Modeling of the ubiquinol oxidation (Q) site of Plasmodium falciparum (3D7) cytochrome bc$_{1}$ complex (Pfbc$_{1}$) - It is known that atovaquone mechanism of action is to disrupt the mitochondrial electrical potential by inhibiting the cytochrome bc$_{1}$ complex at the Q$_{o}$ site (Kessl et al. 2004). Other hydroxynaphthoquinones, such as compounds 1-41, are thought to share the same binding site. Hence, for modeling the Q$_{o}$ site, we have built homology models for cytochrome b (CYB) and Rieske iron-sulfur protein (ISP), the relevant subunits of the *Plasmodium falciparum* bc$_{1}$ complex (Supplementary data, Figs 1-2).

Homology modeling was carried out using the Modeller software version 9v2 (Šali & Blundell 1993). Amino acid sequences for CYB and ISP subunits from *P. falciparum* (isolate 3D7) cytochrome bc1 complex were obtained from the UniprotKB/Swiss-Prot database under accession codes Q02768 and Q8IL75, respectively. The models for CYB and ISP were built based on the crystal structures of stigmatellin bound cytochrome bc from cow (Huang et al. 2005) (PDB code 2A06, 2.1 Å resolution, 41% identity), chicken (Zhang et al. 1998) (PDB code 3H1J, 3.0 Å resolution, 45% identity) and yeast (Lancaster et al. 2007) (PDB code 2IBZ, 2.3 Å resolution, 42% identity). Protein sequence alignments were performed using the T-Coffee server (Notredame et al. 2000). A set of 50 modeled structures was generated and ranked by the objective function and the top-scoring models were checked for their and overall structural quality, using the Procheck (Laskowski et al. 1993), Verify-3D (Luthy et al. 1992), ERRAT (Colovos & Yeates 1993) and Whatcheck (Hoof et al. 1996) computer programs. A single model was selected for further analysis and as a starting structure for docking simulations.

The 3D models obtained for the CYB and ISP subunits were correctly folded as indicated by the presence of 98.8% and 99.3% of the residues in the most favored or allowed regions of the Ramachandran plot, respectively (Supplementary data, Fig. 3). Verify 3D results presented 67.21% and 53.25% of the residues from the CYBandISPsubunits respectively had an averaged 3D-1D score > 0.2; ERRAT results showed the overall quality factor of 74.7 and 63.1 for CYBandISPsubunits respectively. The models were considered useful by validation results.

For the construction of the receptor used in the docking methodology, the Pfbc$_{1}$ subunits composing the Q$_{o}$ site, CYB and ISP, were superimposed on the corresponding polypeptide chains of the biologically functional assembly of the homologous yeast bc$_{1}$ complex (downloaded from PDB under code 4PD4 (Birth et al. 2014)). ISP superimposition was performed such that the extrinsic domain in the *P. falciparum* protein would tightly dock on CYB (b position), as observed in the crystal structures of atovaquone bound bc$_{1}$ complex (Birth et al. 2014). The coordinates of the 2Fe-2S cluster present in the Rieske protein were transferred to the receptor structure due to its proximity to the Q$_{o}$ site pocket.

Predicted binding mode of the hydroxynaphthoquinones on the Q$_{o}$ site of Pfbc$_{1}$ - Compounds structures were constructed in Spartan'10 (Wavefunction, Inc., CA, USA). The initial structure for atovaquone was extracted from the crystallographic complex with cyt bc$_{1}$ from *Saccharomyces cerevisiae* (PDB code 4PD4 (Birth et al. 2014)). Subsequently, all compounds structures were optimised within the AM1 semi-empirical Hamiltonian and atomic point charges calculated by fitting the electrostatic potential. For comparison purposes, docking was performed with two different programs: Surflex-Dock (Jain 2003) v.2.51, as implemented in Sybyl-X 1.2 modeling package (Tripos Int., St. Louis, MO, USA), and AutoDock4.2 (Morris et al. 2009). For Surflex-Dock, the protomol (a representation of an idealised ligand to which putative ligands will be aligned) was derived from atovaquone bound to the Q$_{o}$ site using the parameters threshold and bloat set to 0.50 and 0, respectively. The atovaquone bound model was obtained by transferring coordinates from yeast bc$_{1}$ PDB structure code 4PD4 upon superposition on the *P. falciparum* bc$_{1}$. The Surflex-Dock engine was run with the default parameters set. In addition the Surflex-Dock objective function, compounds poses were ranked for binding affinity with the scoring functions integrated (Jain 1996). For running AutoDock4, the ligand (atovaquone) centre of mass was used as the centre of a grid calculated with 50x50x50 points and grid spacing of 0.375. Compound conformational space was explored employing the Lamarckian genetic algorithm, implemented in AutoDock4.2.

The results of the docking calculations with AutoDock4.2 program provided general support for results described above for Surflex-dock calculations (Supplementary data, Fig. 3). Except for 33 and 37, docking runs resulted in configurations that predominantly grouped together in the cluster containing the best scoring solution, which was also the most populated cluster, indicating good convergence of the Lamarckian genetic algorithm. Surprisingly, all solutions generated by AutoDock for compound 1 is oriented to interact with Glu261, instead of His152. However, as discussed, this compound is very small when compared with the binding site and can assume different orientations inside the Q$_{o}$ site in *P. falciparum* bc$_{1}$.

Compounds 33 and 37 can also assume different orientations inside the Q$_{o}$ site, specially the S enantiomers since some poses presented the polar group (i.e., the nitrobenzyl) oriented to the Hist152. However, this orientation is not favored by others interactions inside the active site.

Supplementary data
Fig. 1: alignment of cytochrome b and iron-sulfur protein subunits and the templates (PDB code 2A06, 2IBZ and 3H1J). Secondary structures of the model were obtained by DSSP program [11]. The identical residues are shadowed in black while conserved residues (75%) are in gray.

### Supplementary data
Fig. 2: ramachandran plot of the model built from cytochrome b (chain A) and iron sulfur protein (chain B) subunits.
Supplementary data

Fig. 3: binding mode analysis for atovaquone analogues in the Q<sub>o</sub> site in <i>P. falciparum bc</i><sub>1</sub>. For each ligand, the top 20 scoring poses generated in AutoDock4.2 are shown. (A) Atovaquone; (B) compound 1; (C) compound 3; (D) compound 33R; (E) compound 33S; (F) compound 37R; (G) compound 37S. For reference, selected protein residues in close contact (3.0 Å) with atovaquone are shown in sticks with carbon atoms colored in salmon (His152 from ISP and the others from CYB subunit). Ligand carbon atoms are colored in white. Other atoms are CPK-colored. Hydrogen bonds are indicated by yellow dotted lines.

<table>
<thead>
<tr>
<th></th>
<th>Atovaquone</th>
<th>1</th>
<th>3</th>
<th>33R</th>
<th>33S</th>
<th>37R</th>
<th>37S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of clusters</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>18</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>Number of conformations in the most populated cluster</td>
<td>47</td>
<td>50</td>
<td>50</td>
<td>11</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Estimated free energy binding (kcal/mol)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-11.21</td>
<td>-7.07</td>
<td>-9.46</td>
<td>-10.28</td>
<td>-10.34</td>
<td>-10.60</td>
<td>-9.99</td>
</tr>
</tbody>
</table>

<sup>a</sup>: the program AutoDockTools (ADT) was used to prepare the protein and the ligand. Gasteiger charges were assigned and nonpolar hydrogen atoms were merged. The grid maps of docking studies were computed using the AutoGrid4. Center of the ligand was used as the center of the grid and it was calculated with 50x50x50 points with grid spacing of 0.375 were calculated. Docking studies were carried out using the program AutoDock4 [13]. Compound conformational space was explored employing the Lamarckian genetic algorithm. The parameters for LGA were defined as follows: a maximum number of 250,000 energy evaluations; mutation and crossover rates of 0.02 and 0.8, respectively. The number of docking runs was set to 50. After docking, all structures generated were assigned to clusters based on a tolerance of 2.0 Å all-atom RMSD from the lowest-energy structure;<br><br><sup>b</sup>: conformations of the lowest docked energy of best scoring cluster.