MOLECULAR FRACTIONATION WITH CONJUGATE CAPS STUDY OF THE INTERACTION OF THE ANACARDIC ACID WITH THE ACTIVE SITE OF TRYPANOSOMA CRUZI GAPDH ENZYME: A QUANTUM INVESTIGATION

MÁRCIA MACHADO MARINHO1*, RICARDO PIRES DOS SANTOS2, EVELINE MATIAS BEZERRA3, RONER FERREIRA COSTA4, CIRO SIQUEIRA FIGUEIRA5, ALICE MARIA COSTA MARTINS6, PEDRO LIMA NETO6, EMMANUEL SILVA MARINHO7, VALDER NOGUEIRA FREIRE8, EUDENILSON LINS ALBUQUERQUE5

1Department of Clinical and Toxicological Analysis, Federal University of Ceará, Ceará, Brazil. 2Department of Computer Engineering/Biotechnology, Federal University of Ceará, Ceará, Brazil. 3Department of Exact and Natural Sciences, Federal Rural University of Semiárid (UFERSA), Mossoró, Rio Grande do Norte, Brazil. 4Department of Analytical Chemistry and Physical-Chemistry, Federal University of Ceará, Ceará, Brazil. 5Dom Aureliano Matos Faculty of Philosophy, Ceará State University, Ceará, Brazil. 6Department of Physics, Federal University of Ceará, Ceará, Brazil. 7Department of Biophysics and Pharmacology, Federal University of Rio Grande do Norte, Rio Grande do Norte, Brazil. Email: marinho.marcia@gmail.com

Received: 27 September 2019, Revised and Accepted: 09 November 2019

ABSTRACT

Objective: The objective of this study was to use the molecular fractionation with conjugate caps (MFCC) method to elucidate the possible interaction mechanism of anacardic acid (AA) with the saturated alkyl chain (AA0) in the Trypanosoma cruzi glyceraldehyde-3-phosphate dehydrogenase (TcGAPDH) enzyme.

Methods: Initially, the geometry optimization of the AA three-dimensional structure (with the pentadecyl chain) was performed using density functional theory (B3LYP) calculations. With the AA0 optimization data, it was possible to plot the molecular electrostatic potential (MESP) surface. Molecular docking simulation was performed using automated coupling with the AutoDock Vina program. The best-fit conformation in the docking simulation of AA0 is the binding site used for the construction of the TcGAPDH-AA0 complex. Interaction energies between the AA0 molecule and the amino acid residues of the TcGAPDH enzyme were estimated using the MFCC strategy.

Results: To obtain more reliable quantitative information on the interaction of AA with the active site of the TcGAPDH enzyme, the fragmentation method was combined with conjugated layers (MFCC) and molecular docking. It can be observed that the AA0 molecule occupies a region near the active site of the chalepin molecule in the TcGAPDH enzyme, and the Ile13 residue has the strongest binding energy of approximately 25 kcal/mol with AA0, through a strong Van der Waals interaction.

Conclusion: The paper presents an improved quantitative analysis approach for assessing the contribution of individual amino acids to the free energy of interaction between AA and TcGAPDH. Specifically, the paper illustrates the advantageous approach of combining molecular docking with the MFCC method.

Keywords: Anacardic acid, Chagas disease, Density functional theory, Glyceraldehyde-3-phosphate dehydrogenase, Trypanosoma cruzi, Molecular electrostatic potential, Molecular fractionation with conjugate caps, molecular docking.

INTRODUCTION

Caused by the parasite protozoan hemoflagellate Trypanosoma cruzi (TC), Chagas disease is commonly transmitted to humans by the insect vector Triatoma infestans [1]. This is an endemic disease in tropical and subtropical regions of Latin America, Africa, and Asia, affecting millions of people each year [2]. Through immigration, the disease has also affected non-endemic regions, such as the United States [3,4]. Benznidazole and nifurtimox, the drugs used to treat Chagas disease, are not effective against certain forms of the disease, in addition to having significant side effects [5,6]. Therefore, it is important to develop new drugs so these problems can be minimized or eliminated.

The mechanism of action of trypanocidal drugs is generally not well elucidated. Many published studies assume that the drug produces free radicals to which TC is particularly sensitive [7], but some of the proposed mechanism of action have been challenged [8]. The inhibition of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from TC (TC glyceraldehyde-3-phosphate dehydrogenase [TcGAPDH]), on the other hand, was also proposed as an explanation for the trypanocidal effect of drugs [9-11]. This enzyme is responsible, in the glycolytic pathway of the parasite, for the conversion of glyceraldehyde-3-phosphate into 1,3-bisphosphoglycerate in the presence of nicotinamide-adenine-dinucleotide (NAD) inorganic phosphate [12,13]. Through anaerobic glycolysis in the cytoplasm, its inhibition can modify the energetic metabolism and the production of adenosine triphosphate, important for the parasite’s life cycle [14,15]. In addition to its primordial role in glycolysis in the cytoplasm, GAPDH has also been associated with several non-metabolic events, such as the onset of apoptosis, axoplasmic transport, activation of transcription, and transport of ER to Golgi [16,17]. Thus, this enzyme can be an important target for the rational planning of new drugs.

Various substances extracted from medicinal plants have been tested for TcGAPDH inhibition [18-20]. For instance, chalepin, a furanocoumarin isolated from species of the Rutaceae family, showed a high inhibitory effect on TcGAPDH [21,22]. Another important natural product inhibitor of TcGAPDH is the 2-hydroxy-6-pentadecyl-benzoic acid, also known as anacardic acid (AA) [23]. It is found in the shells of cashew nuts, the Anacardium occidentale, a plant native to the Northwestern United States.
Brazillian coast, and also found in Central America and countries such as India, Asia, and Africa [24]. Several biological activities of AA have been reported: Antimicrobial, antifungal, antioxidant, antitumoral, antiparasitic, molluscidal, larvicide, and insecticidal [25-29]. The AA is a phenolic lipid, having the structure of a salicylic acid substituted by an alkyl chain with 15 carbons. The alkyl chain may be saturated or unsaturated (monoene, diene, or triene) [30]. All forms of AA (saturated or unsaturated alkyl chain) show an inhibitory effect on TcGAPDH; however, the form with the saturated alkyl chain shows the greatest effects [31]. Because of its high inhibition capacity, the AA has been studied as a promising trypanocidal drug with 1C₅₀ of 28 μM [23]. TcGAPDH has not yet been cocrystalized with AA, and details of its mechanism of interaction with the enzyme site of action are not well elucidated.

Computational chemistry has been widely used for the planning of biologically active compounds and the understanding of their mechanisms of action [32,33]. Density functional theory (DFT) is a quantum computational method that allows the determination (with acceptable accuracy) of various drug properties, such as geometric configurations of the lowest energy, expected positions of the constituent atoms, electronic charge distributions, and vibrational frequencies [34,35]. Changes in the geometric configuration and charge distribution of a drug molecule caused by its interaction with the active site of an enzyme play a key role in its biological activity [36,37]. Molecular docking is a computational method usually based on force fields of classical molecular mechanics that predict the mode of binding of a small-molecule ligand to the active site of a macromolecular ligand to form a stable complex [38], this method is used in the search for new ligands for drug therapy because of their low computational cost when compared to methods based on quantum mechanics [39]. However, the typically used docking methods do not correctly predict the binding energy of a ligand-receptor complex [40,41]; moreover, the molecular fractionation with conjugate caps (MFCC) is a more powerful approach using quantum mechanics calculations to determine interaction energies for protein-ligand systems [42,43]. In this method, the interaction energy is calculated by summing the interactions between the individually capped protein fragments (individual amino acid-based fragments) and the ligands [44]. The MFCC combined with the docking approach can provide more reliable quantitative information on the interaction of drugs with the active site of macromolecular targets.

Devipriya and Kumara have [45] performed molecular docking and the DFT approach to show changes in the conformation and charge density of AA molecules before and after interaction with the active site of the p300 enzyme. Da Silva et al. [46] and Freitas et al. [31] studied the interaction of AA and analogs with TcGAPDH through molecular docking and biochemical assays. However, a more detailed analysis of this interaction (using the MFCC method) together with conformational and charge density changes of the ligand molecule (before and after docking) have not been conducted. To better elucidate the possible interaction mechanism of AA with the saturated alkyl chain (AA0) in the TcGAPDH enzyme, molecular docking and MFCC analyzes of the AA0 molecule were performed at the active site of the enzyme. Results were compared with the binding energy (calculated by MFCC) of NAD+ cofactor crystalized with TcGAPDH. From this perspective, the present work aims using the MFCC quantum method to measure the interaction of AA (with the saturated alkyl chain) and the TcGAPDH enzyme, identifying the component residues of their binding sites, as well as assessing the peculiarities surrounding the drug-receptor complexes, thus elucidating the interaction mechanism.

METHODS

Geometry optimization and molecular electrostatic potential (MESP) surface

Initially, the geometry optimization of the AA three-dimensional (3D) structure (with the pentadecyl chain) obtained from the ZINC database (ZINC code: 8234355) [47] was performed using DFT (B3LYP) calculations. The Gaussian 03 program (Gaussian, Inc.) was used, adopting the B3LYP hybrid functional exchange-correlation and the 6–311+G (d, p) basis set to expand the electronic states. Structural convergence of isolated neutral AA0 (gas phase) was reached when the following thresholds were met: Maximum force per atom <1.5×10⁻² Ha Å⁻¹ and RMS force <1.0×10⁻⁶ Ha Å⁻¹; self-consistent field energy variation <10⁻⁶ Ha, and maximum atomic displacement <6×10⁻⁵ Å. The lowest energy configuration (optimized) was represented as AA0. With the AA0 optimization data, it was possible to plot the MESP surface.

Molecular docking

Molecular docking simulation was performed using automated coupling with the AutoDock Vina program (Molecular Graphics Lab) [48]. The crystallographic structure of TcGAPDH complexed with chalpepin (TcGAPDH-chalpepin) was obtained from RCSB protein data bank (PDB) (code: 1K3T) [49]. The ligand used was the optimized AA0, and the receptor was the chain C of the TcGAPDH enzyme after the removal of the chalpepin of the complex. A box region with 820 nm² centered on and encompassing the active site was defined as the receptor search volume. The best-fit conformation in the docking simulation of AA0 is the binding site used for the construction of the TcGAPDH-AA0 complex. To compare the position of the ligand at the binding site, the Chain A of the TcGAPDH-NAD+ cofactor (PDB code: 1QSX) complexes were used [10]. Molecular graphics and docking were displayed using the UCSF Chimera 1.8 package [50] and Discovery Studio Visualizer.

MFCC

To estimate the total energy of interaction, all 359 amino acid residues of the C chain of the enzyme of the TcGAPDH-AA0 complex were taken into account, seeking to disregard not important interactions. For each amino acid residue, sets of four amino acids were used as caps, two on each side, to increase accuracy when measuring the interaction energy. Interaction energies between the AA0 molecule and the amino acid residues of the TcGAPDH enzyme were estimated using the MFCC strategy [36,45]. The binding energy between the inhibitor molecule AA0 and the amino acid residue Rᵢ is given by E[AA-Rᵢ]:

\[ E(\text{AA-R}) = E(\text{AA-C}) - E(\text{C-R}) - E(\text{AA-C}) + E(\text{C}) \]

where the C (cap) is obtained by attaching a carboxyl or amine group to the dangling bond of the residue R. On the right side of the equation, \( E(\text{AA-C-R}) \) is the total energy of the system formed by the AA molecule and the capped residue; the \( E(\text{C-R}) \) term is the total energy of the blocked residue alone; and E (AA-C-C) is the total energy of the system formed by the AA and the covers alone. To conclude, \( E(\text{C}) \) is the total energy of the system formed only by the molecular caps. For the first-principles calculations, the methodology seen in early works [44,51], the positions of the non-hydrogen atoms were kept fixed while the positions of the hydrogen atom and AA were in principle optimized using the CHARMM force field [52], which has parameters specific for amino acids. Simulations were performed within the formalism of the DFT, using local density approximation for the correlation-change functional with a dispersion correction scheme [DFT/LDA/0BS] was carried out using the Dmol3 code [53,54]. To expand the Kohn-Sham electron orbitals, a set of numerical bases and double polarization (dynamic nuclear polarization [DNP]) was selected, considering all electrons, explicitly, and with restricted spin. The DNP basis set has an accuracy equivalent to the 6–311+G (3df, 2pd) Gaussian basis set [53,55], with negligible basis set superposition error. The orbital cutoff radius for this basis set was set to 3.7 Å and the self-consistent field geometry threshold to 10⁻⁶ Ha. The binding energy of chain A of TcGAPDH-NAD+ cofactor complexes was also calculated (by MFCC) to be compared with the TcGAPDH-AA0 complex.

RESULTS AND DISCUSSION

The MESP surface

The MESP surface allows to visualize the relative polarity of the compounds [56] and to predict the binding site between biological
molecules and their receptors [57,58], as well as being an important tool in the study of new drugs [59,60]. Associated with the dipole moment, the electronegativity and partial charge of the MESP [61], can be used to predict the types of intermolecular interaction, the molecule may carry as well as the sites of these interactions.

Given the electrostatic potential energy between a hypothetically positively charged ion (+1) located at a point (x, y, and z) and the molecule, the electrostatic potential at x, y, and z is calculated. The ion repulsion means the potential is positive, being the ion attracted by the molecule, then the potential is negative. Thus, electron-poor regions usually have positive potentials and electron-rich regions usually have negative potentials, the MEP at a point r in the space around a molecule (in atomic units) can be expressed as follows [62]:

\[
V(r) = \sum_{A} Z_{A} \rho_{A}(r) - \frac{1}{4\pi} \int \frac{\rho(r')}{r} dr'
\]

where \( Z_{A} \) is the charge on nucleus \( A \), located at \( R_{A} \) and \( \rho(r') \) is the electronic density function for the molecule. \( V(r) \) is the resulting electrostatic net effect produced at the point \( r \) by both the electrons and the cores of the molecule, where the first term represents the contributions due to the potential for electrons and the second term due to the cores. The knowledge of the charge distribution, identifying their partial densities, helps in understanding the form of interaction between a molecule and another, and thus, one can identify the molecular reactive site [63,64].

AA (15:0) has in its structure a nucleus derived from salicylic acid substituted with an alkyl chain composed of 15 carbon atoms, linked only by sigma bonds (Fig. 1), the electrostatic potential map is shown in Fig. 2, the 3D charge distributions of the molecule, while negative electrostatic potentials (with intensity proportional to the absolute value of the potential energy) are shown in red, positive electrostatic potentials are visualized in blue and green, areas with potentials close to zero are perceived.

A negative charge partial density in the carboxyl group indicates a region rich in electrons, this fact can be explained by the presence of pairs of non-binding electrons in oxygen, and by the difference of electronegativity in relation to the carbon atom, being indicative of a nucleophilic site. While in the alkyl group, a potential close to zero was observed, characterizing a region predominantly not polar.

**Molecular docking and MFCC**

The best-fit conformation in the docking simulation of AA0 is shown in Fig. 3, together with the chalepin docking configuration (obtained from crystallographic data of the TcGAPDH complex – chalepin, PDB 1K3T) for comparison. It is possible to observe that the AA0 molecule occupies a region near the active site of the chalepin molecule in the TcGAPDH enzyme. The AA0 can interact with a large number of residues due to its long chain. In addition, it consists of a polar head with a benzene ring and a long hydrophobic tail. The polar head favors the formation of hydrogen bonds, while the tail can interact with amino acids that have a hydrophobic character.

Fig. 4 shows the molecular two-dimensional interaction map for AA0 and TcGAPDH complex. As it can be seen, there are many kinds of interaction, since weak interactions by London dispersion forces (pale pink), sigma-pi interactions (lilac), pi-pi stacking (darker pink), and hydrogen bonding (bright green) are found. Those interactions are in agreement with AA0/chalepin docking with TcGAPDH complex. The TcGAPDH-AA0 complex (with the AA0 better conformation in the docking simulation) was used in the MFCC calculations.

Based on the DFT, the interaction energy of the amino acid residues was determined, through computations using the MFCC approach. The amino acids that showed significant interaction energy for stabilizing the AA0 molecule were: Ile13; Met16; Ser134,224; Ala35,164,228; Pro136; Cys165; Thr167,197,226; His194; Gly227; Arg249; Asp334; Asn335; Glu336; and Tyr339. The binding site, interaction energy, and residues domain (BIRD) panel (Fig. 5) show a graphical display with interaction energies between ligand and the most important amino acid residues comprising the binding site of the TcGAPDH enzyme. This panel shows relevant residues interacting with AA contributing to its energy stabilization at the binding site of the gGAPDH enzyme.
This panel is a new version of BIRD [43] and shows: (i) The interaction energy of each residue with the AA0 molecule using horizontal bars, based on which can quantitatively estimate the role of each residue in the binding site, i.e., its efficacy to attract or repel the drug; (ii) the most significant residues for the bond on the left side; (iii) the atoms that show the shortest distance between the amino acid residue and the ligand; (iv) the left side panel containing the minimum distance in angstroms between amino acid ligands; and (v) the bars indicating the interaction strength between ligand and residues of TcGAPDH show colors that depend on the side chain of each amino acid score.

It is possible to see in Fig. 5 the attached BIRD that 22 residues contribute to the stabilization of AA0: Arg12, Ile13, Arg15, Met16, Asp38, Glu109, Pro136, Ala164, Ser165, Cys166, Thr167, His194, Thr197, Thr226, Ala228, Arg249, Asp254, Asp334, Asn335, Glu336, Tyr339, and Arg342, whereas only one Ala135 amino acid residue has positive (repulsive) interaction energy. Pavão et al. [49] reported that the residues Thr167, Cys166, Arg249, and Asp210 of the gGAPDH enzyme have important interactions with chalepin. The residue Thr167 interacts with the oxygen of the CO group of chalepin through W739 [49]. The Cys166 residue shows a characteristic hydrophobic interaction with the 1,1-dimethylallyl group; a hydrogen interaction with the Arg249 is also reported, and finally, a weak interaction of hydrogen with Asp210 is intermediated by W812 water. Due to the structural differences of the ligands, the above amino acids may or may not show molecular interactions relevant to AA0 stabilization, as it can be observed in the BIRD panel.

Residue Ile13 has the strongest binding energy, approximately 25 kcal/mol with AA0, through a strong Van der Walls interaction involving the AA0 tail and the side chain of Ile13, both with hydrophobic characteristics. Although the hydrophobic interactions are admittedly one of the weakest, it is relatively strong because of the proximity...
between the chemical groups of both binder and residue, which is about 2 Å, as well as the extent of the side chain of the Ile13 residue, which is approximately 4 Å.

The second amino acid with the strongest interaction energy is Glu336, with a binding energy around 15.0 kcal/mol, due to the natural formation of an AA0 electrostatic dipole inside the tail-binding site, with absence of electrons, which favors the interaction with amino acids that have more electronegative atoms or side chains, which is the case of Glu336 that has the COO$^-$ group.

The third most important residue, Tyr339, has binding energy with the AA0 of 14.1 kcal/mol due to a sigma-pi interaction, caused by the presence of an aromatic ring rich in electrons in the side chain of that amino acid, as well as positive characteristic of the AA0 tail. We can also observe that the method used to calculate the interaction energy was able to describe a pi-pi interaction between the aromatic rings of AA0 and the side chain of the His194 residue, with an intensity of about 5.0 kcal/mol. The residues Thr167 and Thr226 interact with AA0 through a hydrogen interaction of about 12.0 kcal/mol and 13.0 kcal/mol, respectively. The hydrogen of the OH group of Thr167 interacts with the oxygen of the CO group of AA0, which are separated by a distance of 1.95 Å, whereas the other hydrogen bond is due to the interaction of the hydroxyl hydrogen of the residue Thr226 and the oxygen of the AA0 hydroxyl group. Together, Ser165 and Cys166 create a region with a positive electrostatic potential, which couples with the region of the AA0 hydrophobic tail that has electronegative characteristics, thus producing a relevant attractive interaction between these amino acids of approximately 13.0 kcal/mol.

CONCLUSION

The paper presents an improved quantitative analysis approach for assessing the contribution of individual amino acids to the free energy of interaction between organic ligands (such as AA) and target protein (like GAPDH). Specifically, the paper illustrates the advantageous approach of combining the molecular docking (such as the mostly used AutoDock Vina program) with the MFCC method. It was obtained more reliable quantitative information on the interaction of AA with the active site of the enzyme TcGAPDH by combining MFCC with molecular docking. It has been verified that the AA0 molecule occupies a region near the active site of the chalasin molecule in the TcGAPDH enzyme and that residue Ile13 has the strongest binding energy, approximately 25 kcal/mol with AA0, through a strong Van der Waals interaction.

ACKNOWLEDGMENTS

Acknowledge the financial support received during the development of this work from the Brazilian National Council (CNPq).

AUTHORS’ CONTRIBUTIONS

The research was developed with: Conceptualization and use of the software by (Prof. Dr. Ricardo Santos and Prof. Dr. Valder Nogueira); Methodology (Msc. Marcia Marinho, Prof. Dr. Ricardo Santos, Dr. Eveline Bezerra, PmF Dr. Roner Costa and Prof. Dr. Emmanuel Marinho); Investigations and resources (Prof. Dr. Ricardo Santos and Msc. Marcia Marinho); Supervision and Project administration (Prof. Dr. Ricardo Santos and Prof. Dr. Eudinhaenlo Albuquerque); Funding Acquisition (Prof. Dr. Ricardo Santos, Prof. Dr. Valder Nogueira, PmF Dr. Alice Martins, prof. Dr. Pedro Lima Neto and Prof. Dr. Eudinhaenlo Albuquerque); and Writing-Original Draft Preparation (Msc. Marcia Marinho, Dr. Eveline Bezerra, Msc. Ciro Figueira, Prof Dr. Roner Costa and prof. Dr. Emmanuel Marinho). All authors had a contribution to the final manuscript.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

FUNDING

This research received financial support from Brazilian National Council (CNPq) projects 151182/2013-2 and 447568/2014-9.
study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

REFERENCES


