

# Thermochemical and Quantum Descriptor Calculations for Gaining Insight into Ricin Toxin A (RTA) Inhibitors

Acassio Rocha-Santos, Elton José Ferreira Chaves, Igor Barden Grillo, Amanara Souza de Freitas, Demétrius Antônio Machado Araújo, and Gerd Bruno Rocha\*



Cite This: *ACS Omega* 2021, 6, 8764–8777



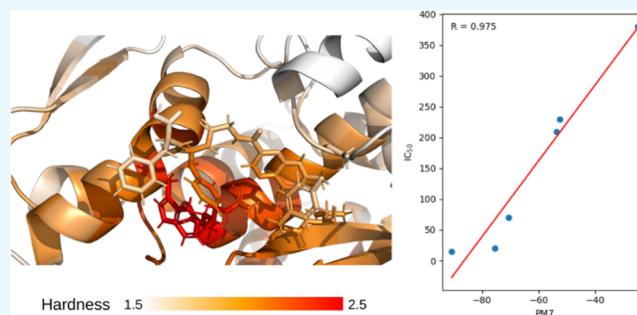
Read Online

ACCESS |

Metrics & More

Article Recommendations

**ABSTRACT:** In this work, we performed a study to assess the interactions between the ricin toxin A (RTA) subunit of ricin and some of its inhibitors using modern semiempirical quantum chemistry and ONIOM quantum mechanics/molecular mechanics (QM/MM) methods. Two approaches were followed (calculation of binding enthalpies,  $\Delta H_{\text{bind}}$ , and reactivity quantum chemical descriptors) and compared with the respective half-maximal inhibitory concentration ( $IC_{50}$ ) experimental data, to gain insight into RTA inhibitors and verify which quantum chemical method would better describe RTA–ligand interactions. The geometries for all RTA–ligand complexes were obtained after running classical molecular dynamics simulations in aqueous media. We found that single-point energy calculations of  $\Delta H_{\text{bind}}$  with the PM6-DH+, PM6-D3H4, and PM7 semiempirical methods and ONIOM QM/MM presented a good correlation with the  $IC_{50}$  data. We also observed, however, that the correlation decreased significantly when we calculated  $\Delta H_{\text{bind}}$  after full-atom geometry optimization with all semiempirical methods. Based on the results from reactivity descriptors calculations for the cases studied, we noted that both types of interactions, molecular overlap and electrostatic interactions, play significant roles in the overall affinity of these ligands for the RTA binding pocket.



## 1. INTRODUCTION

Ricin is a cytotoxic protein produced in the seed of the castor bean plant (*Ricinus communis*), and it belongs to the ribosome-inactivating protein family (type-2) and is one of the most potent biological toxins known.<sup>1</sup> This protein consists of two subunits joined together by a disulfide bond, namely, ricin toxin A (RTA) and ricin toxin B (RTB).<sup>2–4</sup> RTA (267 residues) is an N-glycosidase that inactivates eukaryotic ribosomes via depurination of a specific adenine located in the sarcin-ricin loop (SRL) motif of the 28S rRNA subunit.<sup>5</sup> RTB (262 residues) is a lectin that mediates the uptake of holoricin into cells via recognition of galactose and N-6-acetylgalactosamine.<sup>6</sup> Once internalized, holoricin undergoes vesicular retrograde transport until reaching the endoplasmic reticulum lumen;<sup>7</sup> then, an isomerase reduces the disulfide bond between the subunits, and RTA is translocated to the cytosol and subsequently efficiently attacks ribosomes.<sup>8,9</sup>

There is a global concern by world authorities regarding ricin toxicity due to its potential use as a chemical weapon, mainly by terrorist and activist groups. In addition, the absence of countermeasures to ricin poisoning further contributes to this concern. In this way, theoretical methods, such as molecular docking, have guided research groups to

find antagonist scaffolds for the RTA active site. However, this search has not been a trivial task because the active site of RTA and its surroundings are largely polar, imposing polarity constraints that in turn are not accounted for by the simplest theoretical methods.

To date, theoretical and computational chemistries have played a key role in studying biological and/or biochemical systems, providing proper direction toward drug discovery.<sup>10</sup> Knowledge of the intermolecular interactions of protein–ligand systems is an important feature for the development of new drugs.<sup>11</sup> In this way, some approaches, e.g., molecular docking, have been used to predict the bioactive pose of ligands in the active sites of biological targets with therapeutic interest;<sup>12</sup> however, it is quite ineffective to predict the relative binding affinity and rank ligands according to experimental data.<sup>10,13,14</sup>

Received: June 1, 2020

Accepted: December 30, 2020

Published: March 23, 2021



Currently, the estimation of the binding free energy ( $\Delta G_{\text{bind}}$ ) and binding enthalpy ( $\Delta H_{\text{bind}}$ ) of a ligand in a protein–ligand complex is a major challenge in computational chemistry.<sup>10</sup> To tackle this problem, several groups have been using computational chemistry methods to predict better energy profiles.<sup>10,11,15–33</sup> Among these approaches are those entirely based on classical force fields, such as MM-G(P)BSA,<sup>34,35</sup> LIE,<sup>36</sup> SMD,<sup>37,38</sup> and FEP,<sup>39</sup> as well as their versions involving hybrid potentials (quantum mechanics/molecular mechanics (QM/MM)) or those totally calculated by quantum mechanics (QM), such as QM-MM-G(P)-BSA,<sup>40,41</sup> QM-MM-LIE,<sup>23,42,43</sup> QM-FEP,<sup>25,44</sup> and QM-SMD.<sup>45,46</sup> All of these approaches require protein–ligand flexibility, either at their end-points or at the free-energy surface, resulting in high computational costs. Therefore, the use of such approaches is limited to a small set of ligands and small- and medium-sized complexes.

An alternative that requires lower computational cost is using theoretical methods for a single structure, in which the flexibility of the receptor (R), ligand (L), and receptor–ligand (R–L) complex can be partially accounted for by carrying out geometry optimization for R, L, and R–L. In this scenario, if the interest is to increase the accuracy of the receptor–ligand–binding energy predictions, the use of QM methods is mandatory. There are two approaches for calculating the binding energy of a ligand in a biological target using QM methods. The first approach considers a selection of the R–L complex atoms (QM cluster), and the second considers all atoms of the R–L complex in the QM calculation.

In the QM cluster strategy, a representative set of residues (ranging between 20 and 200 atoms) is selected and cut off from the active site. This set is studied separately by applying QM methods either in vacuum or in a continuous solvent model.<sup>47,48</sup> The advantage of this strategy is that few atoms are considered in the calculation, which allows us to assess a large set of ligands for the same active site. However, the disadvantage of this approach is that removing important residues during QM region selection can lead to inaccurate results.<sup>49,50</sup> In the strategy that considers all R–L atoms for QM calculations, it is possible to both hasten the computation of thermochemical properties and enable exploration of large ligand databases using linear scaling techniques coupled with graphics processing unit (GPU)-type accelerators.<sup>51,52</sup>

In the literature, some studies have performed molecular modeling and simulation involving ricin. Olson<sup>53</sup> applied molecular dynamics (MD) methods to analyze the structural and energetic aspects of three polynucleotide ligands (rRNA substrate analogues) that bound to the RTA active site. The results of the interaction energies showed that the overall binding is dominated by nonspecific interactions that occur in the region with a high degree of protein basicity through specific arginine contacts. The simulations of the three R–L complexes, as well as their comparison with experimental data, allowed a better understanding of the interaction of RTA with rRNA.

In another report, Olson and Cuff<sup>54</sup> expanded the previous study<sup>53</sup> by analyzing the free-energy determinants for the formation of RTA complexes with the rRNA substrate and several small ligands. The authors found that the absolute free energies of formation obtained for the RTA–RNA complex, as well as for several protein mutants, presented good agreement with the experimental data. In addition, it was observed that the terms of free energies presented unfavorable

electrostatic contributions that were balanced by the favorable nonspecific hydrophobic effect, with free energies similar to those of protein–protein complexes. The individual components (by amino acid residue) of the binding free energy of the RTA–RNA complex revealed highly relevant electrostatic interactions arising from the charge–charge complementarity of the interfacial arginines with the RNA phosphate backbone. In addition, it has been observed that the hydrophobic complementarity of the domain is exerted by the base interactions of the GAGA loop structure.

Yan and co-workers<sup>55</sup> conducted studies on the interactions of small rings with the RTA active site to better understand how ricin recognizes adenine rings. The geometries and interaction energies were calculated using MM methods for some complexes between the RTA active site and tautomeric modifications of adenine, formycin, guanine, and pterin. The results indicated that the interaction energies between the pterin ring and RTA are stronger than those of formycin with RTA. It has also been found that formycin binds more strongly to RTA than adenine. This information presented good agreement with the experimental data. In addition, the results of experimental and molecular modeling work suggest that the binding site of ricin is quite rigid and can recognize only a small range of adenine-like rings.

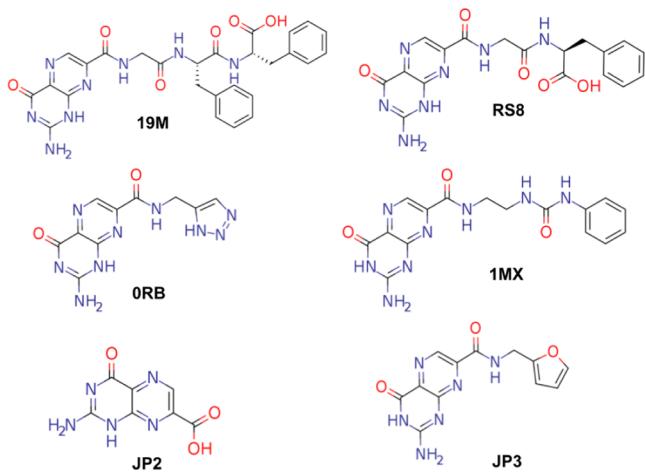
In a recent study, Chaves and co-workers<sup>12</sup> carried out redocking of six known RTA inhibitors and then performed steered molecular dynamics (SMD) simulations to assess the relative binding affinity of these ligands. The molecular docking approach was able to predict the bioactive pose of the ligands; however, the score function was unable to rank these inhibitors according to experimental data. Steered MD simulation was used to decouple these ligands from the binding pocket, and the force profiles were estimated and presented a strong correlation with the experimental data ( $R^2 > 0.9$ ).

As a matter of fact, there are few molecular modeling or simulation studies about ricin, and in our searches, we found no study applying quantum chemical methods for whole RTA–ligand complexes. Thus, in this work, we performed calculations of  $\Delta H_{\text{bind}}$  and quantum descriptors between RTA and some of its inhibitors using semiempirical quantum chemical and QM/MM ONIOM methods to compare the results with experimental data (half-maximal inhibitory concentration ( $IC_{50}$ )) and obtain local interaction information between the RTA residues and RTA inhibitors.

## 2. METHODS

We retrieved the following RTA–ligand complex structures from the Protein Data Bank (PDB): 4HUP (RTA–19M),<sup>56</sup> 4HUO (RTA–RS8),<sup>56</sup> 4ESI (RTA–0RB),<sup>57</sup> 4MX1 (RTA–1MX),<sup>58</sup> 3PX8 (RTA–JP2),<sup>59</sup> and 3PX9 (RTA–JP3).<sup>59</sup> Figure 1 presents the chemical structures of the six ligands of RTA studied in this work.

All RTA–ligand complexes were relaxed and then equilibrated using molecular dynamics simulations. We set the MD simulations according to the settings used in ref 12. The last frame of each MD simulation was used to calculate binding enthalpies using semiempirical quantum mechanical (SQM) methods considering two scenarios: (i) a direct single-point energy calculation with RM1,<sup>60</sup> PM6,<sup>61</sup> PM6-DH+,<sup>62</sup> PM6-D3H4, and PM7<sup>63</sup> and (ii) after full-atom geometry optimization with PM6-DH+, PM6-D3H4, and PM7.



**Figure 1.** Structures of the six ligands of RTA studied in this work. The three-letter code corresponds to the PDB ID codes for these ligands.

For all semiempirical calculations, we used the linear scaling algorithm MOZYME,<sup>64</sup> available in the MOPAC2016 package.<sup>65</sup> For the SCF convergence criteria, we used standard settings and a cutoff radius of 10 Å for the MOZYME algorithm. For full-atom geometry optimizations, we used the limited-memory BFGS algorithm considering a norm of gradient of 5.0 kcal·mol<sup>-1</sup>·Å<sup>-1</sup> as stop criteria for complexes and noncomplexed proteins and 0.01 kcal·mol<sup>-1</sup>·Å<sup>-1</sup> for the free ligands, where no restrictions of movement were considered for the atoms. For all calculations (single-point and geometry optimization calculations), we considered the conductor-like screening model (COSMO) implicit solvent field with a relative permittivity of 78.4 and an effective solvent molecule radius of 1.3 Å. The binding enthalpies for the ligand–RTA complexes were calculated according to eq 1. For all noncomplexed proteins and R–L complexes, we assumed a total charge of +2 e.

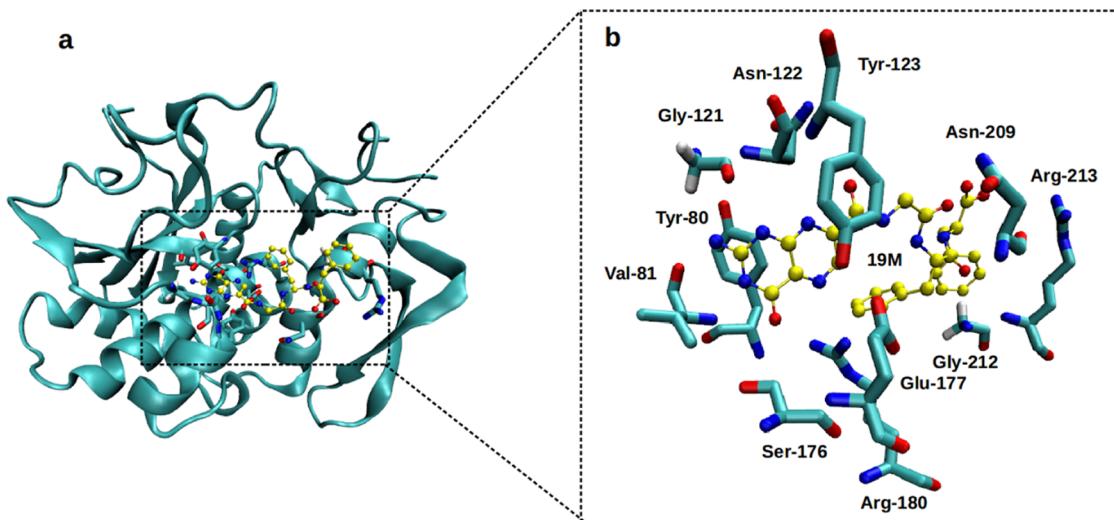
$$\Delta H_{\text{binding}} = \Delta H_f^{\text{ligand-RTA complex}} - (\Delta H_f^{\text{ligand}} + \Delta H_f^{\text{RTA}}) \quad (1)$$

The RTA has 4200 atoms and the ligands 19M, RS8, ORB, 1MX, JP2, and JP3 have 67, 47, 30, 43, 20, and 31 atoms, respectively. The summation between the number of atoms for RTA and each ligand gives the number of atoms for the respective RTA–ligand complex.

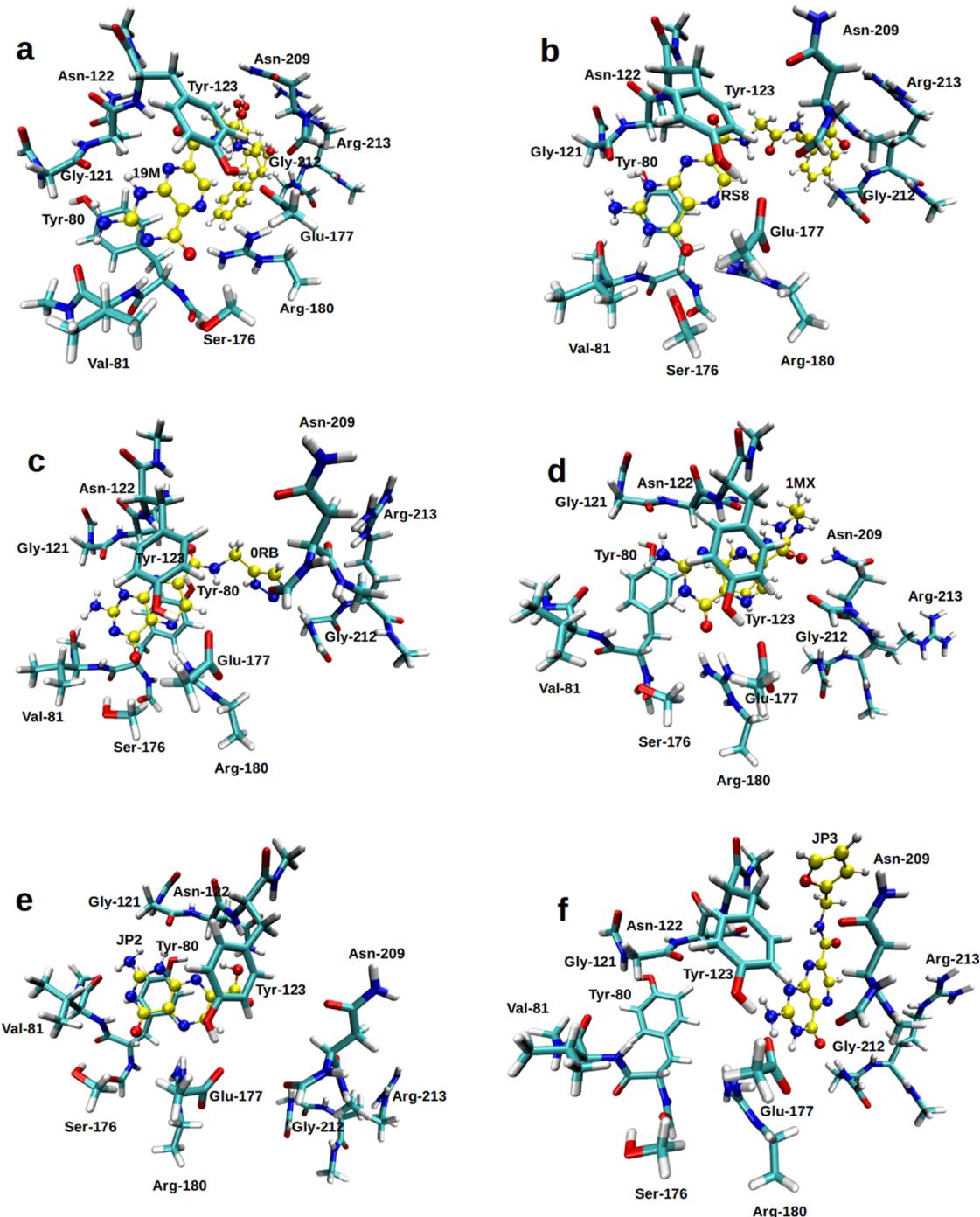
For the sake of comparison and to test the performance of semiempirical methods, we also carried out single-point calculations considering hybrid QM/MM ONIOM method<sup>66–68</sup> using the same geometries from DM for the six RTA–ligand complexes. We used Gaussian 09 program<sup>69</sup> with the hybrid GGA B3LYP functional<sup>70,71</sup> and the hybrid GGA ωB97X-D functional,<sup>72</sup> which includes DFT-D2<sup>73</sup> dispersion correction. In both cases, we used the 6-31+G(d) basis set<sup>74–76</sup> for the QM region in both RTA protein and RTA–ligand complexes. UFF universal force field<sup>77</sup> was used for the remainder of the protein, which is denoted as the MM part. As requested, hydrogen atoms were used as link atoms connecting these two parts. We used all default criteria for ONIOM QM/MM calculations in Gaussian 09 program.

The QM part of the RTA protein includes the following residues: Glu-177, Arg-180 (important for enzyme catalysis), Tyr-80, Val-81, Gly-121, and Tyr-123 (important for attachment and recognition of the ligand at the active site).<sup>78</sup> In addition, we include residues Asn-122, Ser-176, Asn-209, Gly-212, and Arg-213 because they are about 3.0 Å apart from any of the six ligands, thereby forming hydrogen-bond interactions, producing a model with 189 atoms, 716 electrons, and +1 charge. For the RTA–ligand complexes, we also included the respective ligands in the QM part. Therefore, models were generated with (256 atoms, 1008 electrons and +1 charge), (236 atoms, 930 electrons and +1 charge), (219 atoms, 864 electrons and +1 charge), (232 atoms, 908 electrons and +1 charge), (209 atoms, 822 electrons and +1 charge), and (220 atoms, 864 electrons and +1 charge) for RTA–19M, RTA–RS8, RTA–ORB, RTA–1MX, RTA–JP2, and RTA–JP3 complexes, respectively. The binding energies for the ligand–RTA complexes were calculated according to eq 2.

$$\Delta E_{\text{binding}} = \Delta E_{\text{QM/MM}}^{\text{ligand-RTA complex}} - (\Delta E_{\text{QM}}^{\text{ligand}} + \Delta E_{\text{QM/MM}}^{\text{RTA}}) \quad (2)$$



**Figure 2.** (a) Geometry of RTA–19M complex used in the ONIOM QM/MM calculation. (b) Zoomed view of the active site showing the 11 residues and 19M ligand (in yellow).



**Figure 3.** QM models for ONIOM QM/MM calculations for complexes: (a) RTA–19M, (b) RTA–RS8, (c) RTA–0RB, (d) RTA–1MX, (e) RTA–JP2, and (f) RTA–JP3.

In Figure 2, we present the geometry for the RTA–19M complex with emphasis on the residues of the active site and the ligand (in yellow).

In Figure 3, we present QM models used for all complexes. We considered all atoms for residues Tyr-80, Val-81, Gly-121, Asn-122, Tyr-123, Asn-209, Gly-212, and Arg-213 because there are interactions between atoms of protein backbone and the ligands for such residues. For residues Ser-176, Glu-177, and Arg-180, we consider only side chains since intermolecular interactions with ligands occur only in this region of amino acids.

### 3. REACTIVITY DESCRIPTOR CALCULATIONS

The most successful quantum descriptors come from conceptual density functional theory (CDFT), which through the mathematical development of density functional theory has given valid quantitative definitions for well-known chemical concepts,<sup>79</sup> such as hardness and softness. From this theory, the main interactions between two molecules are summarized in two types of processes: mutual polarization followed by molecular orbital overlap interactions and Coulombic forces.<sup>80</sup> The propensity of these effects in molecules can be traced locally using the Fukui function for

the first type of interaction and local hardness for the second.<sup>81</sup>

To locally represent the molecular orbital overlap interactions, we used the Fukui function.<sup>81</sup> This descriptor was further divided into two functions, namely, the left Fukui function ( $f^-$ ), specific for electrophilic attack susceptibility (which for the frozen orbital approximation is equal to the highest-energy occupied molecular orbital (HOMO) density<sup>80</sup>), and the right Fukui function ( $f^+$ ), specific for nucleophilic attack susceptibility (which is equal to the lowest-energy unoccupied molecular orbital (LUMO) density).

A convenient representation of the Fukui functions is where the values are assigned for each  $k$  atom center in the molecule. The  $f^-$  defined in eq 3 is the sum of the squared  $\nu$  atomic orbital (AO) coefficients that comprise the HOMO and belong to the  $k$ th atom, plus the sum of the product between the coefficients of indices  $\nu, \mu$ , and the corresponding overlap matrix element  $S_{\mu\nu}$ .<sup>82</sup> An equivalent definition for the  $f^+$  is presented in eq 4, using the LUMO instead of the HOMO.

$$f^-(k) = \sum_{\nu \in k}^{\text{AO}} |C_{\nu \text{HOMO}}|^2 + \sum_{\mu \notin \nu}^{\text{AO}} |C_{\nu \text{HOMO}} C_{\mu \text{HOMO}}| S_{\mu\nu} \quad (3)$$

$$f^+(k) = \sum_{\nu \in k}^{\text{AO}} |C_{\nu \text{LUMO}}|^2 + \sum_{\mu \notin \nu}^{\text{AO}} |C_{\nu \text{LUMO}} C_{\mu \text{LUMO}}| S_{\mu\nu} \quad (4)$$

In this study, the overall effects tallied in these two Fukui functions were used in combination via the average Fukui function, defined in eq 5.

$$f^0 = \frac{f^+(k) + f^-(k)}{2} \quad (5)$$

These quantum chemical descriptors have been identified as useful for determining the toxicity and biological activity of several potential ligands.<sup>83,84</sup> For the enzymes, these descriptors have been used to determine reactivity sites, as score functions in molecular docking, and to search protein native structures<sup>85</sup> and find key structures in reaction path simulations.<sup>86,87</sup>

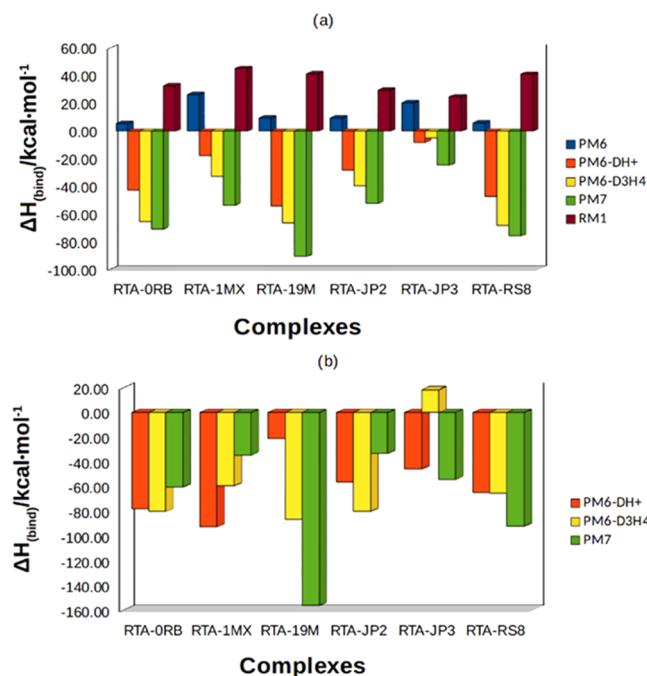
The electronic structure of protein systems presents several relevant molecular orbitals for chemical interactions with the electronic energy near the HOMO and LUMO.<sup>84,86</sup> Thus, in this study, we computed the Fukui functions not only by using the HOMO and LUMO orbitals but also by considering all of the molecular orbitals from a range of 3 eV from the HOMO and LUMO. As Fukushima and co-workers showed in their work, this value can vary from 1 to 5 eV depending on the system under study.<sup>88</sup>

We used the local hardness ( $H(k)$ ) as a quantum descriptor to compute the Coulombic force interactions. This particular definition is based on the electron–electron contribution of the molecular electrostatic potential, as shown in eq 6;<sup>89</sup> the calculation of local hardness at the  $k$ th atomic center is defined as a sum of the left Fukui function for each  $l$ th atomic center divided by their Euclidean distance  $R_{kl}$ . These theoretical quantities were calculated using the PRIMoRDIA software.<sup>90</sup>

$$H(k) = \sum_{l \neq k}^{\text{atoms}} \frac{f^-(l)}{R_{kl}} \quad (6)$$

## 4. RESULTS AND DISCUSSION

In Figure 4, we present the  $\Delta H_{\text{bind}}$  results of all of the RTA–ligand complexes studied in this work, considering both



**Figure 4.** Binding enthalpies,  $\Delta H_{\text{bind}}$ , for the complexes (RTA-0RB, RTA-1MX, RTA-19M, RTA-JP2, RTA-JP3, and RTA-RS8). In (a), the results are depicted from single-point calculations, and in (b), the results are after full-atom geometry optimization of R, L, and R-L.

strategies are carried out: (i) direct single-point energy calculation and (ii) calculation of  $\Delta H_{\text{bind}}$  after full-atom geometry optimization of R, L, and R-L.

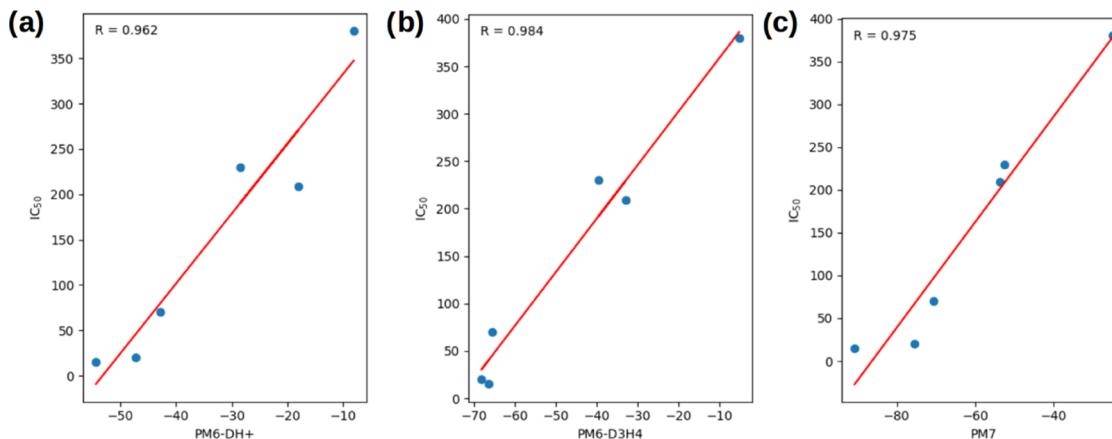
Looking at the results for  $\Delta H_{\text{bind}}$  calculated using the single-point strategy (Figure 4a), we can see that PM6-DH+, PM6-D3H4, and PM7 predicted negative  $\Delta H_{\text{bind}}$  values for all studied RTA–ligand complexes. The range of  $\Delta H_{\text{bind}}$  values is also consistent with the expected results from ligand–enzyme thermochemical assays, approximately a dozen of kcal·mol⁻¹.<sup>91</sup>

PM6 and RM1 showed low performance, presenting positive values for  $\Delta H_{\text{bind}}$  for all RTA–ligand complexes. By comparing the performance among the PM6, PM6-DH+, and PM6-D3H4 methods, we verified that there are differences in binding enthalpies when corrections for dispersion and hydrogen bonding are considered. These results are consistent with recent studies suggesting that the quality of semi-empirical methods presents significant improvements when such corrections are considered.<sup>16,18,92</sup> These studies also indicate that SQM methods incorporating dispersion and hydrogen bonding yield results similar to those of D-type corrections for density functional theory (DFT) methods.

Figure 4b shows the results of  $\Delta H_{\text{bind}}$  after full-atom geometry optimization for all complexes using the PM6-DH+, PM6-D3H4, and PM7 methods. We verified that the methods followed the same tendency of the single-point calculations, i.e., the calculated binding enthalpies were negative for all RTA–ligand complexes. The only exception was the RTA–JP3 ligand, which presented a positive  $\Delta H_{\text{bind}}$  value with the PM6-D3H4 method. This result suggests that the PM6-DH+,

**Table 1.** Binding Enthalpies ( $\text{kcal}\cdot\text{mol}^{-1}$ ) for the Six RTA–Ligand Complexes Evaluated in This Study

ligand	$\text{IC}_{50}$ ( $\mu\text{M}$ )	$\Delta H_{\text{bind}}$ via single-point calculations					$\Delta H_{\text{bind}}$ via geometry optimizations		
		PM6	PM6-DH+	PM6-D3H4	PM7	RM1	PM7	PM6-DH+	PM6-D3H4
19M	15 (1)	9.58 (3)	-54.37 (1)	-66.47 (2)	-90.76 (1)	41.27 (5)	-156.11 (1)	-20.97 (6)	-86.38 (1)
RS8	20 (2)	6.10 (2)	-47.31 (2)	-68.23 (1)	-75.59 (2)	41.12 (4)	-93.33 (2)	-64.95 (3)	-65.33 (4)
0RB	70 (3)	5.55 (1)	-42.88 (3)	-65.53 (3)	-70.5 (3)	32.74 (3)	-60.5 (3)	-78.21 (2)	-80.12 (2)
1MX	209 (4)	26.28 (6)	-18.06 (5)	-33.01 (5)	-53.65 (4)	45.24 (6)	-34.41 (5)	-92.76 (1)	-59.54 (5)
JP2	230 (5)	9.71 (4)	-28.5 (4)	-39.53 (4)	-52.39 (5)	-29.51 (1)	-33.23 (6)	-56.1 (4)	-79.94 (3)
JP3	380 (6)	20.54 (5)	-8.12 (6)	-5.16 (6)	-24.89 (6)	24.78 (2)	-54.16 (4)	-45.87 (5)	18.35 (6)

**Figure 5.** Correlation graphs between  $\text{IC}_{50}$  and  $\Delta H_{\text{bind}}$  data obtained by single-point calculations for the (a) PM6-DH+, (b) PM6-D3H4, and (c) PM7 methods.

PM6-D3H4, and PM7 methods are able to describe, at least qualitatively, the binding enthalpies of the studied systems. We observed that the full-atom geometry optimization showed a tendency to decrease the  $\Delta H_{\text{bind}}$  value of the complexes. From the 18  $\Delta H_{\text{bind}}$  calculations performed (optimization of six complexes with three distinct semiempirical methods: PM6-DH+, PM6-D3H4, and PM7), 12 presented a decreased  $\Delta H_{\text{bind}}$  value compared to those obtained by single-point calculations. Exceptions occurred for the RTA–0RB ( $\Delta H_{\text{bind}} = -60.50 \text{ kcal}\cdot\text{mol}^{-1}$ ), RTA–1MX ( $\Delta H_{\text{bind}} = -34.41 \text{ kcal}\cdot\text{mol}^{-1}$ ) and RTA–JP2 ( $\Delta H_{\text{bind}} = -33.23 \text{ kcal}\cdot\text{mol}^{-1}$ ) complexes with the PM7 method, RTA–JP3 ( $\Delta H_{\text{bind}} = 18.35 \text{ kcal}\cdot\text{mol}^{-1}$ ) and RTA–RS8 ( $\Delta H_{\text{bind}} = -65.33 \text{ kcal}\cdot\text{mol}^{-1}$ ) complexes with the PM6-D3H4 method and the RTA–19M ( $\Delta H_{\text{bind}} = -20.97 \text{ kcal}\cdot\text{mol}^{-1}$ ) complex with the PM6-DH+ method.

In Table 1, we present the  $\text{IC}_{50}$ <sup>56–59</sup> and  $\Delta H_{\text{bind}}$  values for each RTA–ligand complex evaluated in this work.

When comparing the  $\Delta H_{\text{bind}}$  results obtained by single-point calculations using the PM6, PM6-DH+, PM7, and RM1 methods with  $\text{IC}_{50}$  experimental data for the ligands (19M, RS8, 0RB, 1MX, JP2, and JP3),<sup>12</sup> we found that the PM6 method showed correlation of 0.688 and RM1 method showed no correlation with the  $\text{IC}_{50}$ . On the other hand, the PM6-DH+ and PM7 methods were able to identify the 19M ligand, which has the lowest  $\text{IC}_{50}$  value ( $15 \mu\text{M}$ ), as the best ligand.

The PM6-D3H4 method switched the rank order of the 19M and RS8 ligands ( $\Delta H_{\text{bind}} = -66.47$  and  $-68.23 \text{ kcal}\cdot\text{mol}^{-1}$ , respectively), but this may have occurred because the  $\text{IC}_{50}$  values of these two ligands are very close ( $15$  and  $20 \mu\text{M}$ ) so that the PM6-D3H4 method was not sensitive enough to rank the best ligand for small  $\text{IC}_{50}$  variations. The PM6-DH+

+ method also showed inversions of binding enthalpy values between the 1MX ( $\Delta H_{\text{bind}} = -18.06 \text{ kcal}\cdot\text{mol}^{-1}$ ) and JP2 ( $\Delta H_{\text{bind}} = -28.50 \text{ kcal}\cdot\text{mol}^{-1}$ ) methods. The PM7 method was able to correctly rank all of the ligands without inversions, being more sensitive to the small variations of  $\text{IC}_{50}$ .

When we performed the statistical treatment between the  $\text{IC}_{50}$  and  $\Delta H_{\text{bind}}$  data obtained by single-point calculations, we found  $R$  values of 0.962, 0.975, and 0.984 for the PM6-DH+, PM7, and PM6-D3H4 methods, respectively. Thus, the PM6-D3H4 method presented the best correlation with the  $\text{IC}_{50}$  data. So, it is worth mentioning that the results for PM6-DH+ and PM6-D3H4 are quite satisfactory. In Figure 5, we present the correlation graphs between the  $\text{IC}_{50}$  and  $\Delta H_{\text{bind}}$  data obtained by single-point calculations when the PM6-DH+, PM6-D3H4, and PM7 methods were used.

When analyzing the binding enthalpy data with full-atom geometry optimization (Table 1), we verified that the PM6-DH+ method, although presenting good results with single-point calculations, did not show the same performance. This method was not able to identify 19M as the best ligand or present a good correlation with the  $\text{IC}_{50}$  data.

The PM7 method could rank the best ligands (from 1 to 3), but there were inversions in the ranking of the other ligands. Besides, there is a very marked distortion between the variation in the binding enthalpy results and  $\text{IC}_{50}$  values. The PM6-D3H4 method was able to identify the best ligand (19M) but showed inversions in the ranking between the RS8 and 0RB ligands and 1MX and JP2 ligands. However, compared to the PM7 method, the PM6-D3H4 method presented lower distortions between the  $\Delta H_{\text{bind}}$  and  $\text{IC}_{50}$  data.

When the statistical treatment was performed between the  $\text{IC}_{50}$  and  $\Delta H_{\text{bind}}$  data, we found  $R$  values of -0.085, 0.672, and 0.789 for the PM6-DH+, PM7, and PM6-D3H4 methods,

respectively. Therefore, the PM6-D3H4 method also presented the best correlation with the IC<sub>50</sub> when we performed a full-atom geometry optimization strategy. Moreover, we observed that in the full-atom geometry optimization, the correlation between the IC<sub>50</sub> data and binding enthalpies decreased considerably.

Sulimov et al.<sup>93</sup> recalculated the energies of approximately 8000 best-ranked structures in the molecular docking with the MMFF94 force field through single-point calculations and ligand optimization using the PM7 method, considering an implicit COSMO solvent model. The authors observed that the energies obtained by single-point calculations greatly improved the ranking of structures close to the native state. However, when optimizing the ligands with the PM7 method in vacuum and recalculating their energies with the COSMO model, the authors observed a worsening of the results for several complexes compared to those with PM7 (single-point energy calculation). In accordance with those results, we observed more comprehensively that the geometry optimizations reduced the performance of all considered semiempirical methods.

Although the correlation for full-atom geometry optimizations is lower than that obtained through single-point energy calculations, we emphasize that if the molecular dynamics simulation is not carried out properly to find an equilibrated structure, then this optimization is necessary to obtain some correlation with experimental data. We have indicated PM7 and PM6-D3H4 as the best semiempirical methods for obtaining binding enthalpies and ligand rankings. We are not stating that single-point QM calculations on the end frame from MD simulation are better for ligand-binding energies than ones that use geometry optimization at the QM level. However, this was true for our study. Effects due to low sampling, entropy lacking, or cancelation of errors could be behind these results, but the investigation of these issues is outside the scope of this study. Anyhow, an excellent discussion about these issues can be found in the review by Ryde and Soderhjelm.<sup>10</sup>

In Table 2, we present the root-mean-square deviation (RMSD) data between the MD structures and those optimized with the PM6-DH+, PM6-D3H4, and PM7 methods.

When analyzing the data in Table 2, we verified that the PM6-D3H4 method presented the lowest RMSD values, except for the RTA-JP2 complex, where the PM6-DH+ method presented a lower RMSD value. The PM7 method presented the highest RMSD values for all of the analyzed

**Table 2. RMSD Results (in Å) between the End Frame of the MD Simulations and the Optimization of These Same End Frames by Semiempirical Quantum Methods<sup>a</sup>**

complex	PM6-DH+ (Å)	PM6-D3H4 (Å)	PM7 (Å)	crystallographic (Å)
RTA-19M	1.554	1.296	2.108	2.176
RTA-RS8	1.747	1.533	2.024	2.266
RTA-0RB	1.567	1.401	1.950	1.839
RTA-1MX	1.832	1.468	1.919	2.676
RTA-JP2	1.581	1.626	2.036	1.917
RTA-JP3	1.745	1.511	2.057	2.103

<sup>a</sup>In the last column, we present RMSD results (in Å) between the end frame of the MD simulations and each crystallographic structure.

complexes. This observation suggests that the PM7 method is able to relax the initial structure more than the PM6-DH+ and PM6-D3H4 methods. Figure 6 shows the superposition between the initial structure and the structures optimized via the PM7 method.

From the analysis of the RTA structures in the complexes, we verified that there were no significant changes after optimization, with conservation of the secondary structures. All R-L complexes presented RMSDs close to 2.0 Å, which is consistent with the resolution of the experimental data.<sup>94</sup>

When comparing the ligand poses via full-atom geometry optimization with structures from molecular dynamics, we observed that the 0RB ligand (Figure 6a) presented rotation of the triazole group, being approximately perpendicular to the preferred position of this group in the protein. The ligand 1MX (Figure 6b) presented small pterin group torsion and benzene ring variation. Due to its high conformational flexibility, the 19M ligand (Figure 6c) underwent a large modification after geometry optimization, mainly regarding torsions in the pterin moiety, benzene rings, and carboxylic acid group. The JP2 ligand (Figure 6d) exhibited rotation of its carboxylic acid group and small deviations in other groups of the structure. The JP3 ligand (Figure 6e) showed modification of its structure, with displacements in the positions of atoms along the same plane. The RS8 ligand (Figure 6f) exhibited a small variation in comparison to that observed in the last frame obtained from the molecular dynamics simulations.

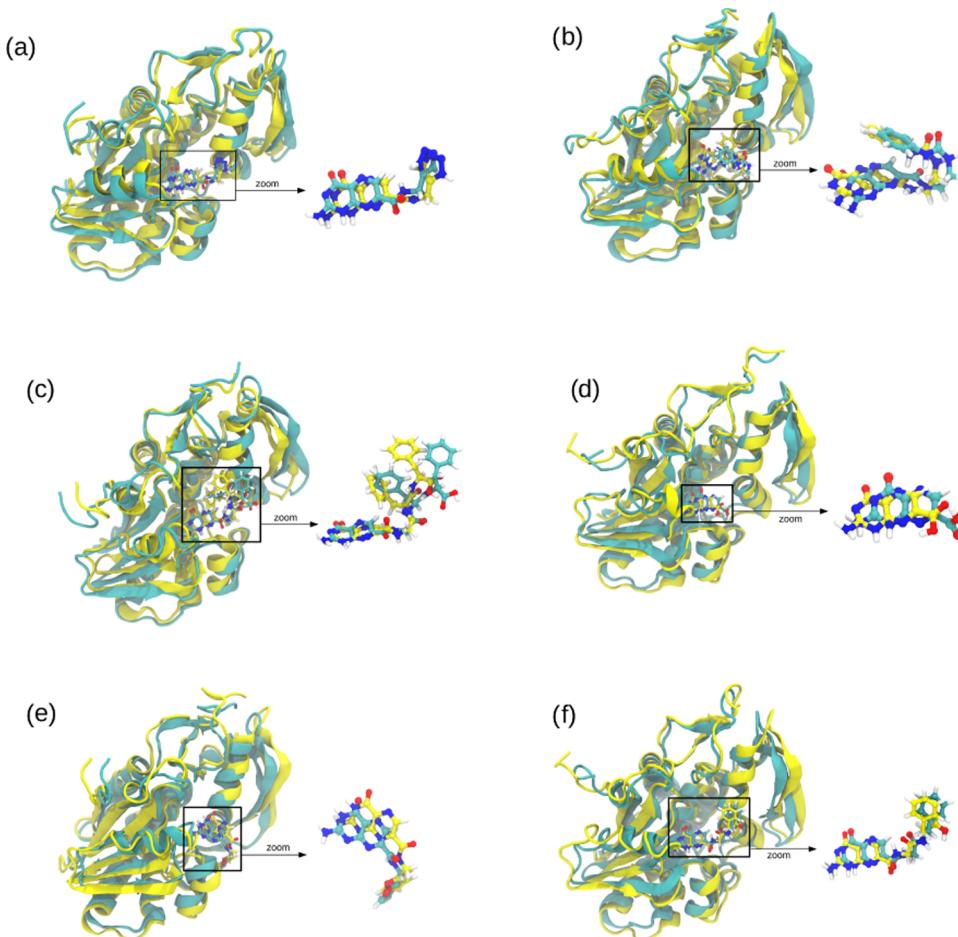
In Table 3, we present the RMSD data between the ligands of complexes from the MD simulations and the ligands of the complexes optimized with the PM6-DH+, PM6-D3H4, and PM7 methods.

When analyzing the data of Table 3, we verified that the poses of the ligands presented several variations after the geometry optimization for all R-L complexes. For the PM6-DH+ method, the 1MX ligand presented the highest variation (RMSD = 0.932 Å), and the JP3 ligand presented the lowest variation for this method (RMSD = 0.581 Å). For the PM6-D3H4 method, it was observed that the RS8 and 1MX ligands presented the highest and lowest variations, respectively, with RMSDs equal to 1.291 and 0.477 Å. For the PM7 method, the 19M ligand showed the highest variation (RMSD = 1.530 Å), and the lowest variation for this method was for the RS8 ligand (RMSD = 0.598 Å). We observed that the PM7 method showed higher RMSD values for four of the six ligands, and the two exceptions were the RS8 and 1MX ligands.

In Table 4, we present the IC<sub>50</sub><sup>56–59</sup> and ΔE<sub>bind</sub> values for each RTA-ligand complex calculated via ONIOM QM/MM with B3LYP and ωB97X-D functionals.

When analyzing the data in Table 4, we found that the single-point ONIOM QM/MM calculations with B3LYP functional were able to identify the three best ligands (19M, RS8, and 0RB). The only exception was the inversion between the ligands 1MX and JP2. The correlation between the IC<sub>50</sub> and ΔE<sub>bind</sub> data presented an R-value of 0.929. Although the ONIOM QM/MM results with B3LYP functional showed good correlation with IC<sub>50</sub>, PM6-DH+, PM7, and PM6-D3H4 showed better results.

For the sake of testing, we also carried out single-point ONIOM QM/MM calculations with the B3LYP functional considering a smaller QM region for each complex. In this test, we considered in the QM part only the six most



**Figure 6.** Superposition of the RTA–ligand optimized complexes: (a) RTA–0RB, (b) RTA–1MX, (c) RTA–19M, (d) RTA–JP2, (e) RTA–JP3, and (f) RTA–RS8. The optimized structures are represented in green, and initial structures are represented in yellow.

**Table 3. RMSD Results (in Å) between the Ligands of the Complexes Optimized by Semiempirical Methods and the Ligands of the Complexes of the MD Simulations<sup>a</sup>**

complex	PM6-DH+ (Å)	PM6-D3H4 (Å)	PM7 (Å)	crystallographic (Å)
19M	0.675	0.657	1.530	6.005
RS8	0.808	1.291	0.598	5.819
0RB	0.761	0.755	1.062	2.856
1MX	0.932	0.477	0.812	4.318
JP2	0.889	0.500	0.909	1.957
JP3	0.581	0.510	0.612	2.597

<sup>a</sup>In the last column, we present RMSD results (in Å) between the ligands of the MD simulations and the ligands from crystallographic structures.

**Table 4. Binding Energies,  $\Delta E_{\text{bind}}$ , (in hartrees) for the Six RTA–Ligand Complexes Calculated with ONIOM QM/MM Single-Point Calculations**

complex	$IC_{50}$ ( $\mu\text{M}$ )	$\Delta E_{\text{bind}}/\text{B3LYP}$ (au)	$\Delta E_{\text{bind}}/\omega\text{B97X-D}$ (au)
19M	15 (1)	-0.136 (1)	-0.230 (1)
RS8	20 (2)	-0.119 (2)	-0.202 (2)
0RB	70 (3)	-0.106 (3)	-0.178 (3)
1MX	209 (4)	-0.062 (6)	-0.130 (4)
JP2	230 (5)	-0.070 (4)	-0.129 (5)
JP3	380 (6)	-0.057 (5)	-0.088 (6)

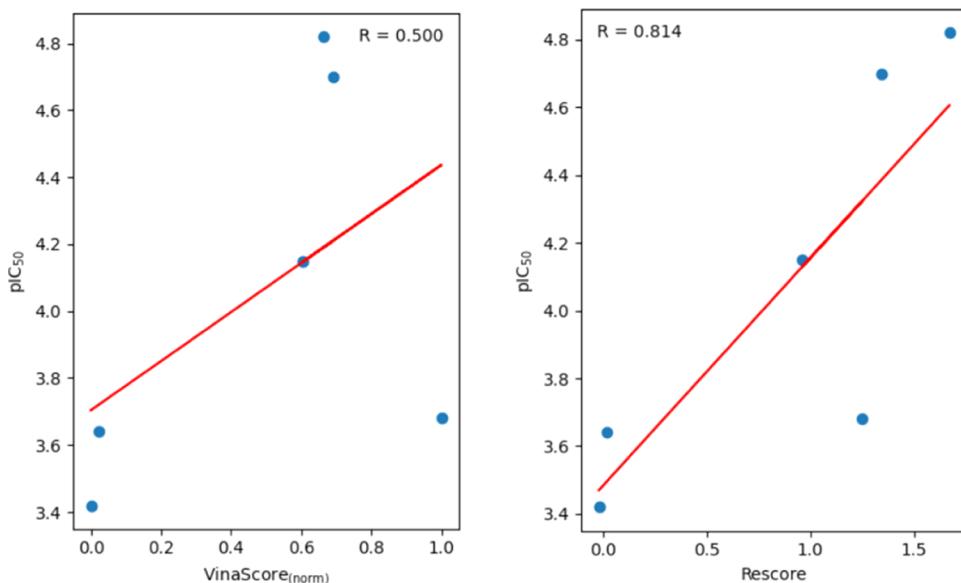
important residues of RTA plus the ligand. When we did this the correlation decreases dramatically to 0.693.

When we changed the B3LYP functional to the  $\omega$ B97X-D functional, which includes DFT-D2<sup>73</sup> dispersion correction, we observe that the R-value increased from 0.929 to 0.972, showing a slight improvement in the correlation with  $IC_{50}$ . In addition, calculated results obtained with this functional, all ligands were ranked correctly according to  $\Delta E_{\text{bind}}$  and  $IC_{50}$  values.

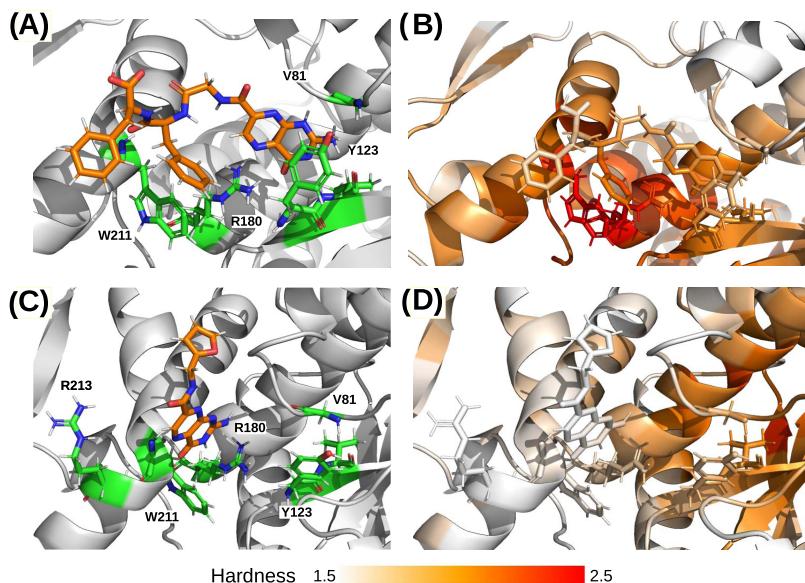
Considering these results and the low computational cost of semiempirical methods, we can state that methods PM7, PM6-DH+, and PM6-D3H4 showed better performance than the ONIOM QM/MM calculation with B3LYP functional for RTA complexes studied in this paper. When using the  $\omega$ B97X-D functional, the QM/MM method presented R-value and performance similar to the PM7 method but with a higher computational cost. The advantage of ONIOM QM/MM is that one can always improve the results by increasing the QM region and using larger basis sets, and/or taking solvent effects and dispersion corrections into account, but this ends up being computationally expensive.

As a final experiment with binding affinities for these RTA complexes, we carried out new calculations using the dataset of binding affinities produced in the study carried out by Chaves and his collaborators.<sup>12</sup>

Our inspiration was a study carried out by Gupta and collaborators, which achieved a significant improvement for



**Figure 7.** Pearson correlation between  $\text{pIC}_{50}$  vs Vina score and  $\text{pIC}_{50}$  vs re-score that is defined in eq 7.



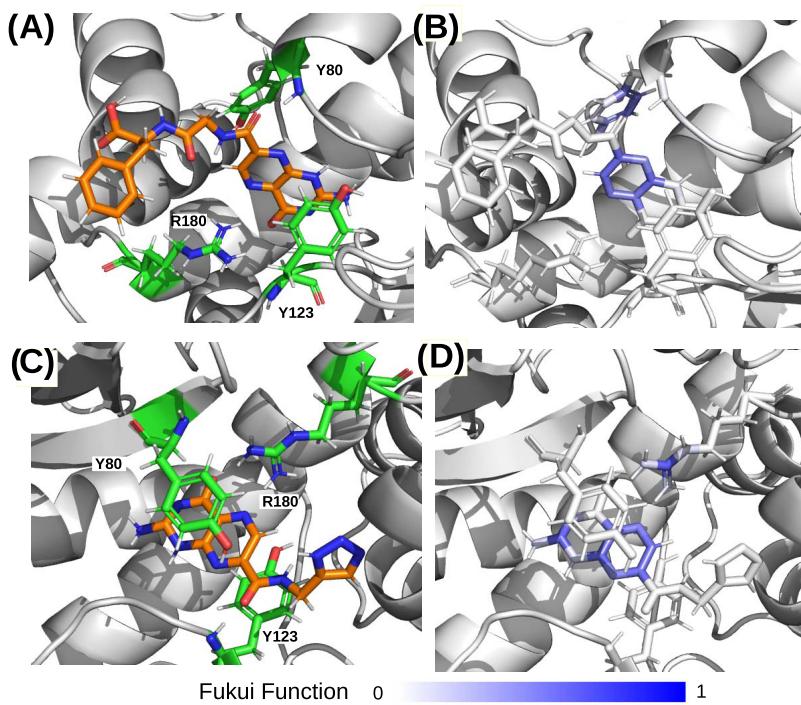
**Figure 8.** Calculated local hardness for RTA complexes with the 19M and JP3 ligands (orange). (A) Labels for the nearest residues (green) from 19M; (B) local hardness for the RTA–19M complex; (C) labels for the nearest residues (green) from JP3; and (D) local hardness for the RTA–JP3 complex.

the binding affinity predictions when they considered molecular descriptors in different molecular docking approaches for the inhibitors of microsomal prostaglandin E synthase-1.<sup>95</sup> In their study, molecular descriptors such as topological polar surface area (TPSA), partition coefficient ( $\log P$ ), volume (Vol), and the number of rotatable bonds (Nrtb) were used as docking scores. With this, the authors took into account desolvation penalties and conformational free-energy changes when a ligand binds to a protein. In their study, the re-score routine was considered as an empirical summation of normalized docking affinities. Similarly, we also incorporated a re-score routine to the binding affinity predictions provided by Vina software<sup>96</sup> for the same set of ligands (19M, RS8, ORB, 1MX, JP2, and JP3). These binding affinity predictions were retrieved from another validation study carried out by our group; therefore, details about

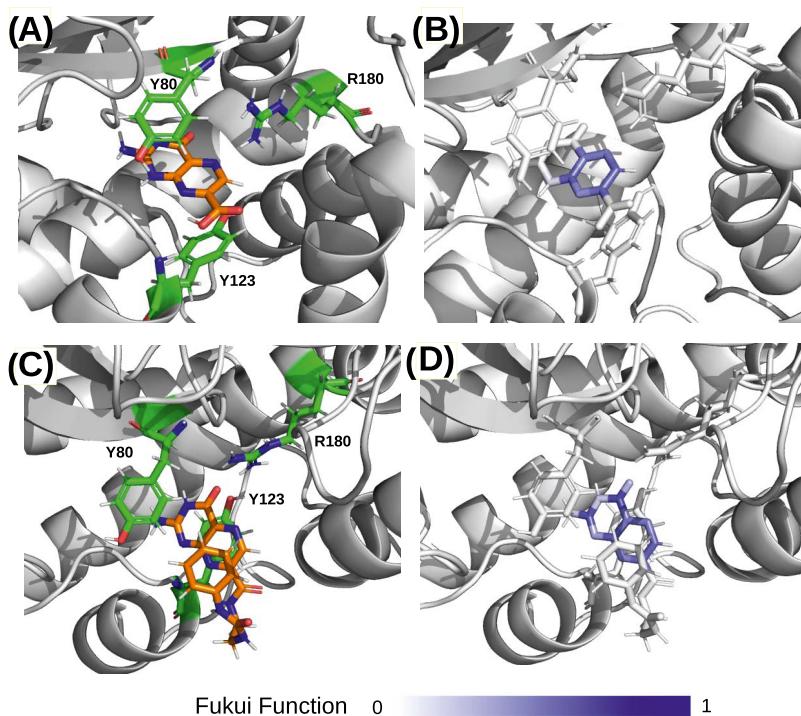
docking protocol can be reviewed in ref 12. In summary, the ligand molecular descriptors such as TPSA, Vol, and Nrtb were calculated using the web service Molinspiration (<https://www.molinspiration.com>). We also normalized the docking scores and the molecular descriptors to values between 0 and 1. Then, the following equation was used to perform re-score calculations

$$\text{rescore} = \text{docking score} + \text{TPSA} + \text{Vol} - \text{Nrtb} \quad (7)$$

We present in Figure 7 the Pearson correlations between  $\text{pIC}_{50}$  vs Vina score and  $\text{pIC}_{50}$  vs re-score, respectively. In short, our results showed that the consideration of such molecular descriptors improves the relative binding affinity predictions provided by Vina software, with an increase from 0.500 to 0.814 in the Pearson correlation.



**Figure 9.** Calculated Fukui function for RTA complexes with the RS8 and ORB ligands (orange). (A) Labels for the nearest residues (green) from RS8; (B) Fukui function for the RTA–RS8 complex; (C) labels for the nearest residues (green) from ORB; and (D) local hardness for the RTA–ORB complex.



**Figure 10.** Calculated Fukui function for RTA complexes with the 1MX and JP2 ligands (orange). (A) Labels for the nearest residues (green) from 1MX; (B) Fukui function for RTA–1MX complex; (C) labels for the nearest residues (green) from JP2; and (D) local hardness for the RTA–JP2 complex.

## 5. REACTIVITY DESCRIPTORS

We calculated the reactivity descriptors for the protein–ligand complexes and used to perform a graphical analysis to theoretically characterize the most important interactions that occur between the inhibitors and the active site residues. As

shown in Figure 8, the local hardness has higher values in the binding pocket for the RTA–19M complex, with its highest value at Arg-180 and Trp-211. For the RTA complex with the JP3 ligand, the electrostatic interactions are not placed within the ligand region. Significant Fukui function values were

computed in the residues of the binding pocket for the three most active ligand complexes and with small values for the JP2 ligand. Figure 9 depicts the Fukui function for the second and third most active cases. For the RTA–RS8 complex, the Fukui function localizes at one of the ligand rings of the pterin group and in Tyr-80. In the RTA–ORB complex, all atoms in the pterin group present high values of the given descriptor, as does Arg-180.

In Figure 10, the Fukui function is presented for 1MX and JP2, showing the distribution of the descriptor at the pterin group and none at the closest residues. The molecular structure of these ligands is based on the pterin group,<sup>58</sup> the same group present in adenine that is hydrolyzed in the ribosome by the action of RTA catalysis.<sup>5</sup> In the complexes considered by calculations, the reactivity in this group is always indicated by the Fukui function.

The models built from thermochemical properties correlations could rank the best inhibitors, but only reactivity descriptors can pose, which are the molecular structure features that new ligands must have to be more efficient.

## 6. CONCLUSIONS

The main objective of this work was to present a study where two computational approaches were applied (calculation of  $\Delta H_{\text{bind}}$  and reactivity quantum chemical descriptors) to theoretically gain insights into the interactions between the RTA subunit of ricin and some of its inhibitors.

In our study, we observed that single-point energy calculations of  $\Delta H_{\text{bind}}$  with the PM6-DH+, PM6-D3H4, and PM7 methods presented an excellent correlation with the IC<sub>50</sub> data. In addition, although the PM7 method presented a slightly lower correlation than the PM6-D3H4 method, only the PM7 method was able to correctly rank all of the ligands analyzed. This result suggests that the PM7 method is more sensitive to small IC<sub>50</sub> variations.

When comparing the IC<sub>50</sub> data with the calculated  $\Delta H_{\text{bind}}$  values obtained after full-atom optimization with the PM6-DH+, PM6-D3H4, and PM7 methods, we find that the correlation decreases significantly. Although the correlation was greatly reduced with the optimization of the structures, the PM6-D3H4 and PM7 methods were able to correctly rank the best ligand. This result suggests that if the molecular dynamics is not carried out properly to obtain an equilibrated structure, it is necessary to carry out a full-atom geometry optimization to obtain some correlation with the experimental data.

We also observed that in geometry optimization, the structure adopts a minimum local conformation that contributes less to the preferential position of the ligand at the active site. On the other hand, the representative structure from molecular dynamics represents the preferred position of the ligand in the active site. Thus, we conclude that for the dataset studied, it is preferable to use equilibrated MD structures to perform single-point energy calculations to obtain  $\Delta H_{\text{bind}}$ .

ONIOM QM/MM single-point calculations with the B3LYP functional showed good correlation with IC<sub>50</sub>; however, the methods PM6-DH+, PM6-D3H4, and PM7 showed better performance. When using the  $\omega$ B97X-D functional that includes DFT-D2<sup>73</sup> dispersion correction, the QM/MM method presented R-value and performance similar to the PM7 method. Besides, ONIOM QM/MM calculations for  $\Delta H_{\text{bind}}$  incurred a higher computational cost compared

with semiempirical methods, about 250 times higher (B3LYP) and 1400 times higher ( $\omega$ B97X-D).

In addition, we found for the cases studied, reactivity descriptors pointed out that both types of interactions, molecular overlap and electrostatic interactions, play an important role in the overall affinity of these ligands for the RTA binding pocket. From a relatively simple computational protocol, this approach opens the possibility to reveal additional information using structures treated with earlier simulations.

## ■ AUTHOR INFORMATION

### Corresponding Author

Gerd Bruno Rocha – Department of Chemistry, Federal University of Paraíba, João Pessoa, PB 58051-900, Brazil;  
 [orcid.org/0000-0001-9805-9497](https://orcid.org/0000-0001-9805-9497); Phone: +55-83-3216-7437; Email: [gbr@quimica.ufpb.br](mailto:gbr@quimica.ufpb.br)

### Authors

Acassio Rocha-Santos – Department of Chemistry, Federal University of Paraíba, João Pessoa, PB 58051-900, Brazil;  
 [orcid.org/0000-0001-9928-9067](https://orcid.org/0000-0001-9928-9067)

Elton José Ferreira Chaves – Department of Biotechnology, Federal University of Paraíba, João Pessoa, PB 58051-900, Brazil

Igor Barden Grillo – Department of Chemistry, Federal University of Paraíba, João Pessoa, PB 58051-900, Brazil  
Amanara Souza de Freitas – Department of Chemical Engineering, Federal University of Paraíba, João Pessoa, PB 58051-900, Brazil

Demétrius Antônio Machado Araújo – Department of Biotechnology, Federal University of Paraíba, João Pessoa, PB 58051-900, Brazil

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acsomeg.0c02588>

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support from the Brazilian agencies, institutes, and networks: Instituto Nacional de Ciéncia e Tecnologia de Nanotecnologia para Marcadores Integrados (INCT-INAMI), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES), Fundação de Apoio à Pesquisa do Estado da Paraíba (FAPESQ-PB), Programa de Apoio a Núcleos de Exceléncia (PRONEX-FACEPE), and Financiadora de Estudos e Projetos (FINEP). They also acknowledge the physical structure and computational support provided by Universidade Federal da Paraíba (UFPB), the computer resources of Centro Nacional de Processamento de Alto Desempenho em São Paulo (CENAPAD-SP), and Núcleo de Processamento de Alto Desempenho da Universidade Federal do Rio Grande do Norte (NPAD/UFRN). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES) through the research project Bioinformática Estrutural de Proteínas: Modelos, Algoritmos e Aplicações Biotecnológicas (Edital Biologia Computacional 51/2013, processo AUXPE1375/2014 da CAPES). G.B.R. acknowledges support from the

Brazilian National Council for Scientific and Technological Development (CNPq grant no. 309761/2017-4).

## ■ REFERENCES

- (1) Bozza, W. P.; Tolleson, W. H.; Rosado, L. A. R.; Zhang, B. Ricin detection: Tracking active toxin. *Biotechnol. Adv.* **2015**, *33*, 117–123.
- (2) Montfort, W.; Villafranca, J. E.; Monzingo, A. F.; Ernst, S. R.; Katzin, B.; Rutenber, E.; Xuong, N. H.; Hamlin, R.; Robertus, J. D. The three-dimensional structure of ricin at 2.8 Å. *J. Biol. Chem.* **1987**, *262*, 5398–5403.
- (3) Rutenber, E.; Katzin, B. J.; Ernst, S.; Collins, E. J.; Mlsna, D.; Ready, M. P.; Robertus, J. D. Crystallographic refinement of ricin to 2.5 Å. *Proteins: Struct., Funct., Bioinf.* **1991**, *10*, 240–250.
- (4) Robertus, J. D.; Monzingo, A. F. The structure of ribosome inactivating proteins. *Mini-Rev. Med. Chem.* **2004**, *4*, 477–486.
- (5) Endo, Y.; Tsurugi, K. RNA N-glycosidase activity of ricin A-chain. Mechanism of action of the toxic lectin ricin on eukaryotic ribosomes. *J. Biol. Chem.* **1987**, *262*, 8128–8130.
- (6) May, K. L.; Yan, Q.; Turner, N. E. Targeting ricin to the ribosome. *Toxicon* **2013**, *69*, 143–151.
- (7) Audi, J.; Belson, M.; Patel, M.; Schier, J.; Osterloh, J. Ricin poisoning: a comprehensive review. *JAMA, J. Am. Med. Assoc.* **2005**, *294*, 2342–2351.
- (8) Argent, R. H.; Parrott, A. M.; Day, P. J.; Roberts, L. M.; Stockley, P. G.; Lord, J. M.; Radford, S. E. Ribosome-mediated folding of partially unfolded ricin A-chain. *J. Biol. Chem.* **2000**, *275*, 9263–9269.
- (9) Simpson, J. C.; Roberts, L. M.; Römisch, K.; Davey, J.; Wolf, D. H.; Lord, J. M. Ricin A chain utilises the endoplasmic reticulum-associated protein degradation pathway to enter the cytosol of yeast. *FEBS Lett.* **1999**, *459*, 80–84.
- (10) Ryde, U.; Söderhjelm, P. Ligand-binding affinity estimates supported by quantum-mechanical methods. *Chem. Rev.* **2016**, *116*, 5520–5566.
- (11) Nikitina, E.; Sulimov, V.; Zayets, V.; Zaitseva, N. Semi-empirical calculations of binding enthalpy for protein-ligand complexes. *Int. J. Quantum Chem.* **2004**, *97*, 747–763.
- (12) Chaves, E. J.; Padilha, I. Q.; Araújo, D. A.; Rocha, G. B. Determining the Relative Binding Affinity of Ricin Toxin A Inhibitors by Using Molecular Docking and Nonequilibrium Work. *J. Chem. Inf. Model.* **2018**, *58*, 1205–1213.
- (13) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.
- (14) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humbert, C. Comparison of several molecular docking programs: pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2009**, *49*, 1455–1474.
- (15) Yilmazer, N. D.; Korth, M. Comparison of molecular mechanics, semi-empirical quantum mechanical, and density functional theory methods for scoring protein-ligand interactions. *J. Phys. Chem. B* **2013**, *117*, 8075–8084.
- (16) Yilmazer, N. D.; Heitel, P.; Schwabe, T.; Korth, M. Benchmark of electronic structure methods for protein-ligand interactions based on high-level reference data. *J. Theor. Comput. Chem.* **2015**, *14*, No. 1540001.
- (17) Fanfrlík, J.; Bronowska, A. K.; Řezáč, J.; Přenosil, O.; Konvalinka, J.; Hobza, P. A reliable docking/scoring scheme based on the semiempirical quantum mechanical PM6-DH2 method accurately covering dispersion and H-bonding: HIV-1 protease with 22 ligands. *J. Phys. Chem. B* **2010**, *114*, 12666–12678.
- (18) Pecina, A.; Brynda, J.; Vrzal, L.; Gnanasekaran, R.; Hořejší, M.; Eryilmaz, S. M.; Řezáč, J.; Lepšík, M.; Řezáčová, P.; Hobza, P.; Majer, P.; Veveřka, V.; Fanfrlík, J. Ranking Power of the SQM/COSMO Scoring Function on Carbonic Anhydrase II-Inhibitor Complexes. *ChemPhysChem* **2018**, *19*, 873–879.
- (19) Mikulskis, P.; Genheden, S.; Wichmann, K.; Ryde, U. A semiempirical approach to ligand-binding affinities: Dependence on the Hamiltonian and corrections. *J. Comput. Chem.* **2012**, *33*, 1179–1189.
- (20) González, R.; Suárez, C. F.; Bohórquez, H. J.; Patarroyo, M. A.; Patarroyo, M. E. Semi-empirical quantum evaluation of peptide-MHC class II binding. *Chem. Phys. Lett.* **2017**, *668*, 29–34.
- (21) Kamel, K.; Kolinski, A. Assessment of the free binding energy of 1,25-dihydroxyvitamin D3 and its analogs with the human VDR receptor model. *Acta Biochim. Pol.* **2012**, *59*, 653–660.
- (22) Wichapong, K.; Rohe, A.; Platzer, C.; Slyko, I.; Erdmann, F.; Schmidt, M.; Sippl, W. Application of docking and QM/MM-GBSA rescoring to screen for novel Myt1 kinase inhibitors. *J. Chem. Inf. Model.* **2014**, *54*, 881–893.
- (23) Natesan, S.; Subramaniam, R.; Bergeron, C.; Balaz, S. Binding affinity prediction for ligands and receptors forming tautomers and ionization species: inhibition of mitogen-activated protein kinase-activated protein kinase 2 (MK2). *J. Med. Chem.* **2012**, *55*, 2035–2047.
- (24) Alves, C. N.; Martí, S.; Castillo, R.; Andres, J.; Moliner, V.; Tunon, I.; Silla, E. A Quantum Mechanics/Molecular Mechanics Study of the Protein-Ligand Interaction for Inhibitors of HIV-1 Integrase. *Chem. – Eur. J.* **2007**, *13*, 7715–7724.
- (25) Reddy, M. R.; Erion, M. D. Relative binding affinities of fructose-1, 6-bisphosphatase inhibitors calculated using a quantum mechanics-based free energy perturbation method. *J. Am. Chem. Soc.* **2007**, *129*, 9296–9297.
- (26) Mikulskis, P.; Cioloboc, D.; Andrejić, M.; Khare, S.; Brorsson, J.; Genheden, S.; Mata, R. A.; Söderhjelm, P.; Ryde, U. Free-energy perturbation and quantum mechanical study of SAMPL4 octa-acid host-guest binding energies. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 375–400.
- (27) Rathore, R.; Reddy, R. N.; Kondapi, A.; Reddanna, P.; Reddy, M. R. Use of quantum mechanics/molecular mechanics-based FEP method for calculating relative binding affinities of FBPase inhibitors for type-2 diabetes. *Theor. Chem. Acc.* **2012**, *131*, No. 1096.
- (28) Świderek, K.; Martí, S.; Moliner, V. Theoretical studies of HIV-1 reverse transcriptase inhibition. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12614–12624.
- (29) Beierlein, F. R.; Michel, J.; Essex, J. W. A simple QM/MM approach for capturing polarization effects in protein-ligand binding free energy calculations. *J. Phys. Chem. B* **2011**, *115*, 4911–4926.
- (30) Woods, C. J.; Shaw, K. E.; Mulholland, A. J. Combined quantum mechanics/molecular mechanics (QM/MM) simulations for protein-ligand complexes: free energies of binding of water molecules in influenza neuraminidase. *J. Phys. Chem. B* **2015**, *119*, 997–1001.
- (31) Genheden, S.; Ryde, U.; Söderhjelm, P. Binding affinities by alchemical perturbation using QM/MM with a large QM system and polarizable MM model. *J. Comput. Chem.* **2015**, *36*, 2114–2124.
- (32) Olsson, M. A.; Söderhjelm, P.; Ryde, U. Converging ligand-binding free energies obtained with free-energy perturbations at the quantum mechanical level. *J. Comput. Chem.* **2016**, *37*, 1589–1600.
- (33) Cournia, Z.; Allen, B.; Sherman, W. Relative binding free energy calculations in drug discovery: recent advances and practical considerations. *J. Chem. Inf. Model.* **2017**, *57*, 2911–2937.
- (34) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897.
- (35) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69–82.
- (36) Brandsdal, B. O.; Österberg, F.; Almlöf, M.; Feierberg, I.; Luzhkov, V. B.; Åqvist, J. Free Energy Calculations and Ligand Binding. *Adv. Protein Chem.* **2003**, *66*, 123–158.

- (37) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.
- (38) Ribeiro, R. F.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Prediction of SAMPL2 aqueous solvation free energies and tautomeric ratios using the SM8, SM8AD, and SMD solvation models. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 317–333.
- (39) Miyamoto, S.; Kollman, P. A. Absolute and relative binding free energy calculations of the interaction of biotin and its analogs with streptavidin using molecular dynamics/free energy perturbation approaches. *Proteins: Struct., Funct., Bioinf.* **1993**, *16*, 226–245.
- (40) Mishra, S. K.; Koča, J. Assessing the Performance of MM/PBSA, MM/GBSA, and QM-MM/GBSA Approaches on Protein/Carbohydrate Complexes: Effect of Implicit Solvent Models, QM Methods, and Entropic Contributions. *J. Phys. Chem. B* **2018**, *122*, 8113–8121.
- (41) Su, P.-C.; Tsai, C.-C.; Mehboob, S.; Hevener, K. E.; Johnson, M. E. Comparison of radii sets, entropy, QM methods, and sampling on MM-PBSA, MM-GBSA, and QM/MM-GBSA ligand binding energies of *F. tularensis* enoyl-ACP reductase (F abl). *J. Comput. Chem.* **2015**, *36*, 1859–1873.
- (42) Khandelwal, A.; Lukacova, V.; Comez, D.; Kroll, D. M.; Raha, S.; Balaz, S. A combination of docking, QM/MM methods, and MD simulation for binding affinity estimation of metalloprotein ligands. *J. Med. Chem.* **2005**, *48*, 5437–5447.
- (43) Xiang, M.; Lin, Y.; He, G.; Chen, L.; Yang, M.; Yang, S.; Mo, Y. Correlation between biological activity and binding energy in systems of integrin with cyclic RGD-containing binders: a QM/MM molecular dynamics study. *J. Mol. Model.* **2012**, *18*, 4917–4927.
- (44) Cave-Ayland, C.; Skylaris, C.-K.; Essex, J. W. Direct validation of the single step classical to quantum free energy perturbation. *J. Phys. Chem. B* **2015**, *119*, 1017–1025.
- (45) Rao, L.; Zhang, I. Y.; Guo, W.; Feng, L.; Meggers, E.; Xu, X. Nonfitting protein-ligand interaction scoring function based on first-principles theoretical chemistry methods: Development and application on kinase inhibitors. *J. Comput. Chem.* **2013**, *34*, 1636–1646.
- (46) Fanfrlík, J.; Ruiz, F. X.; Kadlčíková, A.; Řezáč, J.; Cousido-Siah, A.; Mitschler, A.; Haldar, S.; Lepšík, M.; Kolář, M. H.; Majer, P.; Podjarny, A. D.; Hobza, P. The Effect of Halogen-to-Hydrogen Bond Substitution on Human Aldose Reductase Inhibition. *ACS Chem. Biol.* **2015**, *10*, 1637–1642.
- (47) Blomberg, M. R.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. Quantum chemical studies of mechanisms for metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601–3658.
- (48) Siegbahn, P. E.; Himo, F. Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. *JBIC, J. Biol. Inorg. Chem.* **2009**, *14*, 643–651.
- (49) Sumowski, C. V.; Ochsenfeld, C. A convergence study of QM/MM isomerization energies with the selected size of the QM region for peptidic systems. *J. Phys. Chem. A* **2009**, *113*, 11734–11741.
- (50) Hu, L.; Eliasson, J.; Heimdal, J.; Ryde, U. Do quantum mechanical energies calculated for small models of protein-active sites converge? *J. Phys. Chem. A* **2009**, *113*, 11793–11800.
- (51) De Jong, W. A.; Bylaska, E.; Govind, N.; Janssen, C. L.; Kowalski, K.; Müller, T.; Nielsen, I. M.; Van Dam, H. J.; Veryazov, V.; Lindh, R. Utilizing high performance computing for chemistry: Parallel computational chemistry. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6896–6920.
- (52) Jia, W.; Wang, J.; Chi, X.; Wang, L. W. GPU implementation of the linear scaling three dimensional fragment method for large scale electronic structure calculations. *Comput. Phys. Commun.* **2017**, *211*, 8–15.
- (53) Olson, M. A. A-chain structural determinant for binding substrate analogues: A molecular dynamics simulation analysis. *Proteins: Struct., Funct., Bioinf.* **1997**, *27*, 80–95.
- (54) Olson, M. A.; Cuff, L. Free energy determinants of binding the rRNA substrate and small ligands to ricin A-chain. *Biophys. J.* **1999**, *76*, 28–39.
- (55) Yan, X.; Day, P.; Hollis, T.; Monzingo, A. F.; Schelp, E.; Robertus, J. D.; Milne, G.; Wang, S. Recognition and interaction of small rings with the ricin A-chain binding site. *Proteins: Struct., Funct., Bioinf.* **1998**, *31*, 33–41.
- (56) Saito, R.; Pruet, J. M.; Manzano, L. A.; Jascheway, K.; Monzingo, A. F.; Wiget, P. A.; Kamat, I.; Anslyn, E. V.; Robertus, J. D. Peptide-conjugated pterins as inhibitors of ricin toxin A. *J. Med. Chem.* **2013**, *S6*, 320–329.
- (57) Pruet, J. M.; Saito, R.; Manzano, L. A.; Jascheway, K. R.; Wiget, P. A.; Kamat, I.; Anslyn, E. V.; Robertus, J. D. Optimized 5-membered heterocycle-linked pterins for the inhibition of Ricin Toxin A. *ACS Med. Chem. Lett.* **2012**, *3*, 588–591.
- (58) Wiget, P. A.; Manzano, L. A.; Pruet, J. M.; Gao, G.; Saito, R.; Monzingo, A. F.; Jascheway, K. R.; Robertus, J. D.; Anslyn, E. V. Sulfur incorporation generally improves Ricin inhibition in pterin-appended glycine-phenylalanine dipeptide mimics. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 6799–6804.
- (59) Pruet, J. M.; Jascheway, K. R.; Manzano, L. A.; Bai, Y.; Anslyn, E. V.; Robertus, J. D. 7-Substituted pterins provide a new direction for ricin A chain inhibitors. *Eur. J. Med. Chem.* **2011**, *46*, 3608–3615.
- (60) Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. Rm1: A reparameterization of am1 for h, c, n, o, p, s, f, cl, br, and i. *J. Comput. Chem.* **2006**, *27*, 1101–1111.
- (61) Stewart, J. J. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (62) Korth, M. Third-generation hydrogen-bonding corrections for semiempirical QM methods and force fields. *J. Chem. Theory Comput.* **2010**, *6*, 3808–3816.
- (63) Stewart, J. J. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. *J. Mol. Model.* **2013**, *19*, 1–32.
- (64) Stewart, J. J. Application of localized molecular orbitals to the solution of semiempirical self-consistent field equations. *Int. J. Quantum Chem.* **1996**, *58*, 133–146.
- (65) Stewart, J. J. MOPAC: a semiempirical molecular orbital program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1–103.
- (66) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Mol. Struct.: THEOCHEM* **1999**, *461*–462, 1–21.
- (67) Chung, L. W.; Sameera, W. M.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H. B.; Ding, L.; Morokuma, K. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678–5796.
- (68) Chagas, M. A.; Pereira, E. S.; Godinho, M. P.; Da Silva, J. C. S.; Rocha, W. R. Base Mechanism to the Hydrolysis of Phosphate Triester Promoted by the Cd<sup>2+</sup>/Cd<sup>2+</sup> Active site of Phosphotriesterase: A Computational Study. *Inorg. Chem.* **2018**, *57*, 5888–5902.
- (69) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. et al. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.
- (70) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785.
- (71) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (72) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- (73) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

- (74) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- (75) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650–654.
- (76) Franch, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.
- (77) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.
- (78) Grela, P.; Szajwaj, M.; Horbowicz-Drozdzał, P.; Tchórzewski, M. How ricin damages the ribosome. *Toxins* **2019**, *11*, No. 241.
- (79) Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual density functional theory. *Chem. Rev.* **2003**, *103*, 1793–1873.
- (80) Pearson, R. G. Recent advances in the concept of hard and soft acids and bases. *J. Chem. Educ.* **1987**, *64*, No. 561.
- (81) Torrent-Sucarrat, M.; De Proft, F.; Geerlings, P.; Ayers, P. W. Do the local softness and hardness indicate the softest and hardest regions of a molecule? *Chem. – Eur. J.* **2008**, *14*, 8652–8660.
- (82) Sánchez-Márquez, J.; Zorrilla, D.; García, V.; Fernández, M. Introducing a new methodology for the calculation of local philicity and multiphilic descriptor: an alternative to the finite difference approximation. *Mol. Phys.* **2018**, *116*, 1737–1748.
- (83) Sarkar, U.; Roy, D.; Chattaraj, P.; Parthasarathi, R.; Padmanabhan, J.; Subramanian, V. A conceptual DFT approach towards analysing toxicity. *J. Chem. Sci.* **2005**, *117*, 599–612.
- (84) Grillo, I. B.; Urquiza-Carvalho, G. A.; Chaves, E. J. F.; Rocha, G. B. Semiempirical methods do Fukui functions: Unlocking a modeling framework for biosystems. *J. Comput. Chem.* **2020**, 862.
- (85) Merz, K. M.; Faver, J. Utility of the hard/soft acid-base principle via the fukui function in biological systems. *J. Chem. Theory Comput.* **2010**, *6*, 548–559.
- (86) Grillo, I. B.; Urquiza-Carvalho, G.; Bachega, J. F. R.; Rocha, G. B. Elucidating Enzymatic Catalysis using Fast Quantum Chemical Descriptors. *J. Chem. Inf. Model.* **2020**, 578.
- (87) Grillo, I. B.; Bachega, J. F. R.; Timmers, L. F. S.; Caceres, R. A.; de Souza, O. N.; Field, M. J.; Rocha, G. B. Theoretical characterization of the shikimate 5-dehydrogenase reaction from *Mycobacterium tuberculosis* by hybrid QC/MM simulations and quantum chemical descriptors. *J. Mol. Model.* **2020**, *26*, No. 297.
- (88) Fukushima, K.; Wada, M.; Sakurai, M. An insight into the general relationship between the three dimensional structures of enzymes and their electronic wave functions: Implication for the prediction of functional sites of enzymes. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 1940–1954.
- (89) Torrent-Sucarrat, M.; De Proft, F.; Ayers, P. W.; Geerlings, P. On the applicability of local softness and hardness. *Phys. Chem. Chem. Phys.* **2010**, *12*, 1072–1080.
- (90) Grillo, I. B.; Urquiza-Carvalho, G. A.; Rocha, G. B. PRIMORDiA: A Software to Explore Reactivity and Electronic Structure in Large Biomolecules. *J. Chem. Inf. Model.* **2020**, 5885.
- (91) Gilson, M. K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res.* **2016**, *44*, D1045–D1053.
- (92) Yilmazer, N. D.; Korth, M. Enhanced semiempirical QM methods for biomolecular interactions. *Comput. Struct. Biotechnol. J.* **2015**, *13*, 169–175.
- (93) Sulimov, A. V.; Kutov, D. C.; Katkova, E. V.; Sulimov, V. B. Combined docking with classical force field and quantum chemical semiempirical method PM7. *Adv. Bioinf.* **2017**, *2017*, No. 7167691.
- (94) Carra, J. H.; McHugh, C. A.; Mulligan, S.; Machiesky, L. A. M.; Soares, A. S.; Millard, C. B. Fragment-based identification of determinants of conformational and spectroscopic change at the ricin active site. *BMC Struct. Biol.* **2007**, *7*, No. 72.
- (95) Gupta, A.; Chaudhary, N.; Kakularam, K. R.; Pallu, R.; Polamarasetty, A. The augmenting effects of desolvation and conformational energy terms on the predictions of docking programs against mPGES-1. *PLoS One* **2015**, *10*, No. e0134472.
- (96) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.