Communications in Information and Systems
Volume 18, Number 4, 229–250, 2018

Prediction of molecular energy using deep tensor neural networks

Yan Li, Han-Yi Min, Zi-Bing Dong, Tian Yuan, Xiao-Qi Li, Pei-Jun Xu†, and Guo-Hui Li†

In this paper, we propose a combined scheme called Quantum Mechanics and Deep Tensor Neural Network (QM-DTNN) to address the challenges related to the prompt and accurate calculation of the physicochemical properties of a protein. In QM-DTNN, a protein is decomposed into individual amino acid units that are treated with molecular caps. The physicochemical properties (molecular energy) of amino acid units are predicted using DTNNs, which are trained by QM (*ab Initio*) data. The training, validating, and testing data sets are made of conformations drawn through enhanced sampling in specific collective variable spaces. The inputs of DTNNs include pair-wise inter-distance and nuclear charges of an amino acid unit. The outputs of DTNNs, which are the physicochemical properties of amino acid units, are calculated using QM. The three typical amino acid units (i.e., Arginine, Lysine, and Tryptophan) are used to demonstrate the feasibility of QM-DTNN. The prediction results demonstrated good correlations to QM data. The proposed scheme reduces the computational time considerably compared to that of QM calculation with acceptable precision loss.

1. Introduction

Molecular dynamics (MD) simulations have evolved into mature techniques that can be used to effectively understand structure-to-function relationships. The MD of biological molecules such as proteins, DNA complexity, and RNA complexity, is a focus of study in current research on biological molecules [18]. In addition, physicochemical properties such as molecular energy and atom charge play important roles in providing a description of

Project 31700647 supported by NSFC. Project 21573217 supported by NSFC. Y. Li. and H.-Y. Min contributed equally to this work.
†Pei-Jun Xu and Guo-Hui Li are the corresponding author of this paper.

the protein folding process [36, 47]. Therefore, it can be extremely challenging to rapidly and accurately obtain the physicochemical properties of proteins in biology and computational chemistry.

Because of the large number of atoms in proteins (in addition to biomacromolecules), the standard full quantum mechanical (QM) or *ab initio* calculation cannot be performed by computers. Most theoretical studies on biological molecules generally employ classical force fields that are built on pair-wise atomic interaction potentials. The principle idea governing classical force fields is the decomposition of total energy into low-dimensional bonding two-, three-, and four-body terms that represent covalent bonds, bonding angles, and dihedral angles, respectively. In addition, electrostatic and van der Waals interactions are expressed using Coulomb's law and the Lennard-Jones potential, respectively, which describe non-bond interactions. In terms of low-dimensional bonding, only the immediate environment is considered. Despite the success of classical force field approaches in many applications, limitations that reflect the chemical process, especially the polarized effect and bond breaking reactions, still exist. The polarized force fields [40, 50, 68, 70] were proposed. In addition, the flutenting charge model [17, 25] and other charge-related models [21, 65] were introduced to accurately illustrate the polarized effect in a certain process. To reduce the computational time, the CG model was applied [43]. The QM calculations of interaction energies are often required to illustrate electron transferring. In the hardware aspection, the Anton machine [60, 61] is a special-purpose system used for MD simulations of proteins and other biological macromolecules. The most widely accepted scheme involves the use of a GPU for MD simulation [19, 22, 37, 44, 46, 49, 51, 63] and the simultaneous calculation of biomolecular systems using different GPU acceleration cards.

In recent years, artificial intelligence (AI), particularly deep learning approaches [28, 31, 38, 48], have made major breakthroughs in various fields and have been more successful than those approaches of the past one or two decades. In images processing [12, 42], speech recognition [32], and many other text [3] and video-related studies [26, 59, 69], deep learning has been the most preferred choice of tools [27, 35]. In addition, deep learning techniques have demonstrated tremendous potential in various problems across a range of disciplines, including materials design, chemical synthesis, and drug discovery [6, 10, 23, 30, 33, 58, 66]. Over the past 20 years, researchers have used neural networks (NNs) and deep learning to study the protein folding and reaction processes. Early in 2006, Lorenz and Matthias used NNs to depict the reaction process [41]. In 2013, Montavon, Rupp used machine learning models to simultaneously predict multiple electronic ground-state and excited-state properties, including atomization energy, polarizability, frontier orbital eigenvalues, ionization potential, electron affinity, and excitation energies [45]. Behler used symmetry functions as inputs to NNs to depict high-dimensional *ab initio* potential-energy surfaces [7–9]. Shen, Wu, and Yang predicted the energy differences in two-level QM calculations using NNs and merged the predictions in the MD process [62, 67]. Browning and Nicholas used genetic algorithms to optimize training set compositions consisting of tens of thousands of small organic molecules for predicting small organic molecular properties [11]. Based on a local model of interatomic interactions, Gubeav and Konstantin proposed a machine learning algorithm for predicting molecular properties in 2017 [24]. In 2017, Schutt, Kristof et al. proposed the use of a continuous-filter convolutional network (SchNet) to model quantum interactions among molecules [57]. In 2017, Tobias and Christine combined coarse-grained (CG) simulation models and NNs to extract high-dimensional free energy surfaces (FESs) from MD simulation trajectories [39]. In 2018, Kamath and Aditya compared NNs and the Gaussian process regression to depict potential energy surfaces [34]. However, to the best of the author's knowledge, there are several meaningful references to DTNN that are beyond the scope of this paper.

The structure and properties of molecules are determined by the laws of QMs. With the aim of providing a direct functional relationship between atomic conformation and molecular energy, employing one or more artificial NNs can be a framework for the calculation of molecular energy using a deep learning approach with high-precision QM calculation, which is QM-DTNN.

This paper is organized as follows. In the methodology section, the basic idea and steps of QM-DTNN are described. Numerical validations are performed and the results are reported in the results section; the detailed comparisons of QM-DTNN results are obtained concerning the physicochemical properties of the three typical amino acid units. The discussion section includes a summary and discussion of present studies. Furthermore, improvements in QM-DTNN are included as future work.

2. Methodology:QM-DTNN

The basic principle of the proposed framework involves the partitioning of energy of a large molecule into relatively smaller amino acid energy units that can be easily calculated and predicted using QM-deep learning tools. The dividing scheme in MFCC is used to partition a macromolecule into amino acid units. The enhanced sampling, multi-walk metadynamics, is used to sample sufficient conformations to cover two collective variable (CV) space as convergent as possible. The physicochemical properties (molecular energy) of amino acid units are obtained through QM calculation. In the machine learning module, each DTNN is trained with relative data sets and is applied for energy prediction. The entire framework for predicting energy using deep learning tools is called QM-DTNN.

The schematic flowchart of QM-DTNN is shown in Figure 1.

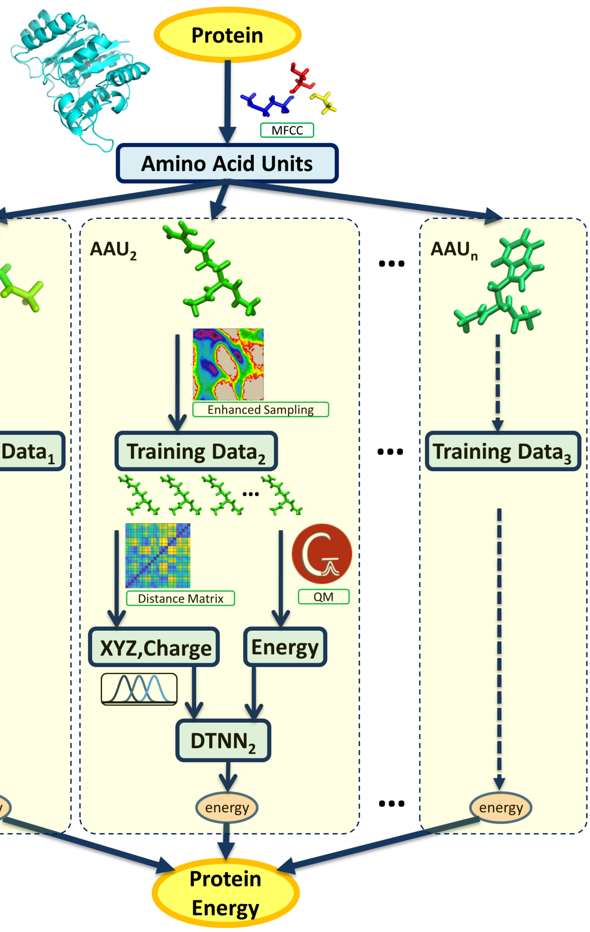

Figure 1: The schematic flowchart of QM-DTNN.

2.1. MFCC to obtain amino acid units

As macromolecules, proteins include linear polymers built from a series of up to 20 different standard amino acids. A protein is split into smaller pieces that are primarily amino acid units. This idea was originally proposed in molecular fractionation using conjugate caps (MFCC), which is another effective *ab initio* framework for calculating the interactions among protein energies [71]. In MFCC, the protein is decomposed into individual amino acid units that are treated with proper molecular caps. In the present model, $ACE-$ and $NME-$ are introduced as caps to both ends of an amino acid unit, which are presented in panel B of Figure 2. Here, we briefly illustrate the partitioning process, and the splitting of a protein (PDBID:1BYI [55]) into amino acid units [55] is shown in panel C of Figure 2.

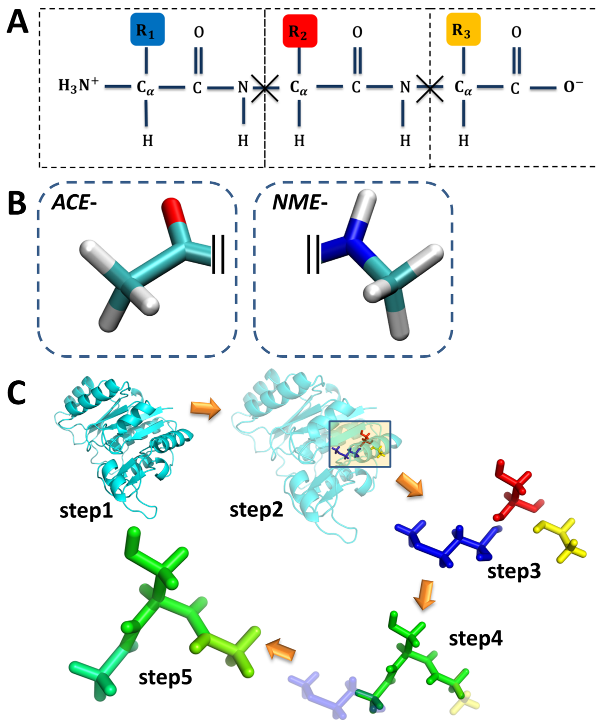

Figure 2: A frame work of the splitting process in MFCC. A: Graphical representation of an extended tripeptide and the locations of the cuts where conjugate caps are introduced. B: The $ACE-$ and $NME-$ are two caps added to the amino acid units. C: Steps for adding conjugate caps to a SER drawn from a typical protein (PDBID:1BYI).

Step 1. Choose a protein conformation from PDB Bank [54].

Step 2. Select the target amino acid. Here, Serine (SER) is randomly chosen as the target amino acid. There are 12 SER in 1BYI. To illustrate the drawing process, the SER at the end of the peptide chain is not selected.

Step 3. In this example, two neighboring amino acid units near the target amino acid unit include Methionine (MET) as a blue stick model and Alanine (ALA) as a yellow stick model.

Step 4. Add conjugate caps on both sides of the target amino acid unit. The MET side is replaced by the $ACE-$ cap. In addition, the ALA side is replaced by the $NME-$ cap. In the $MET-$ end, the CO bond is obtained from SER. The position of C in CH_3 in the $ACE-$ cap is the position C_α of neighboring MET. The bond angles and dihedral angles of atom H in $CH_3 - CO$ are the same as the atom H connected to C_α in MET. The bond length of $C - H$ is set to $1.09\AA$. In the ALA end, the CH_3 is set as the same strategy as the MET end.

Step 5. Output the target amino acid SER with two caps.

2.2. Enhanced sampling to build data sets

In an NN, especially in a deep learning model, the properties of a training set determine the training process and the prediction precision. Ideally, the training set covers the entire conformation space and contains as many conformations as possible. In practice, because of the limitations in computational resources and time, it is unrealistic to include all possible conformations in data sets. However, we can ensure that the conformation in data sets represents the entire conformation space as convergent as possible. We applied conformations in two primary ways. One is the standard MD of the amino acid unit. The other is the enhanced sampling of the amino acid unit.

The basic training set is the amino acid units drawn from 100+ small protein MD trajectories. The mean MD time is 20 ns. Because of high-energy barriers, the traditional MD cannot explore the CV space entirely, even in a long MD runtime. Here, CVs serve as the basis for describing the conformational space.

To enrich the conformational diversity, an enhanced sampling technology is applied. Bias-exchange metadynamics is a powerful technique that can be used to reconstruct the free energy and to enhance the conformational search in complex conformational systems. A simplified version of BEMETA, such as multiple walkers metadynamics (MWM) [5, 52], is applied. It is the simplest way of parallelizing a metadynamics calculation. In MWM, multiple simulations of the same system are implemented in parallel using metadynamics on the same set of CVs. The deposited bias is shared among the replicas in such a way that the history-dependent potential depends on the entire history. For amino acid LYS, about 2 μs MD trajectory of LYS in a water solution is used to draw the conformation for building the data set.

In an amino acid simulation model, two backbone dihedral angles (torsion angles) ϕ and ψ are chosen as CVs. Two replicas of the system are employed, each biased with a different CV. In addition, AMBER99SB [29] is applied as the force field for the entire MD process.

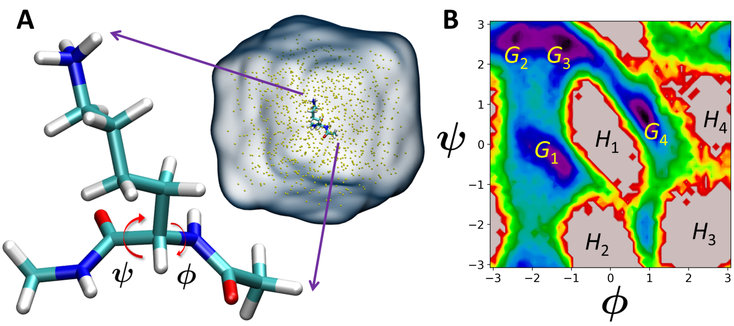

Figure 3: An illustration of enhanced sampling. A: An amino acid unit LYS in a water box. Two torsion angles ϕ and ψ are chosen as CVs. B: The free energy profile of 2 CVs from 2μ s enhanced sampling.

A LYS is placed at the center of a water box. The LYS amino acid unit is surrounded by 867 water molecules. Na and Cl ions are placed in a water box to balance the electron of the entire system. Minimization with 1 ns is implemented to optimize the conformation reasonably.

The free energy profile of two CVs is in panel B of Figure 3. It clearly demonstrates that most of the meaningful conformations in $G_i, i = 1, 2, 3, 4$ are well sampled. The halo part $H_j, j = 1, 2, 3, 4$ are the conformations that rarely appear as high energy. Therefore, the data set is considered sufficiently convergent for deep learning training.

2.3. DTNN to predict energy

Following the splitting process in MFCC and the enhanced sampling process via multi-walk metadynamics, the amino acid unit data sets are ready for further preparation in QM calculation.

In the entire QM-DTNN framework, the deep learning modular, DTNN, is key to predicting physicochemical properties. Therefore, we illustrate this part in detail and list the differences between the original DTNN and QM-DTNN proposed here. DTNN was first proposed by Kristof and Farhad [56]. The purpose of this deep learning framework is to predict the energy of a small molecule (20 is the highest atom number in the GDB9 database [53]) and its physicochemical properties, including the potential of aromatic rings. DTNN has also demonstrated good predicting ability. We extend the DTNN to larger molecules and apply different DTNNs to the relative amino acid units.

2.3.1. The input of DTNN. Because deep learning is effective for handling a considerable size of image files with relative properties, the amino acid unit is converted into an image format. A typical color image is stored in red, green, blue (RGB) pixels, and one image can be replaced by three layers of different color pixels. The amino acid unit is stored in the same way but not limited to three layers. The basis number(layer number) is determined by the atom numbers of the amino acid unit. The amino acid unit is converted into the pair-wise inter-atom distance matrix D, which is shown in panel B in Figure 4. D_{ij} is the distance between $atom_i$ and $atom_j$. This matrix construction satisfies the rotational and translational invariance, which is convenient for MD and net training.

$$D = \begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1n} \\ D_{21} & D_{22} & \cdots & D_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ D_{n1} & D_{n2} & \cdots & D_{nn} \end{bmatrix}_{n\times n} \tag{1}$$

Geometric information is not sufficient to distinguish the conformation. We also feed the atom types, which is a vector of nuclear charges, as the input of DTNN. The vector is shown as Z in panel B of Figure 4.

$$Z = [Z_1, Z_2, \ldots, Z_n] \tag{2}$$

in which n is the number of atoms in an amino acid unit.

2.3.2. The output of DTNN. The molecular energies are the output of DTNN. They are calculated using highly efficient QM soft-package Gaussian 09 [16].

2.3.3. Gaussian feature expansion of D_{ij}. Each element $D_{ij} \in D$ is spread across many dimensions of a uniform grid of Gaussians basis. This

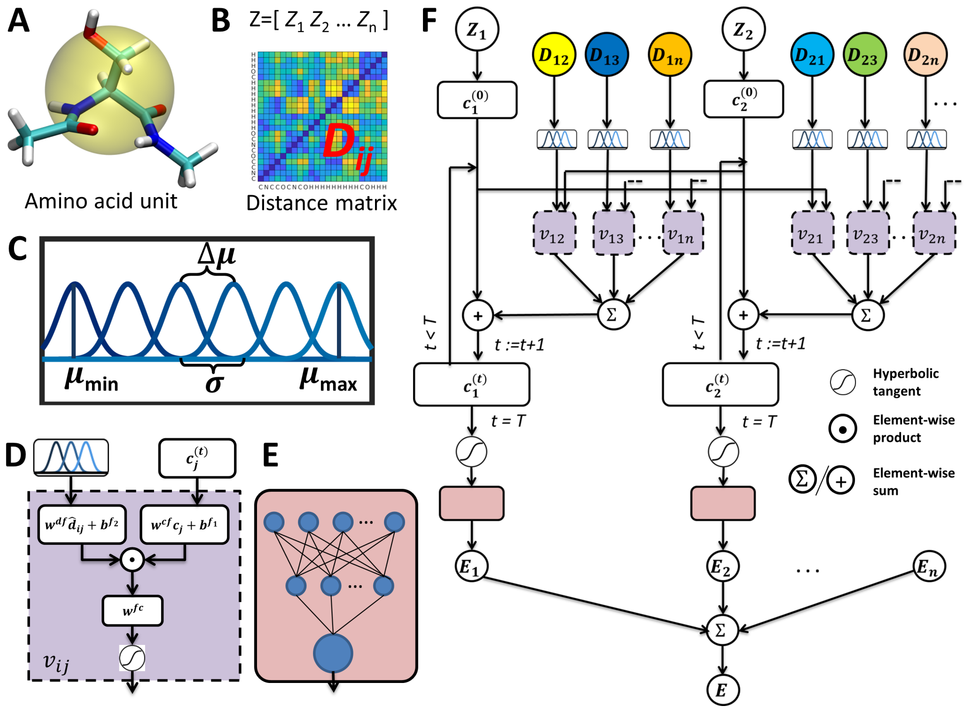

Figure 4: The framework of DTNN. A: An amino acid unit with threshold as cutoff. B: The pair-wise inter-atomic distance matrix D and the Z vector of nuclear charges. C: A series of uniform grids of Gaussians forms the basis. D: The interaction module with element-wise product of two tensors. E: The fully connected layers and the output sum E_i. F: The whole architecture of DTNN with inter-atomic distances and charges of atoms as inputs. The output E is the energy of each amino acid unit.

process is shown in panel C of Figure 4.

$$\hat{d}_{ij} = \left[exp \left(-\frac{D_{ij} - (\mu_{min} + k\Delta\mu)^2}{2\sigma^2} \right) \right] \tag{3}$$

in which $\Delta\mu$ is the gap between two Gaussians with width σ. In this process, a conformation is converted to an image with multi-layers, and the pixel value in each layer is decided by the distance between two atoms in a conformation.

After the decomposition, the atom is presented by atomic descriptors with B coefficients.

$$C_i^{(0)} = C_{Z_i} \in R^B \tag{4}$$

In addition, B is the number of basis functions. The initial value of C_z is set by random coefficients according to

$$C_z \sim N(0, 1/\sqrt{B}) \tag{5}$$

2.3.4. *T* interaction passes. Each coefficient vector $c_i^{(t)}$, corresponding to atom i after t passes, is corrected based on interactions with other atoms of the molecule.

$$c_t^{(t+1)} = c_i^{(t)} + \sum_{j \neq i} v_{ij} \tag{6}$$

in which v_{ij} is defined as

$$v_{ij} = tanh\left[W^{fc}\left((W^{cf}c_j + b^{f_1}) \circ (W^{df}\hat{d}_{ij} + b^{f_2})\right)\right] \tag{7}$$

where '$\circ$' indicates an element-wise multiplication. In addition, W^{cf}, b^{f_1}, W^{df}, b^{f_2}, and W^{fc} are the weight matrices and corresponding biases of atom representations, distances, and resulting factors, respectively [64]. This sub-process is shown in panel D of Figure 4.

2.3.5. Sum of $\boldsymbol{E_i}$. The energy of the entire amino acid unit is the sum of the energies of each atom E_i.

$$o_i = tanh(W^{out_1}c_i^{(T)} + b^{out_1}) \tag{8}$$

The final stage includes n_l fully connected layers. $n_l = 2$. The hidden layer o_i possesses n_{hid} neurons, which is shown in panel E of Figure 4.

$$\hat{E}_i = W^{out_2}o_i + b^{out_2} \tag{9}$$

The $\hat{E}_i$ is shifted to the mean E_μ and scaled by the s.d. E_σ of the energy per atom estimated on the training set.

$$E_i = E_\sigma \hat{E}_i + E_\mu \tag{10}$$

The entire architecture of DTNN is shown in panel F of Figure 4.

3. Results

To demonstrate the versatility of the proposed QM-DTNN, three typical amino acid units were used to train the models with up to 15 interactions. Arginine (ARG), Lysine (LYS), and Tryptophan (TRP) are three big amino acid units with more than 30 atoms (including two caps). Because the network can handle smaller molecules in the training set with a smaller error, the networks, efficiently controlling the three amino acid units, can also provide reasonable prediction of other small amino acid units.

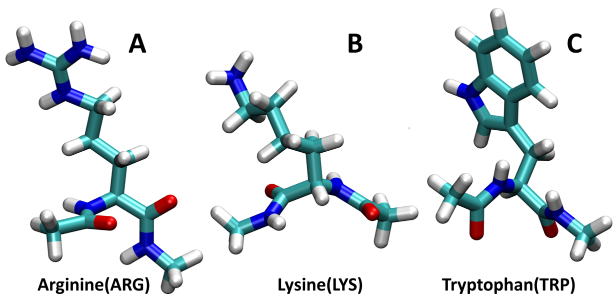

Figure 5: Three typical amino acid units. A: Arginine. B: Lysine. C: Tryptophan.

The parameters of DTNNs are listed in Table 1.

Table 1: Training parameters in DTNNs of 3 Amino Acid Units

DTNN parameter	ARG	LYS	TRP
Training set	15000	54000	54000
Validating set	1000	6000	6000
Live Testing set	1000	2000	10000
Testing set	7271	29591	23630
T passes	12	10	6
Cutoff	3	3	3
Basis number	50	50	60
Hidden layer	100	100	120

We employ three sampled data sets of ARG, LYS, and TRP, including 20000+, 90000+, and 90000+ molecules with up to 20 heavy (C, N, O) atoms. All the conformations in the data set are different from each other.

The training sets are randomly chosen from the data set. The left conformations include the validating set and the testing set. There are no overlaps between any of the three sets.

All network models were trained and executed on an NVIDIA Tesla K80 GPU with deep learning framework TensorFlow [1, 2].

The regressions of the energies calculated from QM and predicted by QM-DTNN in the testing data set are shown in Figure 6. All the energies of QM and QM-DTNN have been transformed to the range of [0-70].

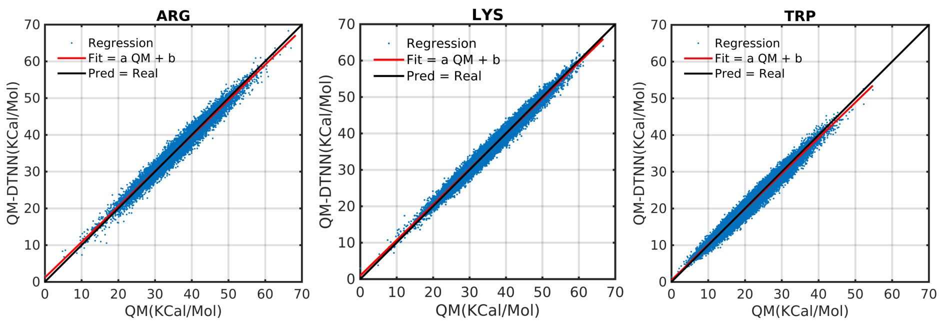

Figure 6: The regressions of QM and QM-DTNN prediction.

The regression is judged by 4 static parameters.

$$Fit = a \times QM + b \tag{11}$$

The parameters a, b are fitted by a least-squares method. All three a are larger than 0.95. In addition, the fit line is shown in red in each sub-figure of Figure 6.

The err_{mean} is the mean value of the absolute prediction error. The err_{std} is the istandard deviation(std) value of the absolute prediction error. All three err_{mean} are approximately 1 $KCal/Mol$. All three err_{std} are less than 0.85 $KCal/Mol$.

The static results of three DTNNs are listed in Table 2.

All four parameters exhibit good correlation between QM-DTNN and QM energies.

4. Discussion

The differences between DTNN and QM-DTNN include the following:

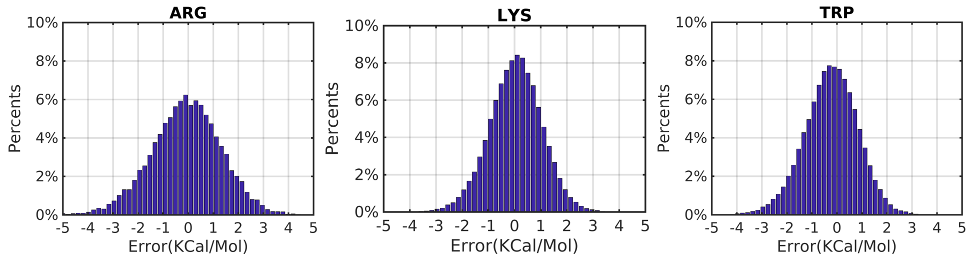

Figure 7: The error distribution of prediction.

Table 2: Prediction results of DTNNs of 3 Amino Acid Units

static parameter	ARG	LYS	TRP
a	0.959	0.962	0.957
b	1.188	1.301	0.582
err_{mean} (KCal/Mol)	1.09	1.01	1.02
err_{std} (KCal/Mol)	0.85	0.82	0.84

- The original DTNN predicts the energy of different small organic molecules in the entire database. Each DTNN of QM-DTNN only handles its own amino acid units with a fixed atom number and the same topology.
- The original DTNN predicts molecules with up to 20 atoms. DTNNs of QM-DTNN predict molecules with larger atomic numbers, with 23 being the least. Commonly, the atom number is approximately 30.
- The DTNN obtains a smaller error compared to QM-DTNN, but it cannot handle large molecules. The QM-DTNN can split large molecules into small pieces and implement the prediction.

Compared to QM calculation, QM-DTNN saves computational resources in at least two magnitudes. If more experienced methods and base sets are applied, QM-DTNN will save more computation time.

Because only three amino acid units, with high-atom numbers, have been introduced to the QM-DTNN model, there still exists considerable scope for improvements in the QM-DTNN model for 20 amino acids. In addition to the energy of amino acid units, long-range interactions will also be included. The properties of the entire protein are determined.

Although the QM-DTNN is a new approach to calculating the molecular energy in QM level precision, considerable scope still exists for its improvement. First, the QM-DTNN cannot handle proteins with non-standard amino acid, which can be addressed by training another DTNN of a non-standard amino acid unit training set. Second, QM-DTNN in this version can only handle standard amino acid units with same atom numbers. Because there exist situations in which the same amino acid units have different number of atoms, future versions of QM-DTNN will include the relative conformations in the training set, and the DTNN module will be retrained. Furthermore, to completely cover the conformational space (in addition to the CV space), making a good sampling is necessary. The quantitative accuracy achieved by DTNN and its size extensibility paves the way for calculating configurational energy differences. With high-QM precision, a considerable number of QM calculations can be implemented. This may be addressed by other deep learning tools such as GAN [14, 15] or autoencode [4, 13, 20].

References

[1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, *TensorFlow: large-scale machine learning on heterogeneous systems* (2015). Software available from `tensorflow.org`.

[2] M. Abadi, M. Isard, and D. G. Murray, *A computational model for TensorFlow: an introduction*, in: Proceedings of the 1st ACM SIGPLAN International Workshop on Machine Learning and Programming Languages, MAPL 2017, 1–7, ACM, New York, NY, USA (2017). `ISBN 978-1-4503-5071-6`.

[3] A. Abdi, S. M. Shamsuddin, S. Hasan, and J. Piran, *Machine learning-based multi-documents sentiment-oriented summarization using linguistic treatment*, Expert Systems with Applications **109** (2018), 66–85.

[4] C. Angermueller, T. Pärnamaa, L. Parts, and O. Stegle, *Deep learning for computational biology*, Molecular Systems Biology **12** (2016), no. 7, 878.

[5] F. Baftizadeh, P. Cossio, F. Pietrucci, and A. Laio, *Protein folding and ligand-enzyme binding from bias-exchange metadynamics simulations*, Current Physical Chemistry **2** (2012), no. 1, 79–91.

[6] A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi, and M. Ceriotti, *Machine learning unifies the modeling of materials and molecules*, Science Advances **3** (2017), no. 12.

[7] J. Behler, *Constructing high-dimensional neural network potentials: A tutorial review*, International Journal of Quantum Chemistry **115** (2015), no. 16, 1032–1050.

[8] J. Behler, *First principles neural network potentials for reactive simulations of large molecular and condensed systems*, Angewandte Chemie International Edition (2017).

[9] J. Behler, S. Lorenz, and K. Reuter, *Representing molecule-surface interactions with symmetry-adapted neural networks*, The Journal of Chemical Physics **127** (2007), no. 1, 014705.

[10] B. E. Bejnordi, M. Mullooly, R. M. Pfeiffer, S. Fan, P. M. Vacek, D. L. Weaver, S. Herschorn, L. A. Brinton, B. van Ginneken, N. Karssemeijer, A. H. Beck, G. L. Gierach, J. A. W. M. van der Laak, and M. E. Sherman, *Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies*, Modern Pathology (2018).

[11] N. J. Browning, R. Ramakrishnan, O. A. von Lilienfeld, and U. Roethlisberger, *Genetic optimization of training sets for improved machine learning models of molecular properties*, The Journal of Physical Chemistry Letters **8** (2017), no. 7, 1351–1359.

[12] D. Bychkov, N. Linder, R. Turkki, S. Nordling, P. E. Kovanen, C. Verrill, M. Walliander, M. Lundin, C. Haglund, and J. Lundin, *Deep learning based tissue analysis predicts outcome in colorectal cancer*, Scientific Reports **8** (2018), no. 1.

[13] W. Chen and A. L. Ferguson, *Molecular enhanced sampling with autoencoders: On-the-fly collective variable discovery and accelerated free energy landscape exploration* (2017).

[14] T. Ching, D. S. Himmelstein, B. K. Beaulieu-jones, A. Kalinin, B. T. Do, G. P. Way, E. Ferrero, W. Xie, G. L. Rosen, B. J. Lengerich, and J. Israeli, *Opportunities and obstacles for deep learning in biology and medicine* (2017).

[15] A. P. N. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A. A. Bharath, and M. Ieee, *Generative adversarial networks: an overview* (2017), no. April, 1–12.

[16] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. T. a, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ã. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, *Gaussian 09 Revision A.1*, Gaussian Inc. Wallingford CT 2009.

[17] J. Gao, D. G. Truhlar, Y. Wang, M. J. M. Mazack, P. Löffler, M. R. Provorse, and P. Rehak, *Explicit polarization: a quantum mechanical framework for developing next generation force fields*, Accounts of Chemical Research **47** (2014), no. 9, 2837–2845. PMID:25098651.

[18] J. Gelpi, A. Hospital, R. Goñi, and M. Orozco, *Molecular dynamics simulations: advances and applications*, Advances and Applications in Bioinformatics and Chemistry (2015), 37.

[19] N. Goga, S. Marrink, R. Cioromela, and F. Moldoveanu, *GPU-SD and DPD parallelization for Gromacs tools for molecular dynamics simulations*, in: 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE), IEEE (2012).

[20] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik, *Automatic chemical design using a data-driven continuous representation of molecules* (2016).

[21] L. Gong, *Development and applications of the ABEEM fluctuating charge molecular force field in the ion-containing systems*, Science China Chemistry **55** (2012), no. 12, 2471–2484.

[22] Y. González, P. Ezzatti, and M. Paulino, *Unleashing the graphic processing units-based version of NAMD*, in: Bioinformatics and Biomedical Engineering, 639–650, Springer International Publishing (2016).

[23] S. Grimme and P. R. Schreiner, *Computational chemistry: the fate of current methods and future challenges*, Angewandte Chemie – International Edition **57** (2018), no. 16, 4170–4176.

[24] K. Gubaev, E. V. Podryabinkin, and A. V. Shapeev, *Machine learning of molecular properties: locality and active learning* (2017), 1–13.

[25] X. Guo, Z. Qu, and J. Gao, *The charger transfer electronic coupling in diabatic perspective: A multi-state density functional theory study*, Chemical Physics Letters **691** (2018), 91–97.

[26] C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, *Deep learning for real-time single-pixel video*, Scientific Reports **8** (2018), no. 1.

[27] C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, *Deep learning for real-time single-pixel video*, Scientific Reports **8** (2018), no. 1.

[28] B. K. Holtzman, A. Paté, J. Paisley, F. Waldhauser, and D. Repetto, *Machine learning reveals cyclic changes in seismic source spectra in Geysers geothermal field*, Science Advances **4** (2018), no. 5.

[29] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, *Comparison of multiple Amber force fields and development of improved protein backbone parameters*, Proteins: Structure, Function, and Bioinformatics **65** (2006), no. 3, 712–725.

[30] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. W. L. Aerts, *Artificial intelligence in radiology*, Nature Reviews Cancer **18** (2018), no. 8, 500–510.

[31] M. R. Hush, *Machine learning for quantum physics*, Science **355** (2017), no. 6325, 580–580.

[32] Y. Ji, L. Liu, H. Wang, Z. Liu, Z. Niu, and B. Denby, *Updating the Silent Speech Challenge benchmark with deep learning*, Speech Communication **98** (2018), 42–50.

[33] M. I. Jordan and T. M. Mitchell, *Machine learning: trends, perspectives, and prospects*, Science **349** (2015), no. 6245, 255–260.

[34] A. Kamath, R. A. Vargas-Hernández, R. V. Krems, T. Carrington, and S. Manzhos, *Neural networks vs Gaussian process regression for representing potential energy surfaces: a comparative study of fit quality and vibrational spectrum accuracy*, The Journal of Chemical Physics **148** (2018), no. 24, 241702.

[35] A. Kan, *Machine learning applications in cell image analysis*, Immunology and Cell Biology **95** (2017), no. 6, 525–530.

[36] M. V. Khan, S. M. Zakariya, and R. H. Khan, *Protein folding, misfolding and aggregation: A tale of constructive to destructive assembly*, International Journal of Biological Macromolecules **112** (2018), 217–229.

[37] C. Kutzner, S. Páll, M. Fechner, A. Esztermann, B. L. de Groot, and H. Grubmüller, *Best bang for your buck: GPU nodes for GROMACS biomolecular simulations*, Journal of Computational Chemistry **36** (2015), no. 26, 1990–2008.

[38] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature **521** (2015), no. 7553, 436–444.

[39] T. Lemke and C. Peter, *Neural network based prediction of conformational free energies — a new route toward Coarse-Grained simulation models* (2017).

[40] H. Li, J. Chowdhary, L. Huang, X. He, A. D. MacKerell, and B. Roux, *Drude Polarizable Force Field for Molecular Dynamics Simulations of Saturated and Unsaturated Zwitterionic Lipids*, Journal of Chemical Theory and Computation **13** (2017), no. 9, 4535–4552. PMID:28731702.

[41] S. Lorenz, M. Scheffler, and A. Gross, *Descriptions of surface chemical reactions using a neural network representation of the potential-energy surface*, Physical Review B **73** (2006), no. 11.

[42] A. Madani, R. Arnaout, M. Mofrad, and R. Arnaout, *Fast and accurate view classification of echocardiograms using deep learning*, npj Digital Medicine **1** (2018), no. 1.

[43] M. J. M. Mazack, A. Cembran, and J. Gao, *Internal dynamics of an analytically coarse-grained protein*, Journal of Chemical Theory and Computation **6** (2010), no. 11, 3601–3612. PMID:21243085.

[44] D. J. Mermelstein, C. Lin, G. Nelson, R. Kretsch, J. A. McCammon, and R. C. Walker, *Fast and flexible gpu accelerated binding free energy*

calculations within the amber molecular dynamics package, Journal of Computational Chemistry **39** (2018), no. 19, 1354–1358.

[45] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K. R. Müller, and O. Anatole Von Lilienfeld, *Machine learning of molecular electronic properties in chemical compound space*, New Journal of Physics **15** (2013).

[46] A. Peramo, *Solvated and generalised Born calculations differences using GPU CUDA and multi-CPU simulations of an antifreeze protein with AMBER*, Molecular Simulation **42** (2016), no. 15, 1263–1273.

[47] U. Perricone, M. R. Gulotta, J. Lombino, B. Parrino, S. Cascioferro, P. Diana, G. Cirrincione, and A. Padova, *An overview of recent molecular dynamics applications as medicinal chemistry tools for the undruggable site challenge*, MedChemComm **9** (2018), no. 6, 920–936.

[48] J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, *Nanophotonic particle simulation and inverse design using artificial neural networks*, Science Advances **4** (2018), no. 6.

[49] J. Phillips, J. Stone, and K. Schulten, *Adapting a message-driven parallel application to GPU-accelerated clusters*, in: 2008 SC — International Conference for High Performance Computing, Networking, Storage and Analysis, IEEE (2008).

[50] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, and T. Head-Gordon, *Current status of the AMOEBA polarizable force field*, The Journal of Physical Chemistry B **114** (2010), no. 8, 2549–2564.

[51] S. Purawat, P. U. Ieong, R. D. Malmstrom, G. J. Chan, A. K. Yeung, R. C. Walker, I. Altintas, and R. E. Amaro, *A Kepler workflow tool for reproducible AMBER GPU molecular dynamics*, Biophysical Journal **112** (2017), no. 12, 2469–2474.

[52] L. A. G. F. L. M. C. P. M. Raiteri, P., *Efficient reconstruction of complex free energy landscapes b y multiple walkers metadynamics*, J. Phys. Chem. B **110** (2006), no. 1, 3533–3539.

[53] R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, *Quantum chemistry structures and properties of 134 kilo molecules*, Scientific Data **1** (2014), 1–7.

[54] P. W. Rose, A. Prlić, A. Altunkaya, C. Bi, A. R. Bradley, C. H. Christie, L. D. Costanzo, J. M. Duarte, S. Dutta, Z. Feng, R. K. Green, D. S. Goodsell, B. Hudson, T. Kalro, R. Lowe, E. Peisach, C. Randle, A. S. Rose, C. Shao, Y.-P. Tao, Y. Valasatava, M. Voigt, J. D. Westbrook, J. Woo, H. Yang, J. Y. Young, C. Zardecki, H. M. Berman, and S. K. Burley, *The RCSB protein data bank: integrative view of protein, gene and 3D structural information*, Nucleic Acids Research **45** (2017), no. D1, D271–D281.

[55] T. Sandalova, G. Schneider, H. Kaeck, and Y. Lindqvist, *Structure of APO-dethiobiotin synthase at* 0.97 *angstroms resolution* (1999).

[56] K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, *Quantum-chemical insights from deep tensor neural networks* **8** (2017), 13890.

[57] K. T. Schütt, P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller, *SchNet: A continuous-filter convolutional neural network for modeling quantum interactions* (2017), no. June.

[58] M. H. S. Segler, M. Preuss, and M. P. Waller, *Planning chemical syntheses with deep neural networks and symbolic AI*, Nature **555** (2018), no. 7698, 604–610.

[59] I. Serrano, O. Deniz, J. L. Espinosa-Aranda, and G. Bueno, *Fight recognition in video using hough forests and 2D convolutional neural network*, IEEE Transactions on Image Processing **27** (2018), no. 10, 4787–4797.

[60] D. E. Shaw, J. C. Chao, M. P. Eastwood, J. Gagliardo, J. P. Grossman, C. R. Ho, D. J. Lerardi, I. Kolossváry, J. L. Klepeis, T. Layman, C. McLeavey, M. M. Deneroff, M. A. Moraes, R. Mueller, E. C. Priest, Y. Shan, J. Spengler, M. Theobald, B. Towles, S. C. Wang, R. O. Dror, J. S. Kuskin, R. H. Larson, J. K. Salmon, C. Young, B. Batson, and K. J. Bowers, *Anton, a special-purpose machine for molecular dynamics simulation*, Communications of the ACM **51** (2008), no. 7, 91.

[61] D. E. Shaw, J. Grossman, J. A. Bank, B. Batson, J. A. Butts, J. C. Chao, M. M. Deneroff, R. O. Dror, A. Even, C. H. Fenton, A. Forte, J. Gagliardo, G. Gill, B. Greskamp, C. R. Ho, D. J. Ierardi, L. Iserovich, J. S. Kuskin, R. H. Larson, T. Layman, L.-S. Lee, A. K. Lerer, C. Li, D. Killebrew, K. M. Mackenzie, S. Y.-H. Mok, M. A. Moraes, R. Mueller, L. J. Nociolo, J. L. Peticolas, T. Quan, D. Ramot, J. K. Salmon, D. P. Scarpazza, U. B. Schafer, N. Siddique, C. W. Snyder,

J. Spengler, P. T. P. Tang, M. Theobald, H. Toma, B. Towles, B. Vitale, S. C. Wang, and C. Young, *Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer*, in: SC14: International Conference for High Performance Computing, Networking, Storage and Analysis, IEEE (2014).

[62] L. Shen, J. Wu, and W. Yang, *Multiscale quantum mechanics/molecular mechanics simulations with neural networks*, Journal of Chemical Theory and Computation **12** (2016), no. 10, 4934–4946.

[63] J. E. Stone, A.-P. Hynninen, J. C. Phillips, and K. Schulten, *Early experiences porting the NAMD and VMD molecular simulation and analysis software to GPU-accelerated OpenPOWER platforms*, in: Lecture Notes in Computer Science, 188–206, Springer International Publishing (2016).

[64] G. W. Taylor and G. E. Hinton, *Factored conditional restricted Boltzmann machines for modeling motion style*, in: Proceedings of the 26th Annual International Conference on Machine Learning, ICML '09, 1025–1032, ACM, New York, NY, USA (2009). ISBN 978-1-60558-516-1.

[65] F.-F. Wang, D.-X. Zhao, and L.-D. Gong, *Ab initio and ABEEM/MM fluctuating charge model studies of dimethyl phosphate anion in a microhydrated environment*, Theoretical Chemistry Accounts **124** (2009), no. 1, 139–150.

[66] J. Wang, H. Cao, J. Z. H. Zhang, and Y. Qi, *Computational protein design with deep learning neural networks*, Scientific Reports **8** (2018), no. 1.

[67] J. Wu, L. Shen, and W. Yang, *Internal force corrections with machine learning for quantum mechanics/molecular mechanics simulations*, Journal of Chemical Physics **147** (2017), no. 16.

[68] J. C. Wu, G. Chattree, and P. Ren, *Automation of AMOEBA polarizable force field parameterization for small molecules*, Theoretical Chemistry Accounts **131** (2012), no. 3.

[69] Y. Xu, Y. Han, R. Hong, and Q. Tian, *Sequential video VLAD: training the aggregation locally and temporally*, IEEE Transactions on Image Processing **27** (2018), no. 10, 4933–4944.

[70] C. Zhang, C. Lu, Z. Jing, C. Wu, J.-P. Piquemal, J. W. Ponder, and P. Ren, *AMOEBA polarizable atomic multipole force field for nucleic*

acids, Journal of Chemical Theory and Computation **14** (2018), no. 4, 2084–2108.

[71] D. W. Zhang, X. H. Chen, and J. Z. H. Zhang, *Molecular caps for full quantum mechanical computation of peptide-water interaction energy*, Journal of Computational Chemistry **24** (2003), no. 15, 1846–1852.

Yan Li and Guohui Li:
State Key Laboratory of Molecular Reaction Dynamics
Dalian Institute of Chemical Physics, Chinese Academy of Sciences
Dalian, Liaoning 116023, China
E-mail address: liyan1982@dicp.ac.cn
E-mail address: ghli@dicp.ac.cn

Hanyi Min:
Chinese Academy of Medical Science
and Peking Union Medical College Hospital, Ophthalmology
Beijing, China
E-mail address: minhy@pumch.cn

Zibing Dong and Tian Yuan:
School of Physics and Electronic Technology
Liaoning Normal University
Dalian, Liaoning 116023, China
E-mail address: 773877191@qq.com
E-mail address: 1002955308@qq.com

Xiaoqi Li and Peijun Xu:
School of Physics and Electronic Technology
Liaoning Normal University
Dalian, Liaoning 116023, China
E-mail address: 1239993859@qq.com
E-mail address: xpj631114@sohu.com