Prediction of Protein-Ligand Binding Sites for Cisplatin and Transplatin based on Hydrogen Bonds

Flávio Almeida Curvelo dos Anjos* and Manoel Victor Frutuoso Barrionuevo

1Department of Physics, Institute of Biosciences, Arts and Exact Sciences, Universidade Estadual Paulista, Sao Jose do Rio Preto, Sao Paulo, Brazil
2Department of Chemistry and Environmental Sciences, Universidade Estadual Paulista, Sao Jose do Rio Preto, Sao Paulo, Brazil

Abstract

Platinum compounds are very important to treatment of various malignant tumors. However, prediction of platinum-binding sites is very hard to be made. Nevertheless, hydrolysis of leaving groups bounded to platinum compounds plays an important role in delivering platinum to a target molecule. Herein, a study in silico provides an understanding of the molecular surface in atomic level of three-dimensional structure of cisplatin and transplatin and their binding-sites in order to offer some insights in drug designing. The goal of this work was to implement a new approach based on geometric and physicochemical parameters to find platinum-binding sites using parallel computing algorithms for graphics processing units (GPUs). These algorithms were tested and validated by analysing platinum-binding sites in five known proteins. The results indicated that these binding sites were predicted with significant success. In our analysis HexServer and PatchDock server did not find putative binding-sites for cisplatin and transplatin as we found for the five chosen proteins. Herein, we have shown that the present method have had a better prediction of platinum-binding site than HexServer and PatchDock methods.

Keywords: Binding sites; Hydrogen bonds; Cisplatin; Transplatin; Molecular surface; GPU computing

Abbreviations: 3D: Three-Dimensional; GPUs: Graphics Processing Units; CUDA: Compute Unified Device Architecture; CPU: Central Processing Unit; PDB: Protein Databank; FFT: Fast Fourier Transform

Introduction

Three-dimensional structures (3D) of proteins and ligands reveal a significant geometrical correspondence between their contact regions [1], i.e., shape complementarity and other physicochemical characteristics of protein surfaces determine the nature of interaction between the protein’s binding-site and the ligand [2]. It is reasonable that analysis of protein’s binding-sites are crucial in order to provide a better understanding of their structure and function, which are important parameters to drug discovery field. As binding-sites contains amino acid residues involved in interactions that require a well-defined arrangement [2], a challenge is to identify groups of interaction between protein and its ligand. Furthermore, binding-sites might be influenced by environment properties, such as pH, ionic strength and high ordering of water molecules [3]. Hence, the process of binding-site analysis relies on identifying atoms likely to establish hydrogen bond, which plays a fundamental role in protein ligand interaction [4] and is pointed as a driven force to the higher degree of selectivity between protein and ligand [4]. For this reason, analysis of binding-site takes into account the atom type and its relative position to neighbouring atoms [5]. In addition, molecular surface characterization at atomic level of platinum compounds affords an opportunity for the development of novel and better drugs. For instance, cisplatin and transplatin are widely used anticancer drugs. They are suitable for many types of cancer, such as testicular, ovarian, cervical, colorectal and relapsed lymphoma cancer [6,7]. Interestingly, the trans isomer of cisplatin known as transplatin is less cytotoxic [8], which is argued as a consequence of steric mechanisms that influence the way transplatin interacts with its target molecules [9]. Generally, methods employed to identify protein binding sites use static 3D protein structures to compute their geometry or energy. Examples of methods that use geometric parameters of protein shape are: CAST [10], PASS [11], PocketPicker [12], and Fpocket [13]. On the other hand, methods that compute the energy simply use van der Waals forces to describe protein shape, i.e., Q-SiteFinder [14]. There are methods that use grid algorithm to compute the energy of an interaction as the sum of Lennard-Jones potential, coulombic attractions and directional hydrogen bond energy terms [15,16]. For instance, Cavbase [17] combine structure and sequence comparison in its analysis process by describing and comparing protein binding pockets on the basis of their geometrical and physicochemical properties. On the other side, to analyse a binding-site, a set of representative conformations of proteins can be obtained from classic molecular dynamics simulations [18-20]. However, static analysis is not always enough to understand the mechanisms of interaction of binding-sites and a representative example of this lack of understanding is cytokine IL-2 [21,22]. Actually, recent advances in parallel computation based on GPUs have improved the computational resources available, allowing GPUs to be used as performance accelerators for a wide variety of scientific applications [23]. In this article, we propose a new approach based on geometric and physicochemical properties to find binding-sites using parallel computing algorithms for GPUs. Herein, cisplatin and transplatin were the chosen ligands for this study in order to analyse how hydrogen bonds drive interactions between protein and ligand. The proposed approach consists in analysing the exposed surface area of each atom of proteins and ligands, as well as its relationship with the point of highest electrical potential on molecular surface. Therefore, to rank the best candidates for binding sites, an empirical scoring function was

*Corresponding author: Flávio Almeida Curvelo dos Anjos, Department of Physics, Institute of Biosciences, Arts and Exact Sciences, Universidade Estadual Paulista, Sao Jose do Rio Preto, Sao Paulo, Brazil, Tel: 0055 17 3221-2240; Fax: 0055 17 3221-2247; E-mail: landnak@gmail.com

Received November 13, 2014; Accepted January 16, 2015; Published January 22, 2015


Copyright: © 2015 Anjos FAC, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited
created. It was made by selecting atoms of highest electrical potential on exposed surface area and adding terms such as lipophilicity [24], which represents the hydrophobic effect [25-27], and geodesic distance between points of highest electrical potential on surface, which exemplifies the geometric complementarity between protein and ligand. A geodesic is a locally length-minimizing curve used for shape analysis, surface evolution and object recognition [28]. This analysis at the atomic level aims to offer new insights towards the understanding of the effectiveness of each atom in establishing hydrogen bonds on molecular surface and inspire development of novel methods to analyse small ligands.

Materials and Methods

Preparation of ligands structures

Cisplatin and transplatin structures were modeled through GaussView/Gaussian09 [29]. These structures were optimized using MPW1MPW91 [30] method with SDD (Stuttgart/Dresden) [31] basis function. Partial charges were calculated using the QTPIE method [32] implemented by using C++ Open Babel library [33]. Cisplatin and transplatin structures are provided in the supplementary information.

Preparation of protein structures

The proteins were selected for the study based on the protein databank [34]. To make sure about the validation of our method we have chosen five proteins that had binding-sites well determined through experimental data and a description of which residues of each binding-site interact with cisplatin and transplatin. For instance, cuprozinc superoxide dismutase (PDB ID: 2AEO) was selected because its crystallization with cisplatin shows an interaction of platinum with His 19 [35], which can be visualized through electron-density map. For chaperone-monomer (PDB ID: 3JWL) the residues Cys 12 and Cys 15 [36] bind platinum, which also can be seen through electron-density map. In the case of myoglobin (PDB ID: 1MBN) the platinum binding-site was found through mass spectrometry technique, and His 116 and His 119 [37] were determined as the residues that bind cisplatin and transplatin. In addition, ubiquitin (PDB ID: IUBQ) had also the platinum binding-sites determined by mass spectrometry. Ubiquitin has two binding-sites in which the residue for the first binding-site is Met 1 and the second is His 68 [38]. Finally, the protein BCL-2 (PDB ID: 1GSM) was chosen because its structure was determined by NMR technique and there was a precise description about the residue that binds to platinum, which is Glu 13 [39]. In this work, all selected proteins have no covalently bound HETATM but only coordinated groups such as cofactors (Cu, Zn or Heme group) or water. In addition, we have added hydrogens in proteins whose PDB were missing hydrogens. That was done by using HADD software [40] and partial charges of all atoms within the proteins were calculated by the same method used to ligand structures.

Mapping molecular surfaces

A software named MSProt based on C++ was developed by using the open source framework Qt version 5.3.2 (http://qt-project.org) with OpenGL (http://opengl.org) in order to map and view molecular surfaces in 3D environment structure. A parallel computing platform called CUDA [41] was used to speed up numerical linear algebra operations. Therefore, to map molecular surface area the skin surface method [42] was used, which was implemented through C++ Computational Geometry Algorithms Library (CGAL), library (http://cgal.org). Kuhn, Schmitt and Klebe have employed a method in which they used vectors to define a set of properties that guide molecular recognition: hydrogen bond acceptor and hydrogen bond donor [17]. Adjustments that were made on Kuhn, Schmitt and Klebe’s method is shown in Figure 1. Details about how every vector was calculated for amino acids and ligands were summarized in the supplementary information. The exposed surface area was calculated by using algorithm 1. Hence, instead of use vector r as proposed by Kuhn, Schmitt and Klebe’s method, a vector p was calculated by taking the point of maximum electric potential on the molecular surface of the region of sphere 3. Further information about this adjustment is provided in Figure 1. To calculate the electric potential at every triangle belonging to the mesh formed by skin surface in sphere 3, the algorithm 2 was used. The sphere 3 represents the region of molecular surface to be coloured according acceptor and donor properties. The function kernelPotenc (explained later on) fixes the exposed area, which can vary between 0-10 Å² (see Figure 6), according to the values of electrical potential calculated through the partial charges. Therefore, the size of the colored area is determined by the electrostatic potential, and the colored area cannot exceed the maximum radius of 2 Å determined by sphere 3 as defined by Kuhn, Schmitt and Klebe’s method. Figure 2 shows an example of mapping of myoglobin. Figure 3 shows an example of mapping of cisplatin.

The geodesic of ligands

Permutation is defined by analytic combinatorics as an ordered sequence without repetition, this concept was used to list a minimum of two points and a maximum of six points as possible sequences, these limits is due to the number of hydrogen bonds that cisplatin and transplatin can establish. Hence, by taking labelled points on ligand surface, sequences from two to six points were made as shown in figure 3. The length of each sequence of points was calculated by geodesic algorithm [43], and the lowest length was stored for further analysis of the distances between the ligand and protein. This was done for each sequence of points generated by permutation.

The geodesic of protein

The same process of permutation described as before was...
Conducted to the points of the protein molecular surface. In order to calculate the shortest path among these points, a virtual sphere was used to select the points that falls within. That virtual sphere represents the ligand (cisplatin or transplatin) whose measured distance between two internal atoms in the optimized ligand structure was measured as 4.94 Å, which was considered to be the diameter. Hence, the used virtual sphere had a radius of 5 Å whose diameter of 10 Å was assumed to be the double of 4.94 Å, i.e., the size of the virtual sphere was enough to roughly mimic the ligands. The points’ selection was made by positioning the center of the virtual sphere at the first referential point of maximum electric potential and all points that dropped within the sphere were collected according figure 4. Thus, it was repeated to each labelled location of maximum electric potential of the protein. The same process of permutation created several sequences, which had not exceeded the maximum number of six elements for the reason described in geodesic of ligands.

**Ranking of binding sites**

The ranking analysis was speeded up by discarding sequences through the employment of four parameters. The first parameter is the geodesic value among points of greatest electrostatic potential between the protein and the ligand. For protein analysis in relation to a chosen ligand, the difference between the paths formed by the geodesic distances was established as 0.4 Å. The second parameter was the perimeter, which was obtained for both the protein and the ligand as the sum of the measured distances of acceptor and donor atoms, respectively. Hence, the difference between the sum of distances among the acceptor atoms of the protein and the sum of distances among the donor atoms of the ligand were calculated and assigned to be lower than 0.2 Å. The third parameter considered was the alignment of these structures. The points of greatest potential were aligned and the root-mean-square deviation (RMSD) [44] was calculated (Figure 5). Herein, the range of 1.1 < RMSD <2.6 was considered satisfactory, whether to determine if the ligand have not superimposed over the protein or if it is in a far region of the binding site. The last parameter was the total acceptor area of the protein and the total donor area of the ligand. In order to calculate the total acceptor area it was summed up the exposed area of each acceptor atom from the sequence that was established by permutation. Thus, the difference between the total acceptor area and total donor area was determined as being lower than 0.4 Å. These parameters are shown in figure 5. Figure 6 shows the relationship between the acceptor exposed area and the angle formed between vectors $v$ and $p$. The present study has considered only the protein’s acceptor atoms and the ligand’s donor atoms to search for binding-sites. Thus, the scoring function was developed based on the relation between acceptor and donor groups found in proteins and ligands respectively. Herein, the goal was to minimize the score function, where the lower the value of the score function the easiest to establish a hydrogen bond, which was in agreement to figure 1. Hence, for ranking of binding-sites it was assumed the following terms to the scoring function:

$$SF=P+ST$$

Figure 2: Mapping molecular surface of the protein myoglobin PDB ID: 1MBN. Donator regions are coloured in dark blue and sums up to a total of 372.9 Å$^2$. Acceptor regions are coloured in red and sums up to a total of 868.6 Å$^2$.

Figure 3: Mapping molecular surface of the cisplatin. The white regions show the points of maximum electric potential. Donor regions are coloured in dark blue. The white triangles represent an example of the geodesic path computed by the algorithm.

Figure 4: Sphere A with 5Å radius represents the virtual sphere, which selects the highest electrical potential points that falls within. Hence, allowing to calculate the shortest geodesic paths. 1-Represents the molecular surface (skin surface).

Figure 5: Discarding sequences parameters.
The term \( P \) represents the effectiveness of each atom in the hydrogen bond on molecular surface.

\[
P = \frac{1}{2n} \sum_{i=1}^{n} \left[ 1 - \frac{A_{\text{max}}}{A_{\text{max}}} c + \frac{\beta_{\text{max}}}{\beta_{\text{max}}} \right]
\]

\( A_{\text{max}} \) corresponds to the maximum exposed area of each atom whose value is 10 Å\(^2\). This value was obtained by our measurements of all exposed areas of the atoms studied (Figure 6). \( A_{\text{max}} \) is the exposed area corresponding to the acceptor atom that interacts with the donor atom whose exposed area is represented by the term \( A_{\text{max}} \). \( \beta_{\text{max}} \) is the maximum angle between the vectors \( v \) and \( p \), which has the value of 100° that was extracted from Kuhn, Schmitt and Klebe’s method. Finally, the sum has \( n \geq 1 \) where \( n \) corresponds to the number of hydrogen bonds. The \( c \) term was prepared to assign a quantitative analysis for the type of each amino acid to be analyzed in relation to its hydrophobicity. This term consists:

\[
c = \frac{V_a}{V_{\text{phe}}} - \frac{L_a}{L_{\text{phe}}}
\]

\( V_a \) is the volume of amino acid divided by its lipophilicity (\( L_a \)). \( L_{\text{phe}} \) values were extracted from table 2 of El Tayar scale [24] of experimental measurements, as well as volume values were obtained from table 1 [24]. Hence, \( V_{\text{phe}} \) and \( L_{\text{phe}} \) stand by volume and lipophilicity of phenylalanine, which was chosen as the parameter of comparison. According to Monera [45], phenylalanine is the most hydrophobic amino acid at pH 7.

The term \( S \) describes the quality of fitting:

\[
S = \frac{Geo_{\text{max}}}{Geo_p} + \frac{A_{\text{max}}}{A_p}
\]

\( Geo_p \) represents the geodesic of protein. \( A_p \) is the total acceptor area of protein. Herein, the value of 8 Å for \( Geo_{\text{max}} \) was the maximum allowed by fitting the ligand. As the total number of allowed hydrogen bonds is six, \( A_{\text{max}} \) assumes the value of 60 Å. That total number of allowed hydrogen bonds corresponds to possible hydrogen bonds of cisplatin or transplatin.

The term \( T \) represents the fit between the protein and the ligand. Ideally, this term is equal to 1, which indicates a good fit.

\[
T = \frac{Geo_p}{Geo_{\text{lig}}}
\]

\( Geo_{\text{lig}} \) is the geodesic of ligand.

**Algorithms for GPUs**

In CUDA’s architecture, Device is the given name to GPUs and Host is the CPU, further information is provided by CUDA C.
Steps of the algorithm 1 in kernelExposeArea

1 - Memory allocation of points from the center of the atom and the vector points \( v \). These data will be shared by each thread of the same block.
2 - Synchronize threads before continuing execution.
3 - Calculate the midpoint representing the center of the triangle, which will be the vector \( b \). The vector \( b \) is only an auxiliary vector for calculating the area of the exposed property.
4 - Check if the triangle is inside the sphere 3 (Figure 1) through the vector \( b \) module. That sphere has a radius of 2 Å, which can be subsequently influenced by the electric potential as described in the mapping of molecular surfaces.
5 - If the vector \( b \) is within the sphere then calculate the angle between \( v \) and \( b \).
6 - If the angle is less than or equal to 100º, the triangle’s vertices are stored in global memory and sent to the Host.

In Host with these selected vertices, the exposed area was calculated. These selected vertices were sent to Device and ran in the kernelExposeArea. Figure 7 shows how the kernelExposeArea was implemented regarding the threads and memory access.

Steps of the algorithm 2 in kernelPotenc

1 - Calculate the midpoint representing the center of a triangle.
2 - Calculate the vector between the triangle’s midpoint and the atom’s centre.
3 - Calculate the magnitude of the vector \( r \).
4 - Read in global memory each atom position and partial charges.
5 - Calculate the electric potential at the midpoint of the triangle by vector \( r \).
6 - Store the electrical potential of the triangle in the global memory to be sent to Host.

The kernelPotenc is performed for all triangles that fall inside of the sphere 3 (Figure 1), and along this calculation the maximum value to form the vector \( p \) is extracted. Figure 8 shows how the kernelPotenc was implemented regarding the threads and memory access. In implementing this work, a computer with 32 GB DDR 3, a core i7 processor 3.40 GHz and a video card Titan Black from Nvidia with 2880 CUDA cores and 6144 MB memory was used.

Compare results with other methods

In order to compare our method to existing methods we have chosen HexServer [46] and PatchDock [47] to run an analysis in which the binding-sites of cisplatin and transplatin were investigated. Protein docking algorithms, such as HexServer and PatchDock, usually assume that the proteins are rigid, though proteins are intrinsically flexible. These algorithms use geometric hashing [48] or fast Fourier transform (FFT) correlation techniques [49] in order to find a reduced number of putative docking orientations that could be enhanced using other techniques. HexServer is a FFT-based protein docking server based in GPUs and PatchDock is based on a rigid-body geometric hashing algorithm. Thus, we ran HexServer and PatchDock using their default conditions and in our analysis and we have compared our method with the ten first results of HexServer and PatchDock.

Results and Discussion

Figure 9 shows the total execution time of kernels. By viewing Figure 9, we note that the kernels execute much faster on GPUs. The number of triangles processed by kernels is shown in Table 1. Given that the binding-sites were found, deletion of HETATM has had no

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>Number of triangles</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1GSM</td>
<td>10466944</td>
<td>1457.64</td>
</tr>
<tr>
<td>1MBN</td>
<td>9422144</td>
<td>1130.57</td>
</tr>
<tr>
<td>2AEO</td>
<td>15059136</td>
<td>3598.5</td>
</tr>
<tr>
<td>1U8Q</td>
<td>4667712</td>
<td>353.12</td>
</tr>
<tr>
<td>3IWL</td>
<td>3426496</td>
<td>227.54</td>
</tr>
</tbody>
</table>

Table 1: Number of triangles of the triangular mesh and execution time for the algorithms on GPUs.
Herein, we showed the applicability of algorithms to determine the binding sites of proteins with platinum compounds. For instance, the PDB of chaperone-monomer (PDB ID: 3IWL) has as HETATM two sulfates, a TCEP ligand (3,3',3''-phosphanetriyltripropanoic acid) and water molecules, and all these HETATM are not covalently bound to chaperone-monomer. For myoglobin (PDB ID: 1MBN) there were water molecules and one heme group, which were not covalently bound to myoglobin and do not interfere in finding the platinum binding site. On the other hand, cuprozinc superoxide dismutase (PDB ID: 2AEO) has metallic cofactors (two of each: Zn and Cu) and water molecules as HETATM. They were removed without disturbing the location of platinum binding-site. Ubiquitin (PDB ID: 1UBQ) had only water molecules as HETATM and their deletion did not bring any concerns to our proposed work. Lastly, BCL-2 (PDB ID 1G5M) had no HETATM on its PDB structure of this protein was defined by NMR technique that provides a high level of resolution.

The next analysed protein was myoglobin [37]. As the literature shows that cisplatin and transplatin can bind to myoglobin, we investigated the binding sites of these compounds. It was seen that transplatin binds at the same region that cisplatin binds in ubiquitin [38].

Table 2: Ranking for each site found. Rank 1-position of the binding site found in relation to the ten best ten candidates for cisplatin. Rank 2-position of the binding site found in relation to the ten best candidates for transplatin.

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>Rank 1</th>
<th>Rank 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1G5M</td>
<td>2</td>
<td>*</td>
</tr>
<tr>
<td>1MBN</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2AEO</td>
<td>3</td>
<td>*</td>
</tr>
<tr>
<td>1UBQ</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3IWL</td>
<td>1</td>
<td>*</td>
</tr>
</tbody>
</table>

*Data not found in the literature about the binding sites of these proteins with transplatin.

The developed method has not found putative binding-sites for cisplatin and transplatin as we found the platinum-binding site on Met 1 and His 68 [38]. The binding-sites were extremely close to Cys 12 and Cys 15 and they were ranked in first place, it was suggested that the binding-site is valid. For BCL-2 protein the analysis were more complicated, because the structure of this protein was defined by NMR technique that provides several models, and only the first model was analysed. According to article [39] there are some potential binding sites for cisplatin such as Glu 13. Indeed, in our method Glu 13 was ranked in second place. Once binding-site for cisplatin was investigated we started to seek for transplatin binding-sites. By analysing how transplatin binds to ubiquitin [38] it was seen that transplatin binds at the same region that cisplatin binds in ubiquitin. In addition, for myoglobin the transplatin binding-site was extremely close to Cys 12 and Cys 15 and they were ranked in first place. The penultimate protein analysed was chaperone-monomer, which the platinum-binding sites are Cys 12 and Cys 15 [36]. Herein, our method found Cys 15 and Lys 60 as platinum-binding sites of chaperone-monomer. As these binding sites were extremely close to Cys 12 and Cys 15 and they were ranked in first place, it was suggested that the binding-site is valid.

Table 3: All binding sites for cisplatin.

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>Binding sites</th>
<th>Acceptor atoms</th>
<th>Donated atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1G5M</td>
<td>Met 16</td>
<td>SD</td>
<td>H5</td>
</tr>
<tr>
<td></td>
<td>Glu 13</td>
<td>O</td>
<td>H2</td>
</tr>
<tr>
<td>1MBN</td>
<td>His 116</td>
<td>O</td>
<td>H1</td>
</tr>
<tr>
<td></td>
<td>His 119</td>
<td>O</td>
<td>H3</td>
</tr>
<tr>
<td></td>
<td>Pro 120</td>
<td>O</td>
<td>H4</td>
</tr>
<tr>
<td>1UBQ</td>
<td>Ile 44</td>
<td>O</td>
<td>H3</td>
</tr>
<tr>
<td></td>
<td>His 68</td>
<td>NE2</td>
<td>H1</td>
</tr>
<tr>
<td>2AEO</td>
<td>Thr 30</td>
<td>OG1</td>
<td>H6</td>
</tr>
<tr>
<td></td>
<td>His 19</td>
<td>NE2</td>
<td>H3</td>
</tr>
<tr>
<td>3IWL</td>
<td>Lys 60</td>
<td>NZ</td>
<td>H3</td>
</tr>
<tr>
<td></td>
<td>Cys 15</td>
<td>SG</td>
<td>H4</td>
</tr>
</tbody>
</table>

Table 4: All binding sites for transplatin.

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>Binding sites</th>
<th>Acceptor atoms</th>
<th>Donated atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1UBQ</td>
<td>Ile 44</td>
<td>O</td>
<td>H6</td>
</tr>
<tr>
<td></td>
<td>His 68</td>
<td>NE2</td>
<td>H4</td>
</tr>
<tr>
<td>1MBN</td>
<td>His 116</td>
<td>NE2</td>
<td>H6</td>
</tr>
<tr>
<td></td>
<td>Gin 128</td>
<td>NE2</td>
<td>H1</td>
</tr>
</tbody>
</table>

Conclusion

In our method we first decide which atoms are important to calculate the vector v, which is used to limit the acceptor and donor areas of each residue. As MSProt is still in development, in future work our method will take into account all hydrogen bonds and interaction forces that hold cofactors seen as HETATM, and we are going to use that approach to increase the range of proteins that could be submitted to our method. Herein, we showed the applicability of algorithms to predicting binding-sites of cisplatin and transplatin compounds. As a perspective of future, an analysis of several models of BCL-2 protein...
shall be made in order to provide a better understanding of how our method works to the many structures obtained through NMR. As shown in results, the binding site Met 1 of ubiquitin was not found due to absence of hydrogens in the first amino acid of the analysed protein, which has misled into a wrong electrical potential of the Met 1 residue. Hence, next studies will seek to include adding/removing hydrogens algorithm in order to take into account the pH range and enhance the mapping of molecular potential surface. This will make the study more realistic, and the scoring function can be used to test other types of ligands. In order to improve the method and wash out the false positives, we are conducting studies to take into account the electronic density on the molecular surface based on quantum chemistry [50]. In addition, we will include in future work a molecular dynamic approach [51] in order to see how our method deals with structures in motion. Therefore, MSProt had a better prediction of cisplatin and transplatin binding-sites than HexServer and PatchDock methods, which provide a more precise insight to assist in new drugs designing.

Acknowledgment

We thank Jose R. Ruggiero for helpful discussions and comments. This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

10. Laurie AT, Jackson RM (2005) Q-SiteFinder: an energy-based method for the mapping of molecular potential surface. This will make the hydrogen docking of balanol to dynamics snapshots of protein kinase A. Proteins 61: 850-858.


