



# GPU Accelerated Quantum Virtual Screening: Application for the Natural Inhibitors of New Delhi Metalloprotein (NDM-1)

Mingsong Shi<sup>1</sup>, Dingguo Xu<sup>1\*</sup> and Jun Zeng<sup>2\*</sup>

<sup>1</sup> College of Chemistry, Sichuan University, Chengdu, China, <sup>2</sup> MedChemSoft Solutions, Wheelers Hill, VIC, Australia

## OPEN ACCESS

### Edited by:

Yong Wang,  
Ningbo University, China

### Reviewed by:

Qing-Chuan Zheng,  
Jilin University, China  
Fei Xia,  
East China Normal University, China  
Jun Gao,  
Huazhong Agricultural University,  
China

### \*Correspondence:

Dingguo Xu  
dgxu@scu.edu.cn  
Jun Zeng  
jun.zeng@medchemsoft.com

### Specialty section:

This article was submitted to  
Theoretical and Computational  
Chemistry,  
a section of the journal  
Frontiers in Chemistry

**Received:** 09 October 2018

**Accepted:** 31 October 2018

**Published:** 20 November 2018

### Citation:

Shi M, Xu D and Zeng J (2018) GPU Accelerated Quantum Virtual Screening: Application for the Natural Inhibitors of New Delhi Metalloprotein (NDM-1). *Front. Chem.* 6:564. doi: 10.3389/fchem.2018.00564

Quantum mechanical approaches for the massive computation on large biological system such as virtual screening in drug design and development have presented a challenge to computational chemists for many years. In this study, we demonstrated that by taking advantage of rapid growth of GPU-based hardware and software (i.e., *teraChem*), it is feasible to perform virtual screening of a refined chemical library at quantum mechanical level in order to identify the lead compounds with improved accuracy, especially for the drug targets such as metalloproteins in which significant charge transfer and polarization occur amongst the metal ions and their coordinated amino acids. Our calculations predicted four nature compounds (i.e., Curcumin, Catechin, menthol, and Ferulic acid) as the suitable inhibitors for antibiotics resistance against New Delhi Metallo- $\beta$ -lactamase-1 (NDM-1). Molecular orbitals (MOs) of the QM region of metal ions and their coordinated residues indicate that the bridged hydroxide ion delocalized the electron over the Zn-OH-Zn group at HOMO, different from MOs when the OH<sup>-</sup> is not presented in NDM-1. This indicates that the bridged hydroxide ion plays an important role in the design of antibiotics and other inhibitors targeting the metalloproteins.

**Keywords:** virtual screening, QM/MM, GPU, NDM-1, natural inhibitors

## INTRODUCTION

The drug discovery and development is an expensive and complex enterprise. The role of the computational chemist within this team has become firmly established as the computer-aided drug design impacts the development across the discovery time line, from target identification and hit discovery to hit-to-lead optimization and safety profiling (Loughney et al., 2011).

A common activity of the computational drug design is the prediction of the bound conformation of a ligand (“Docking”) or millions of ligands to a given receptor binding pocket (“virtual screening”). In general, widely used knowledge-based, empirical or force-field based scoring functions are capable of providing docked poses with reasonable accuracy. Along this line, development of virtual screening has been made in the areas of ligand-based methods including the field based methods (Jilek and Cramer, 2004), substructural descriptors (Vogt et al., 2010) etc., and of structure-based methods with many docking softwares available (see Sukumar and Das, 2011 and its references therein). More recently, the artificial intelligence (AI) and machine learning technologies have also been developed for virtual screening in drug discovery (Goh et al., 2017).

However, these approaches are usually not applicable for the complicated pockets such as those with metal ions or highly polarizable sites (Wang et al., 2011). In these cases, the incorporation of the quantum mechanical calculations into the docking or virtual screening process (Illingworth et al., 2008) becomes essential for obtaining the reasonably accurate prediction which serves as the stepping-stone for further hit-to-lead optimization.

Over the last two decades, there has been a dramatic change in the computational power and accessibility of quantum mechanical software programs, enabling the application of quantum mechanical (QM) scheme to ever larger and more complex systems. It remains however that the highest level of QM can only be applied to modest system sizes of 10s of atoms. A number of general strategies have been developed and employed to further extend the size of systems for drug-protein systems (Mucs and Bryce, 2013). The first is the QM/MM approach that describes a subregion of interest in electronic detail via QM and couples it to a larger environment modeled at the MM level. This approach has been widely used in study of enzyme activity (see review) (Senn and Thiel, 2009). The second approach is in the application of Fragment Molecular Orbital (FMO) (Fedorov et al., 2012). FMO scheme fragments the large molecules into small moieties as isolated monomers by using hybrid projection operators or adaptive frozen orbital. Non-additive terms can be recovered by considering combinations of dimers and trimers of the fragments. The third approach is the linear scaling methods (Zhou et al., 2010) which seek to reduce the scaling of computational cost order of  $N$  basis function. Several approaches have been developed by using localized molecular orbitals and divide-and-conquer method, as well as linear scaling wave plan technique. They have successfully applied to protein-ligand and protein-protein interactions at the QM level (Heady et al., 2006; Cole et al., 2010).

On the other hand, the rapid advances of computer power and technology brought up the importance of the utilization of computational architectures to accelerate the QM calculations. Distribution of the computational overheads of QM algorithms on parallel CPU-based architectures leads to tractable ab initio and density function QM calculations on thousands of basis functions (Valiev et al., 2010). Recently, work has focused on the adaption of QM algorithms to graphics processing units (GPUs) architectures. GPU-based approaches for DFT (Ufimtsev and Martinez, 2008, 2009a,b; Yasuda, 2008) and MP2 (Vogt et al., 2008) calculations have been developed and the software has also been advanced (Genovese et al., 2009; Ufimtsev and Martinez, 2009b; Asadchev and Gordon, 2012; Titov et al., 2013; Frisch et al., 2016).

This study emphasizes on the implementation of QM/MM approach into virtual screening for chemical libraries. By taking advantage of GPU-based architectures, we have developed a combined strategy to screen a library at QM level with a promising speed. To our knowledge, this is the first time to investigate the feasibility of GPU-based virtual screening of a chemical library at the QM/MM level.

We use New Delhi Metallo- $\beta$ -lactamase-1 (NDM-1) as the example. The NDM-1 protein bearing *Klebsiella pneumoniae*

and *Escherichia coli* infections, which have broad spectrum against all beta-lactams tested, were first discovered in India in 2008 (Yong et al., 2009). The occurrence of NDM-1 exaggerates the problem of bacteria resistance, and its rapid spreading also makes it become one of major threats to human health in recent years. However, no clinical effective inhibitors were reported so far. NDM-1 belongs to the family of  $\beta$ -lactamases, which can further be divided into A, B, C, and D subgroups. Basically, subfamily A, C, and D can be grouped into serine proteases, while subfamily B is a metalloprotein containing either one or two zinc ions to keep its activity. NDM-1 belongs to B1 family, in which two zinc ions have been identified with a hydroxide ion bridged between two zinc cofactors (Thomas et al., 2011). Based on the crystal structure NDM-1 complexed with captopril (King et al., 2012), the reaction mechanism of hydrolysis of ampicillin which leads to the drug resistance of bacteria has been studied both experimentally (Zhang and Hao, 2011) and theoretically (Zheng and Xu, 2013).

## METHODOLOGICAL DETAILS

### Model Construction

The structure of NDM-1 was taken from crystal structure of the complex between NDM-1 and hydrolysed ampicillin (PDB 3Q6X, Zhang and Hao, 2011). Previous QM/MM study has suggested a hydroxide ion connecting two zinc cofactors for the hydrolysis of antibiotics and therefore been included in our previous model (Thomas et al., 2011; Zheng and Xu, 2013). Overall, the zinc ions are coordinated with residues ( $\text{OH}^-$ , His120, His122, Asp123), and with residues ( $\text{OH}^-$ , His189, Cys208, His250), respectively. The position of  $\text{OH}^-$  group was taken from previous study (Zheng and Xu, 2013), while the Cys208 is deprotonated with net charge of  $-1$  as shown in **Figure 1**.

### Molecular Mechanical Virtual Screening

Our docking program was used for virtual screening of the natural product library to the NDM-1 protein. The docking program consists of two-steps: (1) performing a multiple-copy simultaneous search "MCSS" (Zeng and Treutlein, 1999; Zeng, 2000) of common fragments such as benzene, pyridine, pyrimidine, pyrazine and phenol on the binding sites of a protein. The details have been described else (Zeng and Treutlein, 1999). In this study, we use the residue Glu123 as the center of the binding site. (2) docking a compound into the binding site of the protein. The fragments of functional groups were used as the guidance to place the compound inside the binding site and a torsional Monte Carlo minimization was performed to optimize the ligand conformation. Details of docking algorithm have been briefly described before (Stehn et al., 2013; Treutlein et al., 2017). Our software package quCBit ([www.medchemsoft.com](http://www.medchemsoft.com)) has been developed and applied in this purpose. All-atom Charmm22 force field (MacKerell et al., 1998) was used and dielectric constant of 10 was used to screen the electrostatic interactions as implemented.

The docked conformations were selected based on its overlay criteria with the MCSS minima and the molecular mechanical interaction energies between the compound and protein. The overlay criterion is designated as following steps. Firstly, the sub-groups correspond to the fragment types within a compound. Secondly, the center of the sub-group of each docked conformation to the center of each MCSS minimum was calculated. If the distance is less than 2.0 Å, the sub-group is considered as overlay with MCSS minima. Only the docked conformations with all the sub-groups overlay were retained. By trial and error, interaction energy of 10.0 kcal/mol was used as threshold to select these conformations that will be subjected to further QM refinement.

## GPU-Based QM Refinement

A QM/MM scheme (Warshel and Levitt, 1976) was used for the QM refinement of all the docked conformations for each compound. In the scheme, the QM region contains residues

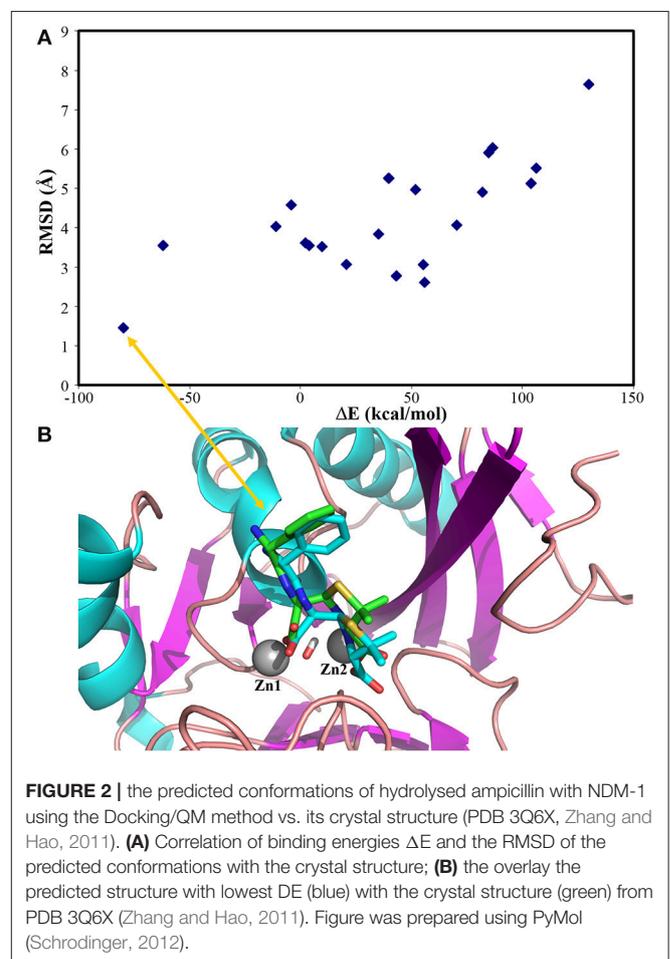
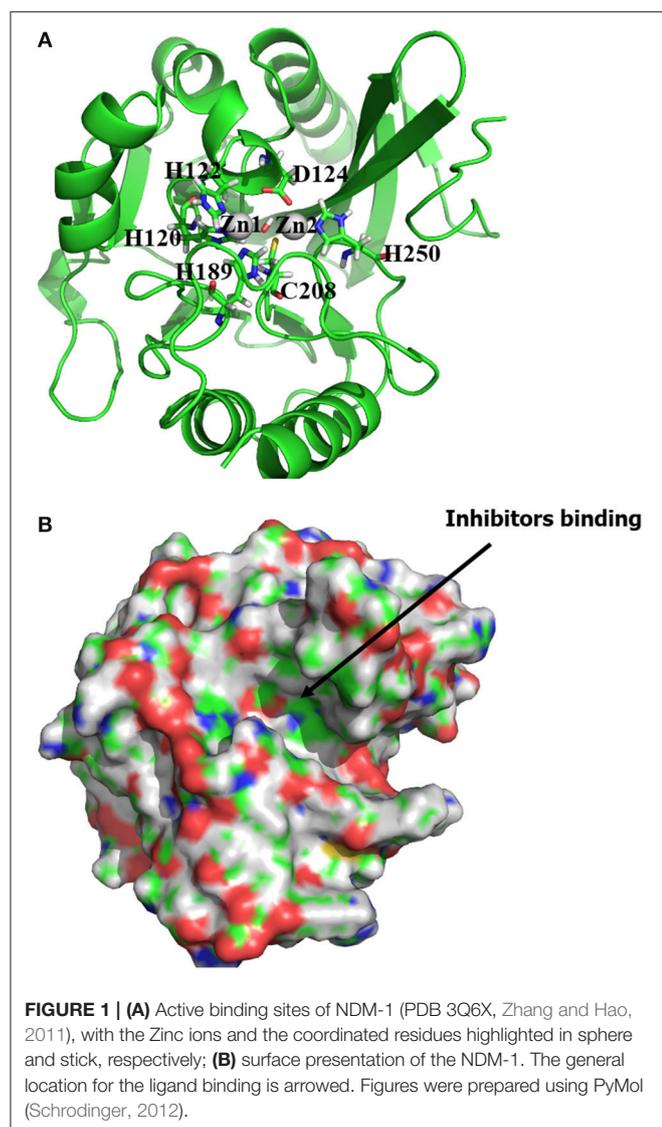
two zinc ions, hydroxide ion, 6 residues (i.e., His120, His122, Asp124, His189, Cys208, and His250), as well as the compound. The remaining part is the MM region. Saturated hydrogen was added for the boundary between QM and MM regions (Field et al., 1990). The total binding energies of protein-compound in solutions were calculated as

$$\Delta E = E_{\text{complex}}^{\text{PCM}} - E_{\text{protein}}^{\text{PCM}} - E_{\text{compound}}^{\text{PCM}} \quad (1)$$

where  $E_{\text{complex}}^{\text{PCM}}$  and  $E_{\text{protein}}^{\text{PCM}}$  should be calculated using QM/MM for the complex and protein with Polarizable continuum model (PCM) to mimic the solvent. However, current version of *teraChem* cannot perform this type of calculations so that we use gas-phase calculation for complex and protein. As the protein atoms are fixed in both complex and protein during the calculations, the errors between the results of gas-phase and PCM could be minor. Therefore, the Equation (1) thus becomes

$$\Delta E = E_{\text{complex}} - E_{\text{protein}} - E_{\text{compound}}^{\text{PCM}} \quad (2)$$

in which the  $E_{\text{complex}}$  and  $E_{\text{protein}}$  are the total energies of complex between protein and compound, and the total energy of the protein only, respectively.  $E_{\text{compound}}^{\text{PCM}}$  is the energy of



each compound in aqueous solution obtained using polarizable continuum model (PCM) (Barone et al., 1997) via Gaussian09 (Frisch et al., 2009) with dielectric constant of 80.0.

The structures of complex and NDM-1 protein were minimized using the combined QM/MM approach in which the QM is calculated using teraChem (Ufimtsev and Martinez, 2009b) and the MM region calculated using Amber12 (Case et al., 2012). Two hundred minimization steps were performed for each docked conformation obtained from section Molecular Mechanical Virtual Screening. During the minimization, DFT method with B3LYP functional (Lee et al., 1988; Becke, 1993) and 6-31G basis set (Ditchfield et al., 1971) was used for the QM region. We choose the standard basis set of 6-31G only as it significantly reduces the computational cost comparing with the 6-31G(d) with the minimum effect on

the energy and structures of the QM region. For the MM region, Amber ff14SB force field (Maier et al., 2015) was used for all amino acids, and general atom force field (GAFF) (Wang et al., 2004) was applied for the compounds. All the teraChem/Amber calculations were carried out on GPU computers (GTX780) with two E3-1200 CPUs and all the Gaussian09/Amber calculations were performed on two E5-2620 CPUs with 12 threads.

## RESULTS AND DISCUSSIONS

### Outcome of Virtual Screening

Our docking protocol was firstly used to predict the binding mode of hydrolysed ampicillin to the NDM-1, in comparison with crystal structure (PDB 3Q6X, Zhang and Hao, 2011).

**TABLE 1** | Calculated binding energies ( $\Delta E$ ) of nature compounds obtained using Equation (2).

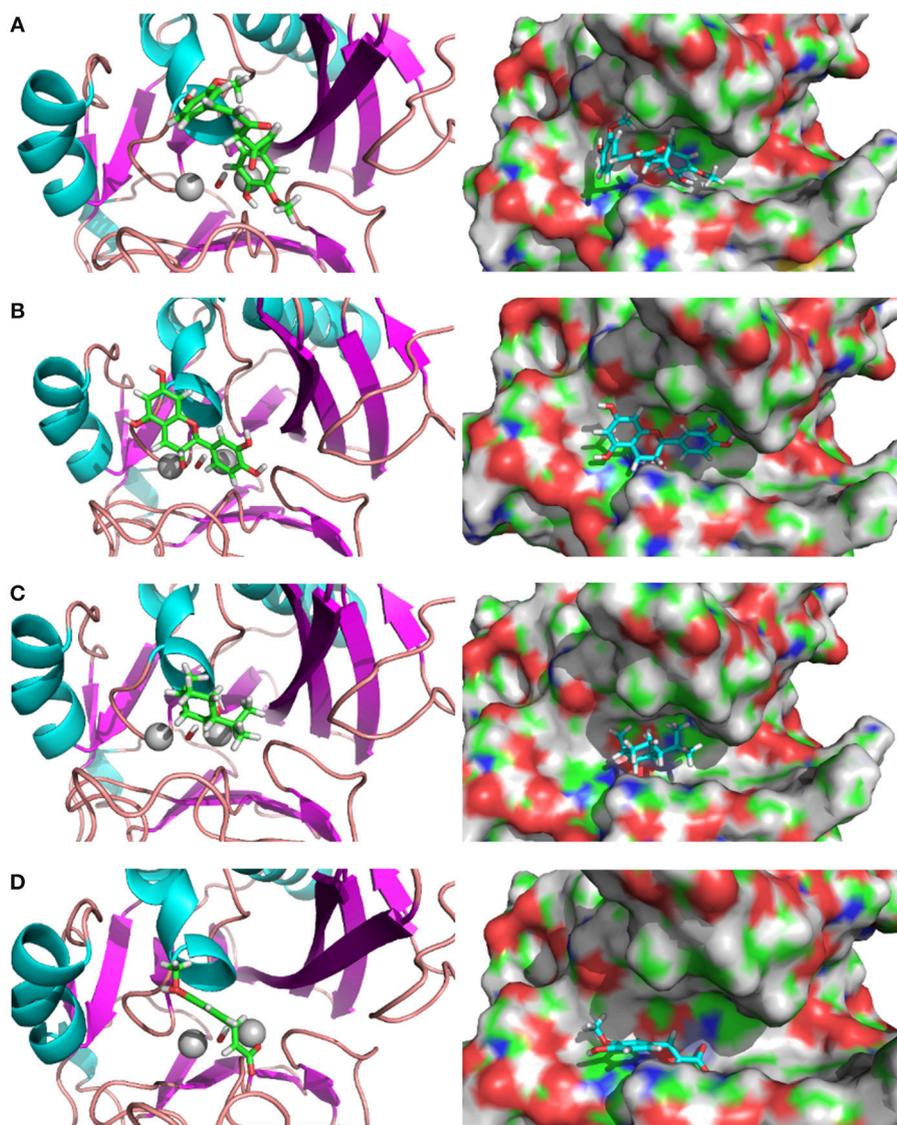
List	Molecules	Source	Formula	Number of atoms	Net charge	$\Delta E$ (kcal/mol)
1	<b>Curcumin (keto form)</b>	<b>Curcuma longa</b>	<b>C<sub>21</sub>H<sub>20</sub>O<sub>6</sub></b>	<b>47</b>	<b>0</b>	<b>-84.25</b>
2	Curcumin (Enol form)	<i>Curcuma longa</i>	C <sub>21</sub> H <sub>20</sub> O <sub>6</sub>	47	0	-14.74
3	Diosgenin	<i>Fenugreek</i>	C <sub>27</sub> H <sub>42</sub> O <sub>3</sub>	72	0	-43.90
4	Yamogenin	<i>Fenugreek</i>	C <sub>27</sub> H <sub>42</sub> O <sub>3</sub>	72	0	-50.18
5	Harmaline	<i>Peganum harmala</i>	C <sub>13</sub> H <sub>16</sub> N <sub>2</sub> O	32	0	-76.56
6	Harmine	<i>Peganum harmala</i>	C <sub>13</sub> H <sub>14</sub> N <sub>2</sub> O <sup>2+</sup>	30	+2	-11.20
7	Harmalol	<i>Peganum harmala</i>	C <sub>12</sub> H <sub>12</sub> N <sub>2</sub> O	27	0	-57.10
8	Tetrahydro-harmine	<i>Peganum harmala</i>	C <sub>13</sub> H <sub>15</sub> N <sub>2</sub> O <sup>+</sup>	31	+1	-71.07
9	Harmalan	<i>Peganum harmala</i>	C <sub>12</sub> H <sub>12</sub> N <sub>2</sub>	26	0	-72.04
10	Taraxerol	<i>Ciltorea Ternatea</i>	C <sub>30</sub> H <sub>50</sub> O	81	0	-
11	Rosmarinic acid	<i>Ocimum sanctum</i>	C <sub>18</sub> H <sub>15</sub> O <sub>8</sub> <sup>-</sup>	41	-1	-42.96
12	Carvacrol	<i>Essential oils of oregano etc.</i>	C <sub>10</sub> H <sub>14</sub> O	25	0	-51.08
13	Caryophyllene	<i>Essential oil of cloves etc</i>	C <sub>15</sub> H <sub>24</sub>	39	0	-
14	Zantrin, Z1		C <sub>20</sub> H <sub>15</sub> Cl <sub>3</sub> O <sub>3</sub>	41	0	-48.76
15	Nimbolide	<i>Azadriachta indica</i>	C <sub>27</sub> H <sub>30</sub> O <sub>7</sub>	44	0	-
16	Gedunin	<i>Azadriachta indica</i>	C <sub>28</sub> H <sub>34</sub> O <sub>7</sub>	69	0	-29.01
17	Mahmoodin	<i>Azadriachta indica</i>	C <sub>30</sub> H <sub>38</sub> O <sub>8</sub>	76	0	-6.50
18	Margolone	<i>Azadriachta indica</i>	C <sub>19</sub> H <sub>23</sub> O <sub>3</sub> <sup>-</sup>	45	-1	-56.88
19	Margonlonone	<i>Azadriachta indica</i>	C <sub>19</sub> H <sub>21</sub> O <sub>4</sub> <sup>-</sup>	44	-1	-25.60
20	Isomargolonone	<i>Azadriachta indica</i>	C <sub>19</sub> H <sub>21</sub> O <sub>4</sub> <sup>-</sup>	44	-1	-26.67
21	<b>Menthol</b>	<b><i>Mentha piperita</i></b>	<b>C<sub>10</sub>H<sub>20</sub>O</b>	<b>31</b>	<b>0</b>	<b>-81.08</b>
22	Campesterol	<i>Croton urucurana</i>	C <sub>28</sub> H <sub>48</sub> O	77	0	-66.43
23	<b>Catechin</b>	<b>Tea leaves</b>	<b>C<sub>15</sub>H<sub>14</sub>O<sub>6</sub></b>	<b>35</b>	<b>0</b>	<b>-83.76</b>
24	Sonderianin	<i>Croton urucurana</i>	C <sub>21</sub> H <sub>26</sub> O <sub>5</sub>	52	0	-4.18
25	Stigmasterol	<i>Various plants and herbs</i>	C <sub>29</sub> H <sub>48</sub> O	78	0	-75.81
26	Gallocatechin	<i>Tea leaves</i>	C <sub>15</sub> H <sub>14</sub> O <sub>7</sub>	36	0	-75.30
27	Acetyl aleuritolic acid	<i>Euphorbiaceae</i>	C <sub>32</sub> H <sub>49</sub> O <sub>4</sub> <sup>-</sup>	85	-1	-
28	<b>Ferulic acid</b>	<b><i>Scrophularia frutescense</i></b>	<b>C<sub>10</sub>H<sub>9</sub>O<sub>4</sub><sup>-</sup></b>	<b>23</b>	<b>-1</b>	<b>-79.12</b>
29	Isovanillic acid		C <sub>8</sub> H <sub>7</sub> O <sub>4</sub> <sup>-</sup>	19	-1	-34.79
30	Chrysin	<i>Tea leaves</i>	C <sub>15</sub> H <sub>10</sub> O <sub>4</sub>	29	0	-28.56
31	Carpaine	<i>Green plants and seeds</i>	C <sub>28</sub> H <sub>50</sub> N <sub>2</sub> O <sub>4</sub>	84	0	-
32	Berberine	<i>Barberry</i>	C <sub>20</sub> H <sub>18</sub> NO <sub>4</sub> <sup>+</sup>	43	1	-24.11
33	Harmane	<i>Various of foods</i>	C <sub>12</sub> H <sub>10</sub> N <sub>2</sub>	24	0	-35.97
34	Warfarin	<i>Rubiaceae</i>	C <sub>19</sub> H <sub>16</sub> O <sub>4</sub>	39	0	-72.69

The selected inhibitors are highlighted in bold.

At first, our MM docking calculations give 21 conformations based on the score of protein-ligand interaction energies less than 10.00 kcal/mol and overlay with the MCSS minima as described in Method section Molecular Mechanical Virtual Screening. Using the GPU-accelerated QM refinement, these conformations were further optimized in the complex with NDM-1. **Figure 2A** plots the correlation between binding energies of docked conformations after QM refinement and its heavy atom RMSD to the crystal structure. Overall, the best conformation (shown in **Figure 2B**) of hydrolysed ampicillin with NDM-1 have the lowest binding energies of  $-79.82$  kcal/mol with RMSD of  $1.42$  Å from the crystal structure. Based the QM protein-ligand binding energies  $\Delta E$ , our combined approach of MM docking and GPU-accelerated QM refinement

is capable to predict the native structure of ligand binding to the protein.

By applying our combined approach to screen the library of 34 nature compounds (Thakur et al., 2013), only 29 compounds were docked into the binding sites. **Table 1** list the best binding energies of these compounds with NDM-1. Using the hydrolysed ampicillin as the reference ( $E = -79.82$  kcal/mol), four compounds are considered as the potential inhibitors because they have value of  $E$  lower than the hydrolysed ampicillin. These hits include three compounds (Curcumin-Keto, Catechin, and Menthol) with strong binding energies of  $-84.25$  kcal/mol,  $-83.76$  kcal/mol, and  $-81.08$  kcal/mol, and one compound (Ferulic acid) with equivalent binding energy of  $-79.12$  kcal/mol, respectively.



**FIGURE 3** | Binding structures of potential inhibitors with the NDM-1 and its location on the surface of NDM-1. **(A)** Keto form of curcumin; **(B)** catechin; **(C)** menthol; **(D)** Ferulic acid. Figures were prepared using PyMol (Schrodinger, 2012).

The complex structures of these inhibitors to NDM-1 are shown in **Figure 3**. The keto form of Curcumin has the strongest binding energy to NDM1 ( $-84.25$  kcal/mol). The binding is dominated by interactions between the Zn-OH-Zn and the hydroxyl group of Curcumin. Similar feature is also found for Catechin and Menthol (**Figures 3B,C**). As shown in **Figure 3D**, Ferulic acid has the acidic group coordinated to Zinc ion directly; in addition, it also form electrostatic interactions to residue Lys393. This reminisces the binding structures of  $\beta$ -lactam antibiotics (e.g., ampicillin) to NDM-1 (**Figure 2**).

## MOs of Divalent Zinc Ions With Bridged Hydroxide Ion

The metalloproteins involved in bacteria infection have been a common target for the antibiotics. As a metalloprotein, NDM-1 can hydrolyse the antibiotics, and therefore, reduces or diminish its therapeutic effects. The hydrolysis reaction involves hydroxide ion between the two divalent Zinc ions (Thomas et al., 2011; Zheng and Xu, 2013). **Figure 4** shows the Molecular Orbitals (MOs) of the QM region of the NDM-1 system with/without hydroxide ion. The MOs are obtained from the Gaussian09 calculations. While the LUMO are similar between two systems with the electron clouds concentrated around Zn2 ion, the OH<sup>-</sup> bridge makes the HOMO delocalized over the Zn-OH-Zn cluster with a through-space  $\pi$ - $\pi$  interactions between OH<sup>-</sup> and S(-) of the deprotonated Cys208. Therefore, the inhibitors targeting on the NDM-1 will be of conjugate nature and interact Zn2 ion between deprotonated Cys208 and residue Gly211. This is consistent to the docked conformation predicted for the potential inhibitors of Curcumin-Keto, Catechin, Menthol, and Ferulic Acid as described above (**Figure 3**). In fact, the bridged hydroxide

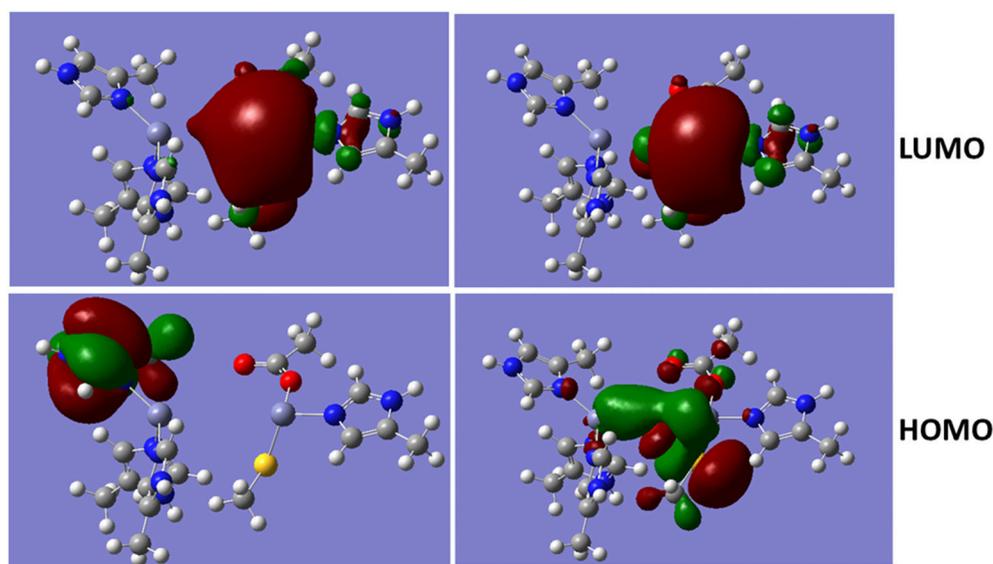
ions have been found in many metalloproteins, including oxygen transporters hemocyanin (Alzuet et al., 1997) and hemerythrin (Thomann et al., 1993), methane monooxygenase hydroxylase (Rosenzweig et al., 1993), and aminopeptidases (Bzymek et al., 2005) etc., and therefore, play critical role in the design of compounds that alter biological functions of metalloproteins.

## GPU vs. CPU Performance

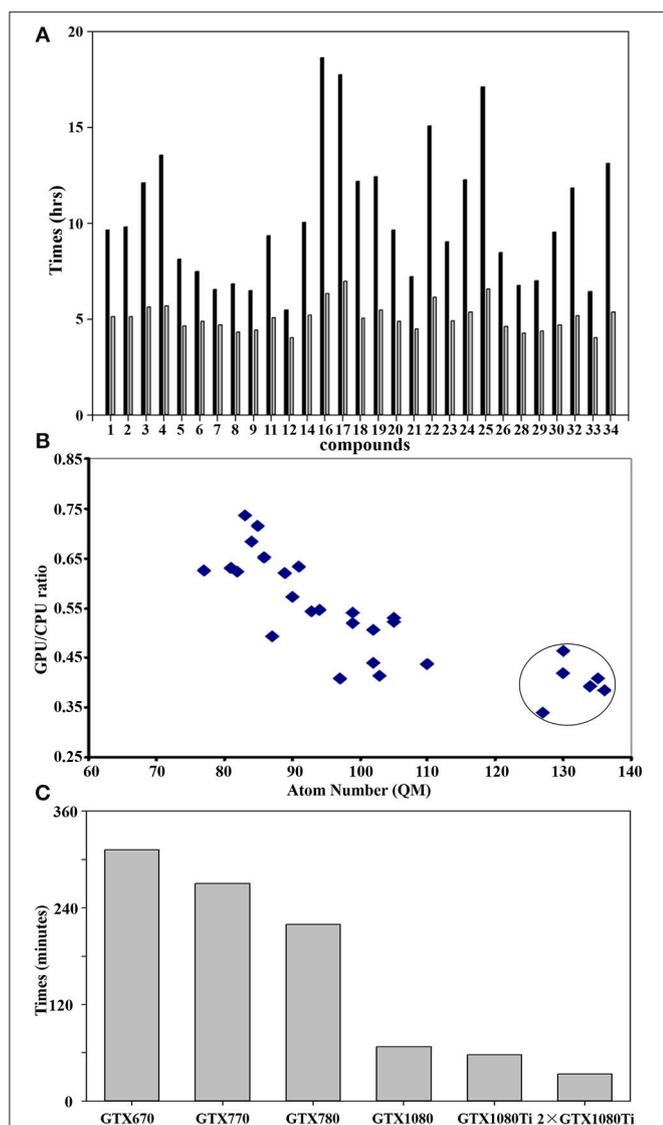
**Figure 5A** shows the difference of QM calculations of a single docked conformation from each nature compound using GPU and CPU (12 threads) via teraChem and Gaussian09, respectively. Overall, the GPU could significantly reduce the time cost up to 4-fold. The computational enhancement also depends on the size of molecules. As shown in **Figure 5B**, the significant reduction (up to 4-folds) of computational costs was achieved when the calculations were performed for the system with a QM region of more than 120 atoms.

GPU hardware also plays an important role in the performance of calculations. As shown in **Figure 5C**, GTX1080/1080ti speeded up the calculation up to 5-fold from GTX670. With two units of the most recent hardware GTX1080ti, maximal computer time cost for QM computation of a docked conformation can be reduced to less than 30 min, making the QM virtual screening of chemical library feasible.

It is noticeable that the computational time cost using GPU are consistent for each compound (**Figure 5A**), while the cost using CPU varies significantly according to its size. This is because the difference of the algorithm used for the computation of integral matrix between the teraChem and others. In order to make the



**FIGURE 4** | Molecular Orbital presentations (LUMO and HOMO) of QM region obtained using Gaussian09. Left panel is NDM-1 without bridged hydroxide ion (PDB 4EXS, King et al., 2012) and Right Panel is NDM-1 in which a hydroxide ion is located between two Zinc ions (PDB 3Q6X, Zhang and Hao, 2011). Figures were prepared using Gaussview (Dennington et al., 2016).



**FIGURE 5 |** GPU accelerated performance of quantum mechanical calculations. **(A)** Comparison of performance of QM/MM optimization between CPU and GPU using Gaussian09 and teraChem for the best docked conformation of the compounds as listed in **Table 1**. The x-axis sequence corresponds to the compound list in **Table 1**; **(B)** Correlation of GPU/CPU ratio (speedups) of the best docked conformation for each compound as function of atom numbers in the QM region. The speedups for the large size of QM region (> 120 atoms) are circled. **(C)** Computer time (in minutes) required for a single QM calculation using teraChem as function of the GPU unit model. Currently, the most advanced GTX is GTX 1080ti.

full utilization of GPU architecture, tetraChem has optimized the computations of the integrals onto GPU units (Ufimtsev and Martinez, 2009b), and reduced certain of computation into single precision (Luehr et al., 2011).

## CONCLUSIONS

QM method has been used for analysis of protein-ligand interaction in drug design and discovery for many years.

However, its expensive cost of computer power has hindered its application for virtual screening in drug design and development. Recent years, utilization of computational architecture to accelerate QM calculations has attracted significant attention (Furukawa et al., 2012; Ratcliff et al., 2017), and GPU-based approaches to perform QM calculations (Ufimtsev and Martinez, 2008) have been developed. Here, we have developed a combined strategy with docking the protein-ligand at MM level and re-rank the binding interaction at QM level. We choose the New Delhi metalloprotein 1 (NDM-1) as the target due to its important metal binding sites in drug resistance and its significant threat to the world health (Johnson and Woodford, 2013).

With the GPU-accelerated QM refinement, our calculations have demonstrated that it is essential to re-rank the docked conformations at the Quantum mechanical level in order to identify the native binding conformation of a ligand close to the crystal structure, as demonstrated from the calculations of the hydrolysed antibiotic ampicillin to NDM-1 (**Figure 2**).

Our calculations predict potential inhibitors different from the previous docking study in which Isomargolone and Nimbolide were estimated to be the inhibitors of NDM-1 (Thakur et al., 2013). The discrepancy could be due to the factors that they used the crystal structure of NDM-1 (PDB 3Q6X, Zhang and Hao, 2011) as target without the bridged hydroxide ion and employed empirical docking and scoring function (Goodsell et al., 1996; Morris et al., 1998) only to estimate the binding affinities of the compounds. From our QM calculations, MOs of the QM region shows that the HOMO with the bridged OH<sup>-</sup> between zinc ions has electron cloud delocalized over the Zn-OH-Zn cluster, while the LUMOs are similar with electron cloud concentrated at Zn<sub>2</sub>, independent of the hydroxide ion. As a result, the inhibitors of antibiotics resistance against the NDM-1 will be of conjugate nature, by binding over the Zn-OH-Zn, such as Curcumin, Catechin, and Ferulic acid (**Figure 3**).

The benchmark calculations presented here also provide some insight on the feasibility of virtual screening at quantum mechanical level. For a large QM region of more than 120 atoms which are very common for the metalloproteins, the GPU calculations can achieve the speedup of 4-fold comparing to the CPU method (**Figure 5A**). Previous study has demonstrated that 200-300 atoms would be needed as the minimum size of QM region in order to obtain accurate activation energies of enzyme reactions in proteins (Kulik et al., 2016). Moreover, the performance based on the GPU hardware infer that two units of the most advanced GPU architecture (GTX1080ti) could achieve the QM computation of single conformation within less than 30 min (**Figure 5C**). Furthermore, the speedup performance can be scaled up ca. twice by using two GPU processes, as demonstrated using two GTX1080ti units (**Figure 5C**).

Overall, our calculations demonstrated that it is feasible to use GPU-accelerated quantum mechanical method for virtual screening of chemical library. With the rapid advance of GPU hardware, quantum mechanical calculation in virtual screening will become more sophisticated and accurate at higher level. Here, we also showed that keto form of Curcumin, Catechin, Menthol, and Ferulic Acid may be considered as inhibitors for the antibiotics resistance against NDM-1. As natural product library

is significantly smaller library than the synthetic chemicals, the approach proposed here will be of significance for accurate prediction of natural product hits as the lead compounds for the drug development.

## AUTHOR CONTRIBUTIONS

MS, DX, and JZ contributed conception and design of the study. MS organized the database. MS and DX performed the statistical analysis. MS wrote the first draft of the manuscript. DX and JZ wrote sections of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

## REFERENCES

- Alzuet, G., Bubacco, L., Casella, L., Rocco, G. P., Salvato, B., and Beltrami, M. (1997). The binding of azide to copper-containing and cobalt-containing forms of hemocyanin from the Mediterranean crab *Carcinus aestuarii*. *Eur. J. Biochem.* 247, 688–694. doi: 10.1111/j.1432-1033.1997.00688.x
- Asadchev, A., and Gordon, M. S. (2012). New multithreaded hybrid CPU/GPU approach to Hartree-Fock. *J. Chem. Theory Comput.* 8, 4166–4176. doi: 10.1021/ct300526w
- Barone, V., Cossi, M., and Tomasi, J. (1997). A new definition of cavities for the computation of solvation free energies by the polarizable continuum model. *J. Chem. Phys.* 107, 3210–3221. doi: 10.1063/1.474671
- Becke, A. D. (1993). Density-functional thermochemistry. 3. The role of exact exchange. *J. Chem. Phys.* 98, 5648–5652. doi: 10.1063/1.464913
- Bzymek, K. P., Swierczek, S. I., Bennett, B., and Holz, R. C. (2005). Spectroscopic and thermodynamic characterization of the E151D and E151A altered leucine aminopeptidases from *Aeromonas proteolytica*. *Inorg. Chem.* 44, 8574–8580. doi: 10.1021/ic051034g
- Case, D. A., Darden, T., Cheatham, T. III, Simmerling, C., Wang, J., Duke, R. E., et al. (2012). *AMBER 12*. San Francisco, CA: University of California.
- Cole, D. J., Skylaris, C. K., Rajendra, E., Venkitaraman, A. R., and Payne, M. C. (2010). Protein-protein interactions from linear-scaling first-principles quantum-mechanical calculations. *Epl* 91:6. doi: 10.1209/0295-5075/91/37004
- Dennington, R., Keith, T. A., and Millam, J. M. (2016). *GaussView, Version 6*, Shawnee, KS: Semichem Inc.
- Ditchfield, R., Hehre, W. J., and Pople, J. A. (1971). Self-consistent molecular orbital methods. 9. Extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* 54:724. doi: 10.1063/1.1674902
- Fedorov, D. G., Nagata, T., and Kitaura, K. (2012). Exploring chemistry with the fragment molecular orbital method. *Phys. Chem. Chem. Phys.* 14, 7562–7577. doi: 10.1039/c2cp23784a
- Field, M. J., Bash, P. A., and Karplus, M. (1990). A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* 11, 700–733. doi: 10.1002/jcc.540110605
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2009). *Gaussian 09 Rev. A01*. Wallingford, CT.
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Gaussian 16 Rev. B.01*. Wallingford, CT.
- Furukawa, Y., Koga, R., and Yasuda, K. (2012). Acceleration of computational quantum chemistry by AVX plus CUDA on heterogeneous computer architectures. *Abstr. Papers Am. Chem. Soc.* 243:1.
- Genovese, L., Ospici, M., Deutsch, T., Mehaut, J. F., Neelov, A., and Goedecker, S. (2009). Density functional theory calculation on many-cores hybrid central processing unit-graphic processing unit architectures. *J. Chem. Phys.* 131:8. doi: 10.1063/1.3166140
- Goh, G. B., Hodas, N. O., and Vishnu, A. (2017). Deep learning for computational chemistry. *J. Comput. Chem.* 38, 1291–1307. doi: 10.1002/jcc.24764
- Goodsell, D. S., Morris, G. M., and Olson, A. J. (1996). Automated docking of flexible ligands: applications of AutoDock. *J. Mol. Recogn.* 9, 1–5. doi: 10.1002/(sici)1099-1352(199601)9:1<1::Aid-jmol241>3.0.Co;2-6
- Heady, L., Fernandez-Serra, M., Mancera, R. L., Joyce, S., Venkitaraman, A. R., Artacho, E., et al. (2006). Novel structural features of CDK inhibition revealed by an ab initio computational method combined with dynamic simulations. *J. Med. Chem.* 49, 5141–5153. doi: 10.1021/jm060190+
- Illingworth, C. J. R., Morris, G. M., Parkes, K. E. B., Snell, C. R., and Reynolds, C. A. (2008). Assessing the role of polarization in docking. *J. Phys. Chem. A* 112, 12157–12163. doi: 10.1021/jp710169m
- Jilek, R. J., and Cramer, R. D. (2004). Topomers: a validated protocol for their self-consistent generation. *J. Chem. Inf. Comput. Sci.* 44, 1221–1227. doi: 10.1021/ci049961d
- Johnson, A. P., and Woodford, N. (2013). Global spread of antibiotic resistance: the example of New Delhi metallo-beta-lactamase (NDM)-mediated carbapenem resistance. *J. Med. Microbiol.* 62, 499–513. doi: 10.1099/jmm.0.052555-0
- King, D. T., Worrall, L. J., Gruninger, R., and Strynadka, N. C. J. (2012). New Delhi Metallo-beta-lactamase: structural insights into beta-lactam recognition and inhibition. *J. Am. Chem. Soc.* 134, 11362–11365. doi: 10.1021/ja303579d
- Kulik, H. J., Zhang, J. Y., Klinman, J. P., and Martinez, T. J. (2016). How large should the QM region be in QM/MM calculations? The case of catechol O-methyltransferase. *J. Phys. Chem. B* 120, 11381–11394. doi: 10.1021/acs.jpcc.6b07814
- Lee, C. T., Yang, W. T., and Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37, 785–789. doi: 10.1103/PhysRevB.37.785
- Loughney, D., Claus, B. L., and Johnson, S. R. (2011). To measure is to know: an approach to CADD performance metrics. *Drug Discov. Today* 16, 548–554. doi: 10.1016/j.drudis.2011.05.003
- Luehr, N., Ufimtsev, I. S., and Martinez, T. J. (2011). Dynamic precision for electron repulsion integral evaluation on graphical processing units (GPUs). *J. Chem. Theory Comput.* 7, 949–954. doi: 10.1021/ct100701w
- MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102, 3586–3616. doi: 10.1021/jp973084f
- Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 11, 3696–3713. doi: 10.1021/acs.jctc.5b00255
- Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., et al. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* 19, 1639–1662. doi: 10.1002/(SICI)1096-987X(199811)19:14<1639::AID-JCC10>3.0.CO;2-B
- Mucs, D., and Bryce, R. A. (2013). The application of quantum mechanics in structure-based drug design. *Expert Opin. Drug Discov.* 8, 263–276. doi: 10.1517/17460441.2013.752812
- Ratcliff, L. E., Mohr, S., Huhs, G., Deutsch, T., Masella, M., and Genovese, L. (2017). Challenges in large scale quantum mechanical calculations. *Wiley Interdiscipl. Rev. Comput. Mol. Sci.* 7:24. doi: 10.1002/wcms.1290
- Rosenzweig, A. C., Frederick, C. A., Lippard, S. J., and Nordlund, P. (1993). Crystal structure of a bacterial non-haem iron hydroxylase that catalyses the biological oxidation of methane. *Nature* 366, 537–543. doi: 10.1038/366537a0

## FUNDING

This work was funded by the National Natural Science Foundation of China (No. 21473117).

## ACKNOWLEDGMENTS

Some of the results described in this work were obtained on the Supercomputing Center of Chinese Academy of Science and Supercomputing Center of Sichuan University. We also thank Prof. Todd Martinez of Stanford University, USA for sharing the software of teraChem.

- Schrodinger (2012). *PyMol: The PyMol Molecular Graphics System, Version 1.5.0.4*. Schrodinger LLC.
- Senn, H. M., and Thiel, W. (2009). QM/MM methods for biomolecular systems. *Angew. Chem. Int. Edn.* 48, 1198–1229. doi: 10.1002/anie.200802019
- Stehn, J. R., Haass, N. K., Bonello, T., Desouza, M., Kottyan, G., Treutlein, H., et al. (2013). A novel class of anticancer compounds targets the actin cytoskeleton in tumor cells. *Cancer Res.* 73, 5169–5182. doi: 10.1158/0008-5472.Can-12-4501
- Sukumar, N., and Das, S. (2011). Current trends in virtual high throughput screening using ligand-based and structure-based methods. *Comb. Chem. High Throughput Screen.* 14, 872–888. doi: 10.2174/138620711797537120
- Thakur, P. K., Kumar, J., Ray, D., Anjum, F., and Hassan, M. I. (2013). Search of potential inhibitor against New Delhi metallo-beta-lactamase 1 from a series of antibacterial natural compounds. *J. Nat. Sci. Biol. Med.* 4, 51–56. doi: 10.4103/0976-9668.107260
- Thomann, H., Bernardo, M., McCormick, J. M., Pulver, S., Andersson, K. K., Lipscomb, J. D., et al. (1993). Pulsed EPR studies of mixed valent [Fe(II)Fe(III)] forms of hemerythrin and methane monooxygenase: evidence for a hydroxide bridge. *J. Am. Chem. Soc.* 115, 8881–8882. doi: 10.1021/ja00072a068
- Thomas, P. W., Zheng, M., Wu, S. S., Guo, H., Liu, D. L., Xu, D. G., et al. (2011). Characterization of purified New Delhi Metallo-beta-lactamase-1. *Biochemistry* 50, 10102–10113. doi: 10.1021/bi201449r
- Titov, A. V., Ufimtsev, I. S., Luehr, N., and Martinez, T. J. (2013). Generating efficient quantum chemistry codes for novel architectures. *J. Chem. Theory Comput.* 9, 213–221. doi: 10.1021/ct300321a
- Treutlein, H. R., Styles, M., and Zeng, J. (2017). *Successful Computational Prediction of the Structure-Activity Relationship of a Potent JAK2 Inhibitor*. doi: 10.4225/03/5a1372ea64301
- Ufimtsev, I. S., and Martinez, T. J. (2008). Quantum chemistry on graphical processing units. 1. Strategies for two-electron integral evaluation. *J. Chem. Theory Comput.* 4, 222–231. doi: 10.1021/ct700268q
- Ufimtsev, I. S., and Martinez, T. J. (2009a). Quantum chemistry on graphical processing units. 2. Direct self-consistent-field implementation. *J. Chem. Theory Comput.* 5, 1004–1015. doi: 10.1021/ct800526s
- Ufimtsev, I. S., and Martinez, T. J. (2009b). Quantum chemistry on graphical processing units. 3. Analytical energy gradients, geometry optimization, and first principles molecular dynamics. *J. Chem. Theory Comput.* 5, 2619–2628. doi: 10.1021/ct9003004
- Valiev, M., Bylaska, E. J., Govind, N., Kowalski, K., Straatsma, T. P., Van Dam, H. J. J., et al. (2010). NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* 181, 1477–1489. doi: 10.1016/j.cpc.2010.04.018
- Vogt, L., Olivares-Amaya, R., Kermes, S., Shao, Y., Amador-Bedolla, C., and Aspuru-Guzik, A. (2008). Accelerating resolution-of-the-identity second-order Moller-Plesset quantum chemistry calculations with graphical processing units. *J. Phys. Chem. A* 112, 2049–2057. doi: 10.1021/jp0776762
- Vogt, M., Stumpfe, D., Geppert, H., and Bajorath, J. (2010). Scaffold hopping using two-dimensional fingerprints: true potential, black magic, or a hopeless endeavor? Guidelines for virtual screening. *J. Med. Chem.* 53, 5707–5715. doi: 10.1021/jm100492z
- Wang, J. C., Lin, J. H., Chen, C. M., Perryman, A. L., and Olson, A. J. (2011). Robust scoring functions for protein-ligand interactions with quantum chemical charge models. *J. Chem. Inf. Model.* 51, 2528–2537. doi: 10.1021/ci200220v
- Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and testing of a general amber force field. *J. Comput. Chem.* 25, 1157–1174. doi: 10.1002/jcc.20035
- Warshel, A., and Levitt, M. (1976). Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 103, 227–249. doi: 10.1016/0022-2836(76)90311-9
- Yasuda, K. (2008). Accelerating density functional calculations with graphics processing unit. *J. Chem. Theory Comput.* 4, 1230–1236. doi: 10.1021/ct8001046
- Yong, D., Toleman, M. A., Giske, C. G., Cho, H. S., Sundman, K., Lee, K., et al. (2009). Characterization of a New Metallo-beta-Lactamase gene, bla(NDM-1), and a novel erythromycin esterase gene carried on a unique genetic structure in *Klebsiella pneumoniae* sequence type 14 from India. *Antimicrob. Agents Chemother.* 53, 5046–5054. doi: 10.1128/aac.00774-09
- Zeng, J. (2000). Mini-review: computational structure-based design of inhibitors that target protein surfaces. *Comb. Chem. High Throughput Screen.* 3, 355–362. doi: 10.2174/1386207003331490
- Zeng, J., and Treutlein, H. R. (1999). A method for computational combinatorial peptide design of inhibitors of Ras protein. *Protein Eng.* 12, 457–468. doi: 10.1093/protein/12.6.457
- Zhang, H. M., and Hao, Q. (2011). Crystal structure of NDM-1 reveals a common beta-lactam hydrolysis mechanism. *Faseb J.* 25, 2574–2582. doi: 10.1096/fj.11-184036
- Zheng, M., and Xu, D. G. (2013). New Delhi Metallo-beta-lactamase I: substrate binding and catalytic mechanism. *J. Phys. Chem. B* 117, 11596–11607. doi: 10.1021/jp4065906
- Zhou, T., Huang, D. Z., and Caffisch, A. (2010). Quantum mechanical methods for drug design. *Curr. Top. Med. Chem.* 10, 33–45. doi: 10.2174/156802610790232242

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Shi, Xu and Zeng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.