. 49.

Discussion

After initial excitement about great promises AI holds for substituting QM methods, the focus is shifting towards tighter integration of AI with QM instead of substituting QM altogether. This shift is motivated by the need to incorporate correct physical behavior of QM methods, while at the same time exploiting great ability of AI to improve low-level QM methods' accuracy without compromising their speed.

In this work, we have made a step towards creating general purpose AI-improved QM methods useful for a variety of applications out-of-the-box. Our approach AIQM1 synergistically combines the best of two worlds — transferability of QM and high accuracy of AI approaches. The success of this approach only became possible with great advances

over recent years in methodology development of both QM and AI components as well as generation of numerous carefully curated, high quality reference data. Thus, AIQM1 allows very accurate prediction of ground-state properties such as energies and geometries of closed-shell, neutral organic compounds approaching the gold-standard CCSD(T)/CBS at the speed of semiempirical QM methods. Remarkably, it has improved accuracy also for other cases, not explicitly considered during training of its NN part, e.g., for charged species, showcasing the benefits of using physically-motivated AI. Thus, AIQM1 method has the potential to become a very useful tool for routine simulations with high accuracy.

It is only the beginning of the exciting road for AI-improved QM methods for general-purpose applications. In the near future we expect tighter integration of AI with QM, further optimizing both AI and QM parts, training on more and higher quality reference data, and further extending transferability and accuracy for all properties of interest to chemists and physicists.

Methods

Neural network training

The neural network training and evaluation was performed with the TorchANI software⁵¹. Each NN-part of AIQM1@DFT* consists of an ensemble of eight ANI-type NNs, which provides better accuracy according to our tests. The ensemble was trained similar to the previous procedure,¹³ i.e., the data set was split into 9 equal parts, with one part held out for testing and the remaining 8 parts were used as cross-validation splits for training eight networks. Each network was trained on 7 cross-validation splits and validated on one split using standard rotation of splits. During the training of AIQM1@DFT*, we stopped training NN after 1000 epochs, because we found that longer training does not improve much the performance for the validation set, but deteriorates performance for some of the external data sets. Transfer learning was then used to refit above eight ANI-type networks to 80% of the entire set with CCSD(T)*/CBS values to obtain the final NN part of AIQM1 consisting of ensemble of 8 NNs; other 10% were used as the validation set and remaining 10% as the hold-out test set. The atomic contributions obtained by linear fitting¹³ are listed in the Supplementary Information for our methods.

Calculation of enthalpies

The enthalpies at 298 K were calculated within harmonic oscillator and rigid rotor

approximation in our locally modified version of the MNDO program⁵². Calculating heats of formation requires the evaluation of the atomization energies, which depend on the choice of the atomic energies. Atomic energies calculated with CCSD(T)*/CBS used for fitting NN-part of AIQM1 lead to large errors in atomization energies even for moderate-sized molecules such as naphthalene (error of 25.4 kcal/mol with respect to CCSD(T)/CBS, where the two-point extrapolation scheme was used with cc-pVDZ cc-pVTZ basis sets), thus we fitted atomic energies of H, C, N, and O elements to reduce the error in heats of formation in the CHNO set. Atomic energies are reported in the Supplementary Information.

Electronic structure and benchmark calculations

All ODM2 and ODM2* calculations were carried out with the MNDO program.⁵² CCSD(T)*/CBS calculations were performed with the ORCA 4.2.0 software package^{53,54} using the procedure described previously.¹³ The ω B97X-D4 calculations were performed with ORCA 4.2.0, and ω B97X-D calculations were performed with Gaussian 16⁵⁵. The ω B97X/6-31G* calculations were performed with Gaussian 16, while ω B97X/def2-TZVPP calculations were performed with ORCA 4.2.0. D4-dispersion corrections were calculated with the dftd4 program.⁵⁶ We performed benchmarks of AIQM1 and AIQM1@DFT with the locally modified version of the MNDO program⁵² interfaced to TorchANI⁵¹ and dftd4⁵⁶. All the data for energy benchmarks can be found in the Supplementary Information.

References

1. Raghavachari, K., Trucks, G. W., Pople, J. A. & Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **157**, 479-483 (1989).
2. Thomas, J. R. *et al.* The balance between theoretical method and basis set quality: A systematic study of equilibrium geometries, dipole moments, harmonic vibrational frequencies, and infrared intensities. *The Journal of Chemical Physics* **99**, 403-416 (1993).
3. Helgaker, T., Gauss, J., Jørgensen, P. & Olsen, J. The prediction of molecular equilibrium structures by the standard electronic wave functions. *The Journal of Chemical Physics* **106**, 6430-6440 (1997).
4. Bak, K. L. *et al.* The accurate determination of molecular equilibrium structures. *The Journal of Chemical Physics* **114**, 6548-6556 (2001).
5. Husch, T., Vaucher, A. C. & Reiher, M. Semiempirical molecular orbital models based on the neglect of diatomic differential overlap approximation. *Int. J. Quantum Chem.* **118**, e25799 (2018).
6. Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **87**, 897-923 (2015).
7. Dral, P. O. Quantum Chemistry in the Age of Machine Learning. *J. Phys. Chem. Lett.* **11**, 2336-2347 (2020).
8. von Lilienfeld, O. A., Müller, K.-R. & Tkatchenko, A. Exploring chemical compound space with quantum-based machine learning. *Nat. Rev. Chem.* **4**, 347-358 (2020).
9. Devereux, C. *et al.* Extending the Applicability of the ANI Deep Learning Molecular Potential to Sulfur and Halogens. *J. Chem. Theory Comput.* **16**, 4192-4202 (2020).
10. Smith, J. S., Isayev, O. & Roitberg, A. E. ANI-1, A data set of 20 million calculated off-equilibrium conformations for organic molecules. *Sci. Data* **4**, 170193 (2017).
11. 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.* **8**, 3192-3203 (2017).

12. Smith, J. S., Nebgen, B., Lubbers, N., Isayev, O. & Roitberg, A. E. Less is more: Sampling chemical space with active learning. *J. Chem. Phys.* **148**, 241733 (2018).
13. Smith, J. S. *et al.* Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nat. Commun.* **10**, 2903 (2019).
14. Manzhos, S. Machine learning for the solution of the Schrödinger equation. *Mach. Learn.: Sci. Technol.* **1**, 013002 (2020).
15. Westermayr, J., Gastegger, M., Schütt, K. T. & Maurer, R. J. Perspective on integrating machine learning into computational chemistry and materials science. *J. Chem. Phys.* **154**, 230903 (2021).
16. Zubatyuk, T. & Isayev, O. Development of Multimodal Machine Learning Potentials: Toward a Physics-Aware Artificial Intelligence. *Acc. Chem. Res.* **ASAP** (2021).
17. Dral, P. O., Wu, X. & Thiel, W. Semiempirical Quantum-Chemical Methods with Orthogonalization and Dispersion Corrections. *J. Chem. Theory Comput.* **15**, 1743–1760 (2019).
18. Caldeweyher, E., Bannwarth, C. & Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **147**, 034112 (2017).
19. Caldeweyher, E. *et al.* A generally applicable atomic-charge dependent London dispersion correction. *J. Chem. Phys.* **150**, 154122 (2019).
20. Axilrod, B. M. & Teller, E. Interaction of the van der Waals Type Between Three Atoms. *J. Chem. Phys.* **11**, 299–300 (1943).
21. Muto, Y. Force between nonpolar molecules. *Proc. Phys. Math. Soc. Jpn.* **17**, 629–631 (1943).
22. Folmsbee, D. & Hutchison, G. Assessing conformer energies using electronic structure and machine learning methods. *Int. J. Quantum Chem.* **121**, e26381 (2020).
23. Zubatyuk, R., Smith, J. S., Leszczynski, J. & Isayev, O. Accurate and transferable multitask prediction of chemical properties with an atoms-in-molecules neural network. *Sci. Adv.* **5**, eaav6490 (2019).
24. Ramakrishnan, R., Dral, P. O., Rupp, M. & von Lilienfeld, O. A. Big Data Meets Quantum Chemistry Approximations: The Δ -Machine Learning Approach. *J. Chem. Theory Comput.* **11**, 2087–2096 (2015).
25. Pan, S. J. & Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **22**, 1345–1359 (2010).
26. Hu, L. H., Wang, X. J., Wong, L. H. & Chen, G. H. Combined first-principles calculation and neural-network correction approach for heat of formation. *J. Chem. Phys.* **119**, 11501–11507 (2003).
27. Wu, J. & Xu, X. The X1 method for accurate and efficient prediction of heats of formation. *J. Chem. Phys.* **127**, 214105 (2007).
28. Dandu, N. *et al.* Quantum-Chemically Informed Machine Learning: Prediction of Energies of Organic Molecules with 10 to 14 Non-hydrogen Atoms. *J. Phys. Chem. A* **124**, 5804–5811 (2020).
29. Wan, Z., Wang, Q. D. & Liang, J. Accurate prediction of standard enthalpy of formation based on semiempirical quantum chemistry methods with artificial neural network and molecular descriptors. *Int. J. Quantum Chem.* **121**, e26441 (2021).
30. Dral, P. O. *et al.* Semiempirical Quantum-Chemical Orthogonalization-Corrected Methods: Theory, Implementation, and Parameters. *J. Chem. Theory Comput.* **12**, 1082–1096 (2016).
31. Dral, P. O., Wu, X., Spörkel, L., Koslowski, A. & Thiel, W. Semiempirical Quantum-Chemical Orthogonalization-Corrected Methods: Benchmarks for Ground-State Properties. *J. Chem. Theory Comput.* **12**, 1097–1120 (2016).
32. Curtiss, L. A., Raghavachari, K., Redfern, P. C. & Pople, J. A. Assessment of Gaussian-3 and density functional theories for a larger experimental test set. *J. Chem. Phys.* **112**, 7374–7383 (2000).
33. Curtiss, L. A., Redfern, P. C. & Raghavachari, K. Gaussian-4 theory. *J. Chem. Phys.* **126**, 084108 (2007).
34. Curtiss, L. A., Redfern, P. C. & Raghavachari, K. Gaussian-4 theory using reduced order perturbation theory. *J. Chem. Phys.* **127**, 124105 (2007).
35. Weber, W. *Ein neues semiempirisches NDDO-Verfahren mit Orthogonalisierungskorrekturen: Entwicklung des Modells, Implementierung, Parametrisierung und Anwendung* DOI, Universität Zürich, (1996).
36. Luo, S., Zhao, Y. & Truhlar, D. G. Validation of electronic structure methods for isomerization reactions of large organic molecules. *Phys. Chem. Chem. Phys.* **13**, 13683–13689 (2011).
37. Peverati, R., Zhao, Y. & Truhlar, D. G. Generalized Gradient Approximation That Recovers the Second-Order Density-Gradient Expansion with Optimized Across-the-Board Performance. *J. Phys. Chem. Lett.* **2**, 1991–1997 (2011).
38. Sellers, B. D., James, N. C. & Gobbi, A. A Comparison of Quantum and Molecular Mechanical Methods to Estimate Strain Energy in Druglike Fragments. *J. Chem. Inf. Model.* **57**, 1265–1275 (2017).
39. Hawkins, J. M., Meyer, A., Lewis, T. A., Loren, S. & Hollander, F. J. Crystal structure of osmylated C₆₀: confirmation of the soccer ball framework. *Science* **252**, 312–313 (1991).

40. Hedberg, K. *et al.* Bond lengths in free molecules of buckminsterfullerene, C₆₀, from gas-phase electron diffraction. *Science* **254**, 410–412 (1991).
41. Liu, S., Lu, Y. J., Kappes, M. M. & Ibers, J. A. The structure of the C₆₀ molecule: X-ray crystal structure determination of a twin at 110 K. *Science* **254**, 408–410 (1991).
42. Yannoni, C. S., Bernier, P. P., Bethune, D. S., Meijer, G. & Salem, J. R. NMR determination of the bond lengths in C₆₀. *J. Am. Chem. Soc.* **113**, 3190–3192 (2002).
43. Zhao, Y. & Truhlar, D. G. Construction of a generalized gradient approximation by restoring the density-gradient expansion and enforcing a tight Lieb-Oxford bound. *J. Chem. Phys.* **128**, 184109 (2008).
44. Peverati, R. & Truhlar, D. G. Exchange-Correlation Functional with Good Accuracy for Both Structural and Energetic Properties while Depending Only on the Density and Its Gradient. *J. Chem. Theory Comput.* **8**, 2310–2319 (2012).
45. Rezac, J., Riley, K. E. & Hobza, P. S66: A Well-balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **7**, 2427–2438 (2011).
46. Goerigk, L. & Grimme, S. Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals-Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **7**, 291–309 (2011).
47. Anacker, T. & Friedrich, J. New accurate benchmark energies for large water clusters: DFT is better than expected. *J. Comput. Chem.* **35**, 634–643 (2014).
48. Kolb, M. *Ein neues semiempirisches Verfahren auf Grundlage der NDDO-Näherung: Entwicklung der Methode, Parametrisierung und Anwendung* DOI, Belgische Universität-Gesamthochschule Wuppertal, (1991).
49. Tuna, D., Lu, Y., Kosłowski, A. & Thiel, W. Semiempirical Quantum-Chemical Orthogonalization-Corrected Methods: Benchmarks of Electronically Excited States. *J. Chem. Theory Comput.* **12**, 4400–4422 (2016).
50. Dral, P. O. & Barbatti, M. Molecular Excited States Through a Machine Learning Lens. *Nat. Rev. Chem.* **5**, 388–405 (2021).
51. 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.* **60**, 3408–3415 (2020).
52. Thiel, W. MNDO, development version (Max-Planck-Institut für Kohlenforschung, Mülheim an der Ruhr, 2019).
53. Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1327 (2018).
54. Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 73–78 (2012).
55. Frisch, M. J. *et al.* Gaussian 16, Rev. A.01 (Wallingford, CT, 2016).
56. Caldeweyher, E., Ehlert, S. & Grimme, S. DFT-D4, Version 2.5.0 (Mulliken Center for Theoretical Chemistry, University of Bonn, 2020).

Acknowledgements

P.O.D. acknowledges funding by the National Natural Science Foundation of China (No. 22003051) and via the Lab project of the State Key Laboratory of Physical Chemistry of Solid Surfaces. O.I. acknowledges support from the National Science Foundation (NSF) CHE-1802789 and CHE-2041108. O.I. and R.Z. acknowledge Extreme Science and Engineering Discovery Environment (XSEDE) award CHE200122, which is supported by NSF grant number ACI-1053575. This research is part of the Frontera computing project at the Texas Advanced Computing Center. Frontera is made possible by the National Science Foundation award OAC-1818253.

Author contributions

P.O.D. conceived the idea. P.Z. carried out method implementation with the help from R.Z. P.Z. carried out all calculations. P.Z. and P.O.D. performed data analysis and visualization. R.Z. generated the NN training data. P.O.D. wrote the manuscript with assistance from P.Z. All authors provided critical feedback and helped shape the research, analysis, and manuscript. P.O.D., W.W. and O.I. supervised and acquired funding for the project.

Supplementary Information

The atomic contributions obtained by linear fitting. Atomic energies to calculate AIQM1 enthalpies of formation. Raw data with energies for benchmarks.

Competing interests

The authors declare no competing interests.

Materials & Correspondence

Correspondence and requests for materials should be addressed to P.O.D. and O.I.