



Cite this: *Phys. Chem. Chem. Phys.*,
2019, 21, 13545

New insights about the monomer and homodimer structures of the human AOX1†

P. Ferreira,^a N. M. F. S. A. Cerqueira,^a C. Coelho,^b P. A. Fernandes,^a
M. J. Romão^b and M. J. Ramos^{a*}

Human aldehyde oxidase (hAOX1) is a molybdenum dependent enzyme that plays an important role in the metabolism of various compounds either endogenous or xenobiotics. Due to its promiscuity, hAOX1 plays a major role in the pharmacokinetics of many drugs and therefore has gathered a lot of attention from the scientific community and, particularly, from the pharmaceutical industry. In this work, homology modelling, molecular docking and molecular dynamics simulations were used to study the structure of the monomer and dimer of human AOX. The results with the monomer of hAOX1 allowed to shed some light on the role played by thioridazine and two malonate ions that are co-crystallized in the recent X-ray structure of hAOX1. The results show that these molecules endorse several conformational rearrangements in the binding pocket of the enzyme and these changes have an impact in the active site topology as well as in the stability of the substrate (phthalazine). The results show that the presence of both molecules open two gates located at the entrance of the binding pocket, from which results the flooding of the active site. They also endorse several modifications in the shape of the binding pocket (namely the position of Lys893) that, together with the presence of the solvent molecules, favour the release of the substrate to the solvent. Further insights were also obtained with the assembled homodimer of hAOX1. The allosteric inhibitor (THI) binds closely to the region where the dimerization of both monomers occur. These findings suggest that THI can interfere with protein dimerization.

Received 21st February 2019,
Accepted 3rd June 2019

DOI: 10.1039/c9cp01040h

rsc.li/pccp

1. Introduction

The metabolism of xenobiotics is a critical aspect of drug discovery.¹ In the last decades, massive investments have been made toward this research area, which enabled the development of new and more efficient *in silico* and *in vitro* screening systems.

This led to a small (10%) but still important increase in the number of compounds that reach the final stages of the drug development pipeline. Although these investments started to pay off the initial investments, they are still insufficient to make drug discovery efficient. This happens mainly because the metabolism of these compounds is very complex and can invoke multiple enzymes.¹

Along with the well-studied endoplasmic cytochrome P450 system (CYP450), the cytosolic aldehyde oxidase (AOX) has recently emerged as a new key metabolic enzyme in the hepatic phase I metabolism of numerous xenobiotics and in the activation of several prodrugs.^{2–5} However, while the mechanism of action of CYP450 has been intensively studied, the one played by AOX remains poorly understood. This occurs due to the lack of experimental data regarding this enzyme, and that only now starts to become available. An example of this fact is the X-ray structure of the human AOX (hAOX1, EC 1.2.3.1) that has only been characterized very recently by X-ray crystallography.⁶

The available X-ray structure of hAOX1 shows a monomer that is composed by three subdomains: a 20 kDa N-terminal domain containing the two Fe₂S₂ clusters, a 40 kDa central FAD-containing domain and an 85 kDa C-terminal domain containing the Moco.⁶

^a UCIBIO@REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal. E-mail: mjramos@fc.up.pt

^b UCIBIO@REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

† Electronic supplementary information (ESI) available: A table with the mean RMSd values of all ligands of the hAOX treated in the present work; a figure with the 2D representations with information regarding the formal charge of each ligand of hAOX studied in the present work; a figure with information about the composition of the catalytic tunnel of the enzyme; a figure comparing the structure of catalytic tunnel in model 2 and model 5; a figure with the 3D representations of all the structures obtained from the protein–protein docking protocol and the frcmod files with force field parameters used to describe the FES, FAD, MLI, THI and PHT. A gif file comparing the behaviour of gate 1 and gate 2 as well as the substrate, Moco and the loop 1017–1025 located at the entrance of the catalytic tunnel, in model 2 (represented in red) and in model 5 (represented in blue) during the first 20 ns of the MD simulation. See DOI: 10.1039/c9cp01040h

Recent studies have also shown that hAOX1 has a broad substrate affinity and catalyses not only the oxidation of aldehydes and the hydroxylation of heterocycles but also reduction reactions of nitro compounds and N- and S-oxides.^{2,5} The substrates differ substantially from the ones that are catalysed by CYP450, but still little is known on which substrates are specific for each enzyme. It is therefore very difficult to predict when AOX or CYP450 are responsible for the failures of specific clinical candidates.

Taking into account what has been described, new developments regarding hAOX1 are urgent and of great interest. Since only little experimental data regarding hAOX1 is available, computational methods can be used to unveil important features at the atomic level that can help to shed some light on the structure of this enzyme. The insights that computational approaches can provide are very important to complement the available experimental information and to overcome some limitations that constrain those experiments.

In this article, molecular dynamics simulations (MD) were carried out to provide a comprehensive study on the behaviour of hAOX1 over time, and from which significant insights regarding important structural features of the protein will be analysed. In addition, it was also studied the impact that the allosteric inhibitor thioridazine (THI) and the malonate ions (MLI) have in the behaviour of the enzyme. These compounds are found in specific regions of the co-crystallized X-ray structure of the hAOX1 and, while the inhibitor THI was found in an unexpected groove at the protein surface, no information about the putative role of MLI was provided.⁶ The presence of malonate ions is essential for protein crystallization. The other co-crystallized ligands in the X-ray structure are the molybdenum cofactor (Moco) and the substrate phthalazine (PHT) involved in the catalytic reaction of the enzyme, and the FAD and two iron-sulfur (Fe_2S_2) clusters involved in the transfer of electrons between the Moco and the FAD molecule.

As the active form of hAOX1 is a homodimer,⁷ protein-protein docking together with long MD simulations were also performed to analyse the homodimer complex and study the effect of thioridazine in the dimerization process.

2. Methodology

2.1. Modelling

The crystallographic structures of the hAOX1 available in the protein databank (PDB: 4UHX; 4UHW)⁶ are not complete, the coordinates of some residues located on the protein surface, namely residues 167 to 199, 570 to 571, 655 to 660, 713 to 715 and 881 to 882) are missing. In this work, the PDB structure of 4UHX was used since it contains all the ligands (THI, MLI and PHT) whose impact on the protein structure was evaluated. In order to complete the three-dimensional structure of this enzyme, modelling techniques were applied. The modelling was conducted using the SWISS-MODEL^{8–11} modelling server. From all the models that were generated, the one that provided the best Global Model Quality Estimation (GMQE) and QMEAN

scores⁹ was selected for further analysis and used in the molecular dynamics simulations. The values of the GMQE and QMEAN scores of the selected structure were 0.98 and -1.25 respectively.

2.2. Parametrization of Moco, thioridazine and malonate ions.

One of the major limitations when studying metalloproteins, using computational approaches, is the lack of force field parameters of the cofactors. This work is no exception and we had to parameterize all the force constants and equilibrium values for all the bonds and angles involving the Mo ion as well as the charges of the Moco. Initially, the geometry of the Moco was optimized using density functional theory (DFT) calculations and then, potential energy surface scans were conducted in all bonds and angles involving Mo. The determination of the force-field parameters was based on a semi-flexible model approach. This means that the position of some atoms were frozen during the calculations in order to prevent unrealistic fluctuations of the coordination sphere and, thus, to mimic the conformation the cofactor adopts inside the protein.

In the proposed scheme, bond and angle parameters were determined from linear transit scans along the internal coordinates associated with metal-ligand interactions. Electrostatic charges were determined from a RESP fitting of Merz-Kollman charges. The atomic charges were calculated from the optimized structure, at the B3LYP/SDD:6-311++G(3df,2pd) level of theory. More details about the methodology of the development and testing of the parameters can be found in our article.¹²

The force field parameters of the substrate (phthalazine – PHT), the allosteric inhibitor (thioridazine – THI), malonate ions (MLI) and the FAD were parameterized using the ANTE-CHAMBER module¹³ of the AMBER12 package.¹⁴ The parameters of the Fe_2S_2 clusters (FES) were taken from the literature.¹⁵ The THI molecule was parameterized in its neutral form, even though its pK_a in aqueous solution is reported to be 9.5,¹⁶ because, as it can be observed in the X-ray structure, it sits in a hydrophobic pocket. Additionally, it can be seen that several hydrogen bond acceptors are present around the nitrogen atom of the methyl piperidine ring and free to reorient. However, not even one orients itself towards the piperidinic nitrogen, indicating that the thioridazine binds the active site cavity mostly in the neutral form.

The resulting *frmod* and *prep/lib* files are transcribed in the *ESI*.[†] The parameters of the Moco are shown in Table SI-2 (*ESI*†). The behaviour of all ligands whose parameters were either adapted from existing force fields (THI, MLI, PHT and FAD) or from the literature (FES), or were determined by us (Moco) was carefully examined to assure they maintained a coherent structure throughout the simulations, we also calculated the mean RMSd of all ligands (Table SI-1, *ESI*†) which returned small values reinforcing our confidence in the adopted parameters.

All calculations were carried out with the B3LYP/SDD:6-31G(d) level of theory using the Gaussian09 package.¹⁷ B3LYP^{18–20} is currently the most used functional, yielding good results on the estimate of several thermodynamic parameters

and the use of the effective core potentials (ECPs) is a common feature in the optimization of systems with transition metals. The SDD^{21,22} pseudopotential implemented in the Gaussian software, has been widely used in several studies.^{23–25}

2.2.1. Molecular dynamics protocol. The x-leap module of AMBER12 was used to protonate the complex, assuming that all residues were in their physiological protonation states, and 13 Na⁺ counter-ions were added to neutralize the charge of the system. Additionally, the system was surrounded with water molecules using a truncated rectangular box of TIP3P water molecules with a minimum distance of 12 Å between any atom of the protein and the faces of the box. The ff10 force field and the general Amber force field (GAFF)²⁶ were used to describe the protein, MLI, THI, PHT and FAD. The total number of atoms in the system (including waters) varied (according to the number of ligands included in each model) between 143 819 in model 1 and 150 564 in model 2.

Before the MD simulations, four step minimizations cycles, using the steepest descent algorithm, were carried out to remove eventual tensions or clashes in the structure. In the first step, the water molecules added by x-leap were minimized, in the second step, the hydrogen atoms were minimized, in the third step, the side chains of the protein were minimized and, finally, in the fourth step, all the system was minimized.

After the minimization procedure, the MD simulation protocol was conducted. Starting from the structure obtained after the minimization procedure, the system was gradually warmed to 310.15 K, under constant volume periodic boundary conditions (NVT), using the Langevin thermostat for that purpose, in a 20 picoseconds long equilibration step. Then, a 300 nanoseconds long production step was conducted under constant pressure and temperature (NPT) using the Langevin thermostat and the Berendsen barostat for that purpose. The cut-off for the Lennard-Jones interactions was set to 10 Å. The Coulomb interactions were treated using the Particle Mesh-Ewald (PME) method, with a cut-off of 10 Å for the real part of the sum. The time step of numerical integration was 2 fs. This protocol was conducted in three replicates for each model. Thus, the total simulation time for each model was 900 ns.

The trajectories obtained with the MD simulations were visualized using VMD.²⁷ The numerical results of the simulations were interpreted using the cpptraj program in the AMBER 12 package.

2.3 Protein–protein docking protocol

In order to generate the homodimer complex of hAOX1, protein::protein docking calculations, using the ClusPro 2.0 webserver,^{28–31} were carried out.

The ClusPro2.0 performs a rigid-body docking using a PIPER-based algorithm, a docking program based on the Fast Fourier Transform (FFT) correlation approach that has performed well in the CAPRI (Critical Assessment of Prediction of Interactions) docking challenge and in other docking benchmarks. We used the “balanced” energy function coefficients as described in ref. 31. The number of solutions of each cluster, together with the visual analysis of the resulting complexes, was chosen as criteria to select the final dimer model.

To analyse the robustness of the homodimer, a MD simulation of 50 ns was performed using the protocol that was described in the previous sections.

3. Results and discussion

Despite the key role played by AOX in human metabolism only fragmentary information about the enzyme is available. The recent X-ray structure of hAOX1 homodimer⁶ (crystallographic symmetry) revealed the position of all the four prosthetic groups: one Moco cofactor, two Fe₂S₂ clusters and one FAD molecule, as well as one of the substrates of the enzyme, phthalazine (PHT), which is located very close to Moco. In the deposited structure, two additional molecules are identified: the non-competitive inhibitor thioridazine (THI) that becomes lodged in a particular region at the protein surface and two malonate ions (MLI) that are used as crystallization additive. The two MLI are located nearby the Fe₂S₂ cofactors.

In order to shed some light on the role played by these molecules in the enzyme and the impact that they have in the protein structure, several models of the hAOX1 monomer were created containing different combinations (presence/absence) of THI, MLI and PHT. These results will be presented in the first subsection of the discussion section, where several MD simulations and structural analyses were carried out. Subsequently, we will discuss the assemblage of the dimer of hAOX1 using protein::protein docking methodologies that were followed by a 50 ns long MD simulation. These results allowed to provide additional details about the role played by THI and the amino acid residues that are important for the dimerization process.

3.1. Computational insights regarding the hAOX1 monomer

Both X-ray structures deposited in the protein data bank (PDB code: 4UHX and 4UHW) lack the coordinates of some amino acid residues that correspond to exposed mobile regions. To complete these gaps, modelling techniques were used to produce an integral structural model of the monomer of hAOX1.

From the resulting model, six different complexes were generated, regarding the presence/absence of substrate (PHT), malonate ions (MLI) and the allosteric inhibitor (THI), as illustrated in Fig. 1. Each of those models was subjected to three replicates of 300 ns long molecular dynamics simulation and thereby used to obtain important functional information. The different analyses made to these structures are discussed in the next sections.

3.1.1. Global analysis of the MD simulations. The root mean square deviation (RMSd) plots determined from all the MD simulations are displayed in Fig. 2. These plots show that, after the first 10 ns, all models were equilibrated. In the following frames (290 ns), the overall structure of the crystallographic residues remained almost constant and exhibited an overall RMSd value between 2.68 and 2.98 Å. Most of the structural fluctuations on the protein structure came from the very mobile loop (residues 167–199) that was modelled (Fig. 2), as well as from the binding pocket where some conformation

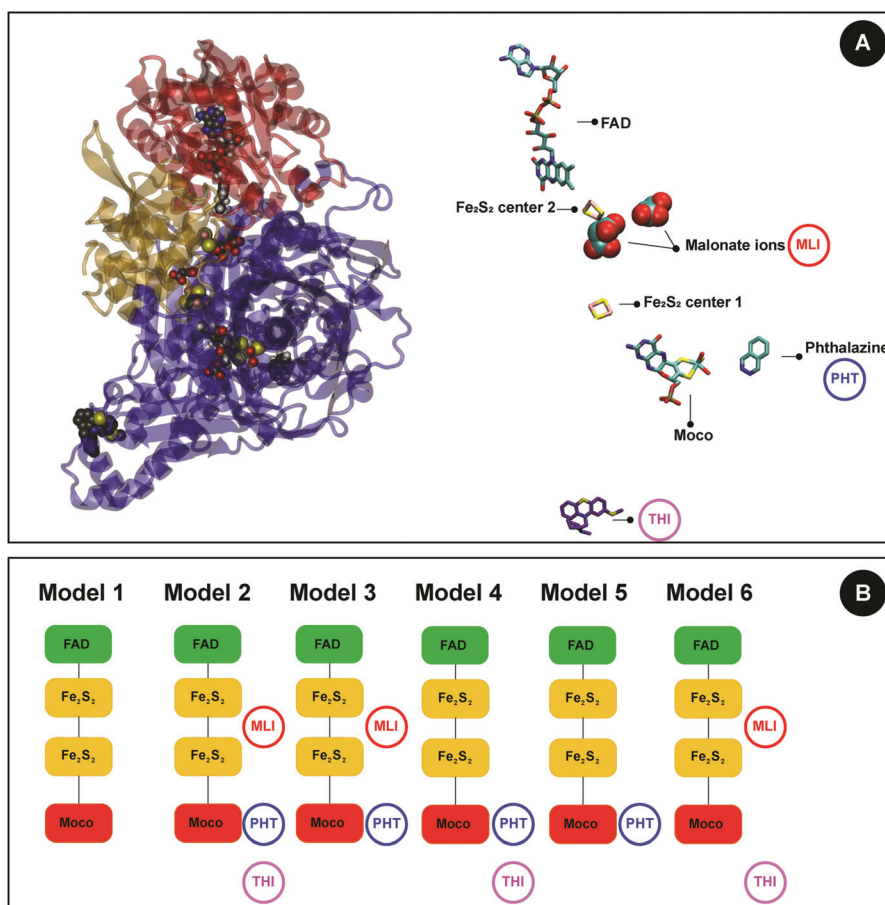


Fig. 1 (A) 3D representation of the X-ray structure of hAOX1 (monomer) and relative positions of all the ligands present in the crystallographic structure. In the 3D structure the blue colored domain represent the 85 kDa C-terminal domain containing the Moco, the dark yellow colored domain represent the 20 kDa N-terminal domain containing the two Fe₂S₂ clusters and the red colored domain represent the 40 kDa central FAD-containing domain. (B) Schematic representation of the composition of the six models of hAOX1 that were created. The FAD, the Fe₂S₂ clusters and the Moco were present in all models. 2D structures of all ligands are available as ESI,† (Fig. SI-1). MLI = malonate ion; PHT = phthalazine; THI = thioridazine.

rearrangements are observed. The overall RMSd values after the first 50 ns range between 2.73 and 3.03 Å. Although there is a small increase in the mean values comparatively to those obtained from 10 ns to the end, they stay within standard deviation values indicating that the overall structure remains stable.

The mean RMSd values with the contribution of the modelled residues taken into account stayed between 3.61 and 4.16 Å (after 10 ns) and between 3.63 and 4.32 Å (after 50 ns) (see Table 1). Although these values are not small, they are acceptable for regions that do not have a defined secondary structure and are very exposed to the solvent, as it happens to be the case.

The conformational rearrangements in the binding pocket have a lower contribution to the total RMSd values and result from the absence of some molecules that were initially present in the X-ray structure of hAOX1 (used to build all the models). These differences are those that we were seeking since they will tell us the impact that each molecule has in the protein structure.

The behaviour of the Moco was carefully examined in all the MD simulations to ensure that all the parameters adopted were

reliable and did not introduce any instability into the system. The RMSd of the Moco remained below 1.0 Å in all models throughout the full simulation in all replicates (Table 1), confirming that the determined parameters are robust and reliable. As expected, despite these small conformational rearrangements in the Moco cofactor, when it is superimposed to the X-ray structure of hAOX1, it occupies the same region of the active site and only a small translation is observed.

Similar conclusions can also be drawn analysing the root mean square fluctuations (RMSF) of all the residues present in each model. As expected, the majority of the residues that exhibited high fluctuations belong to the 32 AA exposed region that was modelled, comprising the residues 167 to 199 as it is clearly seen in Fig. 3. Some fluctuations are also found around the entrance of the binding pocket and in solvent exposed areas of the enzyme. This is clear in the RMSF values of the residues from 1120 to 1170.

3.1.2. Distribution of the water molecules in the binding pocket of hAOX1. The distribution of the water molecules in the binding pocket of hAOX1 was also evaluated in all models. This was achieved through the calculation of the radial distribution function (RDF) near the Moco.

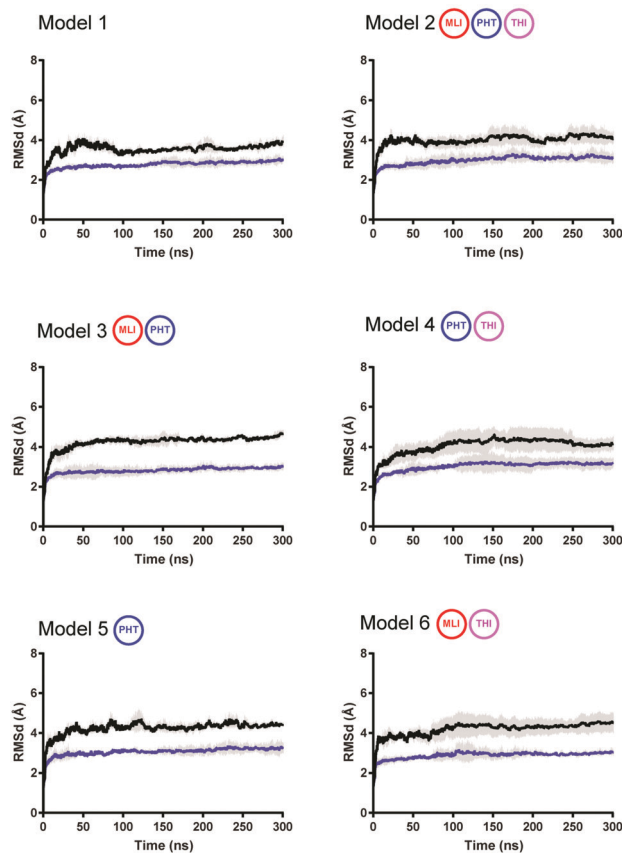


Fig. 2 Representation of the mean RMSd values \pm SEM of the three 300 ns long MD simulation replicates conducted for the Models 1 to 6, of the whole protein (black line) and of the protein without accounting the residues of the modelled mobile loop (indigo line). To facilitate the analysis of the charts, the ligands present in each model are indicated. The standard error of the mean is represented as a gray area above and below the lines.

Table 1 Mean of the root mean square deviation (RMSd) values from the 10 ns and from 50 ns until the end (300 ns) of each MD simulation replicate, obtained from each model of the protein, with and without the unstructured region (loop that was modelled by homology modelling (residues 167–199))

Model	RMSd (Å)		
	Protein with loop	Protein without loop	Moco
1	3.61 ± 0.21^a	2.68 ± 0.11^a	0.75 ± 0.08^a
	3.63 ± 0.18^b	2.73 ± 0.10^b	0.71 ± 0.08^b
2	3.94 ± 0.14^a	2.85 ± 0.15^a	0.84 ± 0.06^a
	3.94 ± 0.13^b	2.93 ± 0.12^b	0.84 ± 0.07^b
3	4.16 ± 0.26^a	2.75 ± 0.08^a	0.86 ± 0.10^a
	4.32 ± 0.11^b	2.78 ± 0.08^b	0.88 ± 0.11^b
4	3.82 ± 0.32^a	2.89 ± 0.17^a	0.77 ± 0.11^a
	4.01 ± 0.20^b	2.99 ± 0.11^b	0.72 ± 0.07^b
5	4.16 ± 0.25^a	2.98 ± 0.12^a	0.85 ± 0.07^a
	4.28 ± 0.15^b	3.03 ± 0.10^b	0.84 ± 0.08^b
6	3.99 ± 0.24^a	2.79 ± 0.12^a	0.91 ± 0.10^a
	4.11 ± 0.10^b	2.86 ± 0.09^b	0.92 ± 0.10^b

^a Mean value calculated after the 10 ns mark. ^b Mean value calculated after the 50 ns mark.

In all models, the RDF analysis shows that only a few water molecules can be found near the Moco. This occurs due to the

hydrophobic nature of the binding pocket that does not allow the concentration of many water molecules in that region of the protein (Fig. 4). But even those that are able to penetrate into the binding pocket, do not stay for a long time, as suggested by the analysis of the MD trajectories and corroborated by the low RDF values (always below 0.6, Fig. 4A). A different pattern is observed above 4 Å from the Moco. For example, in the models that do not contain PHT or those that contain PHT together with THI or MLI or both, at the end of the MD simulation, the binding pocket is soaked with water molecules above the Moco region. In model 5, that only contains the substrate, the binding pocket is more shielded from water molecules.

The reason behind this behaviour can be found by looking at models 2 and 5 that include the substrate (PHT) and only differ in the presence or absence of the MLI and THI molecules. In these models, it can be identified the presence of two loops near the entrance of the active site that were proposed by the crystallographers⁶ to behave as two gates (Gate 1: 652 to 661 and Gate 2: 708 to 719) that control the access of substrate/solvent and release of product to or from the binding pocket. Loop regions that behave as gates for the access and release of substrates, products, inhibitors and solvent at the entrance of the catalytic tunnel are also found in other molybdenum containing enzymes.⁶

In the models that contain THI and MLI (for example model 2, (Fig. 4A) these two gates adopt an open configuration that make the binding pocket more permeable to the entrance of water molecules (GIF-1.SI, ESI†).

In model 5, the two gates adopt a closed configuration blocking the access of the solvent molecules to the binding pocket (Fig. 4B) and contributing to a tighter catalytic tunnel. This closed configuration was only observed, throughout the simulation, in model 5. In the gif file available as ESI,† which compares the behaviour of important of three loops of the catalytic tunnel in model 2 (loops in red) and model 5 (loops in blue) at the beginning of the MD simulation, a different behaviour of the two gates in model 2 and in model 5 (GIF-1.SI, ESI†) can be clearly seen. These findings are interesting because model 5 includes only the natural substrate of the enzyme (PHT) and should mimic the wild type enzyme. The obtained results go therefore in line with the available experimental data, which shows that the substrates of hAOX1 are very hydrophobic and planar and only a hydrophobic active site with few water molecules can sustain those types of substrates close to Moco.

Comparing the RDF results of models 1 and 6, being the difference between the two the presence of MLI and THI in model 6, we can conclude that the presence of the ligands in model 6 made the active site even more permeable to water molecules (Fig. 4A). The reason why the catalytic tunnel of model 1 is permeable in the first place is due to the absence of substrate, which is clear by comparing the RDF results of models 1 and 5 (Fig. 4A).

3.1.3 Integrity of the binding pocket of hAOX1. The analysis of the MD simulations also show that the integrity of the catalytic tunnel is one of the ways through which THI and MLI affect the

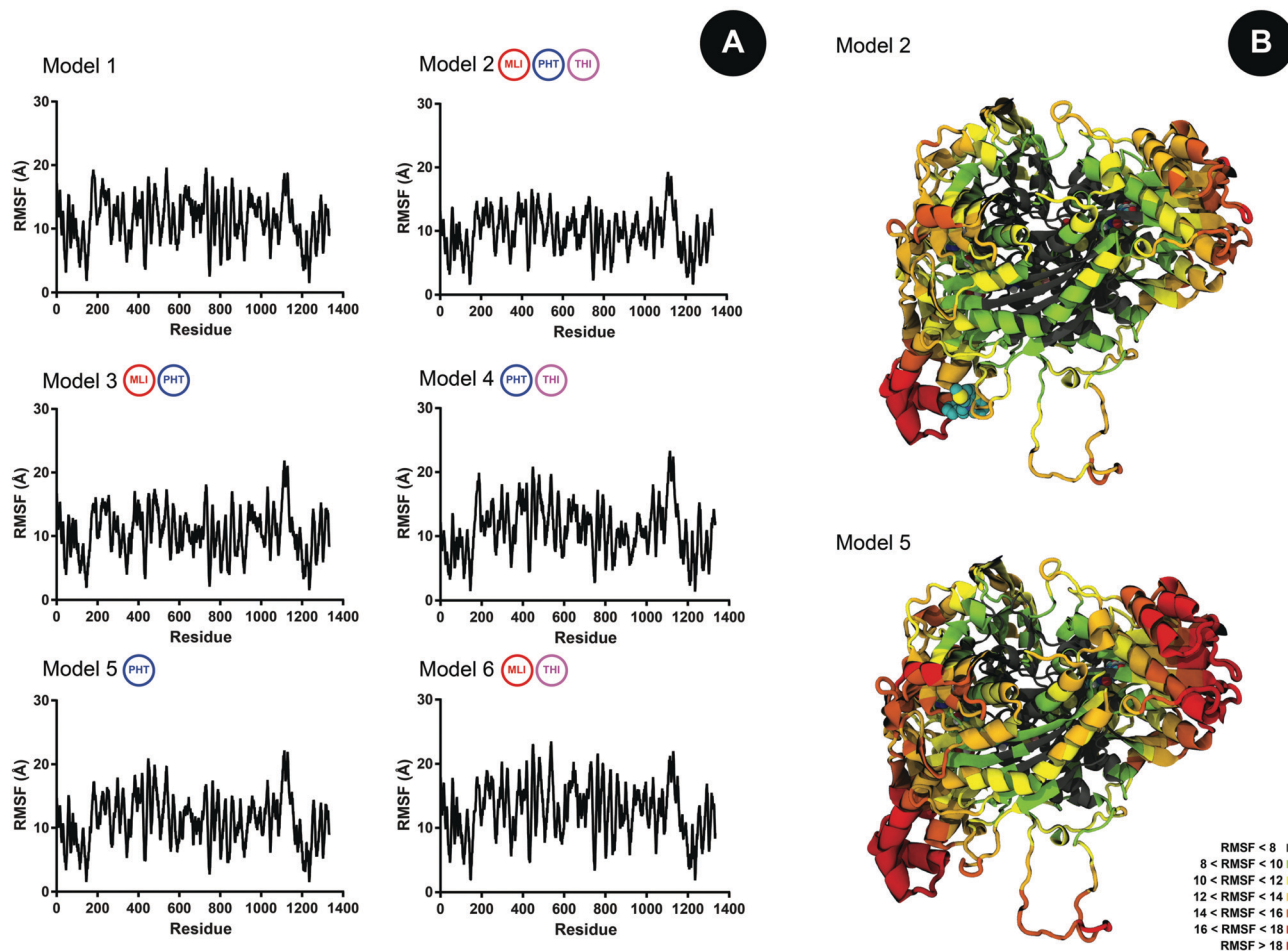


Fig. 3 (A) Representation of the mean RMSF values per residue of the three MD simulation replicates for all models (1 to 6) taken in 0.2 ns intervals. Standard deviation bars are not shown in the charts for easier viewing. To facilitate the analysis of the charts, the ligands present in each model are shown. (B) 3D representation of models 2 and 5, highlighting the amino acid residues where the RMSF values are the highest ones.

normal function of hAOX1. To analyse this aspect, the position and conformation of the residues that compose the walls of the binding pocket were analysed. The catalytic tunnel of hAOX is essentially composed by hydrophobic residues, namely Phe635, Leu812, Val811, Ala919, Phe920, Phe807, Phe923, Ile1085, Phe777, Phe885, Leu884 and Leu880. The polar and charged residues in the walls of the tunnel are essentially found near the Moco or near the entrance of the pocket Fig. SI-2 (ESI[†]).

The results show that the average position and rearrangement of these residues differ from model to model. However, there is a constant pattern among them, depending whether the THI or MLI molecules are present or not. In the models that contain the THI and MLI molecules, the binding pocket becomes wider (Fig. 5, model 2; Fig. SI3, ESI[†]) and, similarly to what was found in the RDFs analysis, the two gates that are located at the entrance of the binding pocket adopt an open conformation. This open configuration, exposes the binding pocket to the solvent, the catalytic tunnel becomes wider which induces the displacement of the substrate away from the catalytic site. In model 5, the gates are closed, the solvent does

not have access to the active site and the binding pocket near the Moco adopts a narrower conformation that maintains the substrate close to the metallic cofactor (see gif file submitted as ESI[†], for a structural comparison of the gates in model 2 and in model 5 in the beginning of the MD simulation).

It is worth mentioning that although THI and MLI are not close to the region of the binding pocket (they are 31 Å and 28 Å away from the Moco, respectively), they induce conformational rearrangements in the structure of the protein that affect this region. This is visible in the loops comprising the residues 791 to 815 and 913 to 924 that are close to the active site and modulate the shape of the binding pocket. In the models that contain the THI and MLI molecules (such as model 2) these loops are displaced and turn the binding pocket wider than, for example, what is observed in model 5 that does not contain THI or MLI (only contains the substrate PHT). In this case, the binding pocket also becomes more constrained near the Moco region turning it more suited to bind the substrate.

The MD simulations have also shown that the modifications endorsed by the MLI and THI molecules in the binding

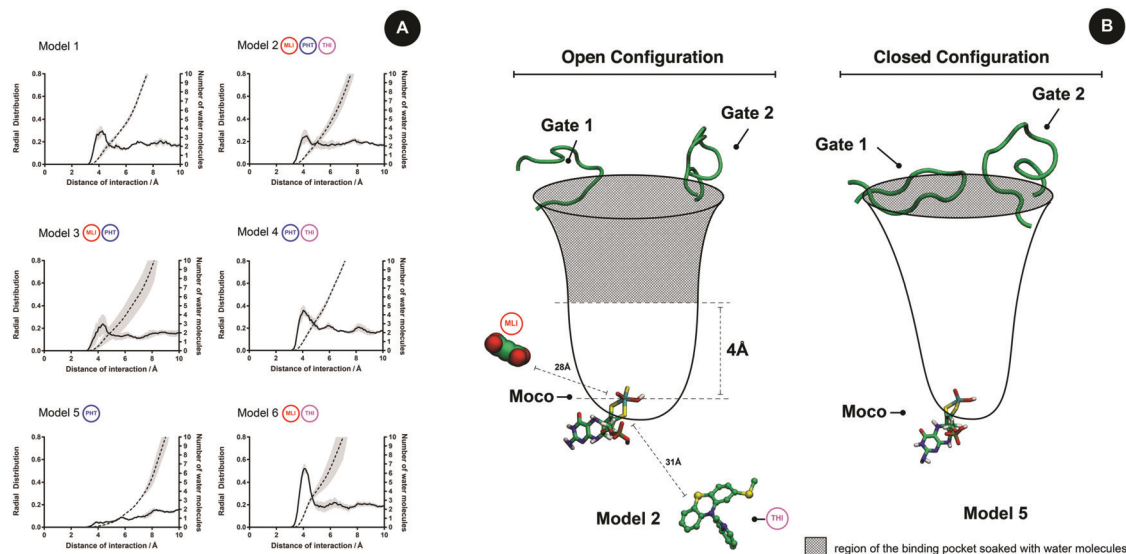


Fig. 4 (A) Mean RDF values of the three replicates of MD simulation of all models (1 to 6) taken in 0.2 ns intervals. To facilitate the analysis of the charts, the molecules present in each model are indicated. The standard error of the mean is represented as a gray area above and below the lines. (B) Schematic representation of the binding pocket of hAOX1, of models 2 and 5, indicating the loops that behave as gates (Gate 1: 652 to 661 and Gate 2: 708 to 719) and control the access of the solvent to the binding pocket.

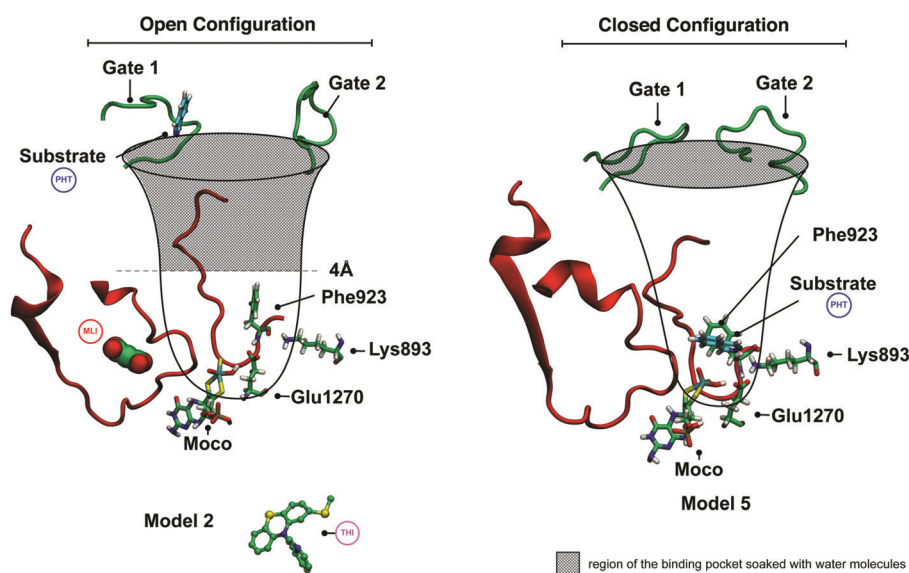


Fig. 5 Schematic representation of the binding pocket of hAOX1 of models 2 and 5, indicating the loops that behave as gates (Gate 1: 652 to 661 and Gate 2: 708 to 719) and control the access of the solvent to the binding pocket, the amino acid residues that modulate the binding pocket shape and the amino acid residues that are closer to the Moco.

pocket also affect the position of the substrate PHT (Fig. SI3, ESI[†]). For example, in the models that contain the substrate with the MLI and THI or only with MLI or THI (models 2, 3 and 4 respectively), the substrate PHT dissociates from the region nearby the Moco and becomes lodged either in the middle of the binding pocket (10.96 Å away from the Moco (mean distance of the three replicates) in model 4) or completely exits from the protein (in all replicates in model 2 and in two replicates of model 3). Only in the model 5 that does not contain the THI and MLI molecules, the substrate stays closer

to the Moco throughout the entire simulation in all replicates (Fig. 5). The mean distances between the Moco and the substrate in the three replicates, over the 300 ns, are 32.89 Å in model 2; 10.50 Å in model 3; 10.96 Å in model 4 and 6.86 Å in model 5.

The explanation for these observations relies on the fact that the tunnel that provides the access to the active site of hAOX1 is very hydrophobic and, although the substrates of hAOX1 are also hydrophobic, they contain a small polar region that needs to be stabilized in the region near the

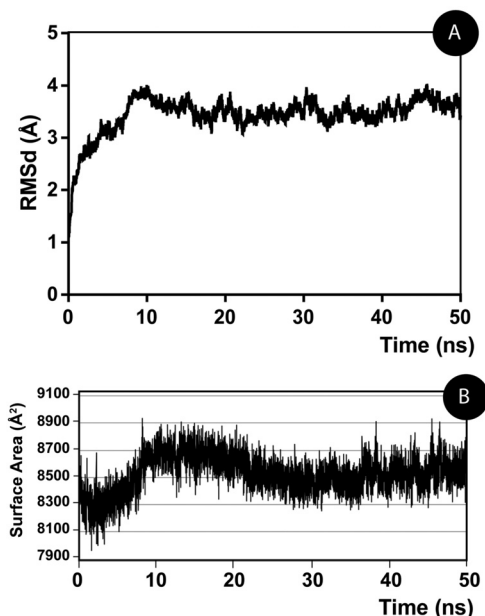


Fig. 6 Representation of the RMSd of the backbone of the hAOX1 homodimer (A) and the surface area of contact between both monomers (B) during the full MD simulation.

Moco (Fig. 5 and Fig. SI-2, ESI[†]). This region is the one that endorses the correct orientation of the substrate inside the binding pocket and makes it ready to react with the Moco. Looking at the region of the binding pocket near the Moco, there are only two amino acid residues that have such characteristics and can play such role. They are Glu1270 and Lys893. Glu1270 is inaccessible to the substrate because it lies at the bottom of the cofactor. Lys893 is, on the other hand, rather accessible. However, the accessibility of Lys893 to the substrate in the binding pocket varies from model to model. In model 5, which resembles the wild type configuration of the enzyme, this residue is adequately accessible and interacts very closely with the substrate PHT at the beginning of the MD simulation. In the simulations in which the PHT dissociates from the enzyme, the Lys893 became buried in the protein structure and unavailable to interact with PHT at the start of the simulation. This is particularly evident in the models where the inhibitors THI and MLI are present, which strongly suggests that their presence introduces structural changes in the binding pocket (in particular in Lys893 and Phe923) that prevent a favourable interaction between the substrate and the active site of the enzyme (Fig. 5).

3.2. Computational insights regarding the hAOX1 homodimer

Since hAOX1 is active as a homodimer, we also modelled computationally the structure of the hAOX1 homodimer. To this end, molecular docking methodologies were applied. From these calculations, a total of 29 complexes were generated (Fig. SI4, ESI[†]). From the visual analysis of all the homodimers combined with the docking scores that were obtained, the ones with the worst scores and unfeasible complexes were discarded.

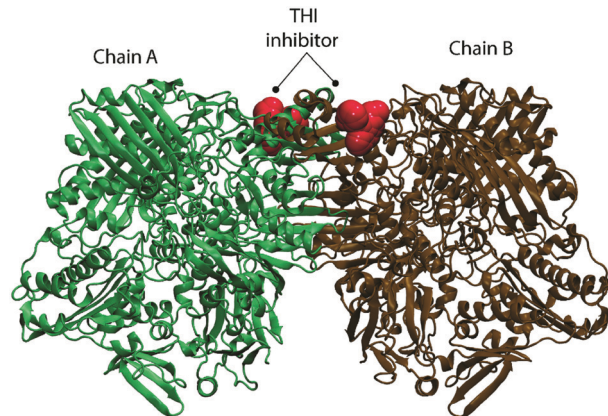


Fig. 7 Structure of the hAOX1 homodimer with the position of the allosteric inhibitor THI highlighted in red.

Among the best scored ones, we analysed the contact surface area between both monomers and the one where this value was maximum was chosen for further analysis.

The selected structure was then subjected to a long molecular dynamics simulation of 50 ns. After an initial adjustment between both monomers, the full system stabilized after 20 ns with an RMSd around 4 Å (Fig. 6A). The full stabilization of the dimer interface was however only obtained after 30 ns as it can be seen through the analysis of the area of contact between both monomers during the full MD simulation (which was calculated using the Volarea plugin³² for VMD, Fig. 6B).

The stabilized structure of the hAOX1 homodimer revealed that both monomers bind in an almost symmetrical way and a perfect contact between both surfaces was obtained.

Another interesting information that can be obtained from the hAOX1 homodimer complex concerns the location of THI. By superimposing the X-ray structure of hAOX1 with the generated homodimer, it becomes clear that THI binds very closely to the dimerization region (Fig. 7). This means that THI can interfere with the dimerization of the enzyme, and therefore besides interfering with the structure of the binding pocket (discussed before), it may also influence the formation of an active hAOX1.

4. Conclusions

In this work, homology modelling, molecular docking and MD simulations were used to gather a detailed atomistic analysis of the monomer and homodimer of hAOX1. With these studies, we were able to investigate the impact that the substrate, the MLI ions and the allosteric inhibitor THI have in the dynamics of the enzyme.

The performed analysis allowed concluding that the THI and MLI molecules interfere with the structure of the enzyme, with emphasis in the region of the binding pocket. Although these two molecules are far away from this region of the protein, they endorse specific modifications in some loops of the protein, in particular loops 791–815 and 913–924 that are close to the binding pocket. These readjustments influence the

configuration of other two loops that are located at the entrance of the binding pocket and behave as gates (Gate 1: 652 to 661 and Gate 2: 708 to 719) that control the entrance of the solvent and possibly the substrate. In the models that contain THI and MLI, these gates adopt an open configuration and the binding pocket becomes more permeable to water molecules. When only the substrate is present, the gates adopt a close configuration and the metallic cofactor becomes more shielded from water molecules.

The presence of THI and MLI species also reshape the region of the binding pocket that is close to the Moco. In particular, they turn the region near the Moco wider and Lys893 remains buried in the protein surface. When this happens, the substrate no longer interacts efficiently with the Moco and tends to dissociate to the solvent. This result highlights the importance of Lys893 in the active site, which plays an important role in stabilizing the substrate and turn it ready for catalysis. Similar conclusions were also observed in the AOX from mouse.³³

The assemblage of the hAOX1 homodimer provided new and important insights about the dimerization process that is required for the full activity of the enzyme.

The hAOX1 homodimer has also revealed that the allosteric inhibitor THI binds very closely to the region where contact between the two monomers occurs. These results suggest that in addition to the structural modification that THI makes in the binding pocket, it can also interfere with protein dimerization.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

The work was supported by UID/MULTI/04378/2019 with funding from FCT/MCTES through national funds. Additionally, we acknowledge PRACE for awarding us access to resource MARENOSTRUM based in Spain at Barcelona Supercomputing Center. The support of David Vicente from the High-Level Support Team at the Barcelona Supercomputing Centre, Spain, to the technical work is gratefully acknowledged. NMFSAC thanks FCT for the Project IF, with the reference IF/01310/2013 and Pedro Ferreira for his PhD Grant with the reference SFRH/BD/119206/2016.

References

- L. G. Yengi, L. Leung and J. Kao, *Pharm. Res.*, 2007, **24**, 842–858.
- D. C. Pryde, D. Dalvie, Q. Y. Hu, P. Jones, R. S. Obach and T. D. Tran, *J. Med. Chem.*, 2010, **53**, 8441–8460.
- E. Garattini and M. Terao, *Drug Metab. Rev.*, 2011, **43**, 374–386.
- C. Fu, L. Di, X. Han, C. Soderstrom, M. Snyder, M. D. Troutman, R. S. Obach and H. Zhang, *Drug Metab. Dispos.*, 2013, **41**, 1797–1804.
- M. J. Romao, C. Coelho, T. Santos-Silva, A. Foti, M. Terao, E. Garattini and S. Leimkuhler, *Curr. Opin. Chem. Biol.*, 2017, **37**, 39–47.
- C. Coelho, A. Foti, T. Hartmann, T. Santos-Silva, S. Leimkuhler and M. J. Romao, *Nat. Chem. Biol.*, 2015, **11**, 779–783.
- E. Garattini and M. Terao, *Expert Opin. Drug Metab. Toxicol.*, 2012, **8**, 487–503.
- K. Arnold, L. Bordoli, J. Kopp and T. Schwede, *Bioinformatics*, 2006, **22**, 195–201.
- M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli and T. Schwede, *Nucleic Acids Res.*, 2014, **42**, W252–W258.
- L. Bordoli, F. Kiefer, K. Arnold, P. Benkert, J. Battey and T. Schwede, *Nat. Protoc.*, 2009, **4**, 1–13.
- N. Guex, M. C. Peitsch and T. Schwede, *Electrophoresis*, 2009, **30**, S162–S173.
- P. Ferreira, N. M. F. S. A. Cerqueira, N. F. Brás, P. A. Fernandes and M. J. Ramos, *J. Chem. Theory Comput.*, 2018, **14**(5), 2538–2548.
- J. M. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graphics Modell.*, 2006, **25**, 247–260.
- T. A. D. D. A. Case, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, *AMBER 2012*, University of California, San Francisco, 2012.
- A. T. P. Carvalho, A. F. S. Teixeira and M. J. Ramos, *J. Comput. Chem.*, 2013, **34**, 1540–1548.
- J. Swarbrick, *Clarke's isolation and identification of drugs*, The Pharmaceutical Press, London, 2nd edn, 1986.
- M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision D*, Gaussian Inc., Wallingford CT, 2009.
- A. D. Becke, *J. Chem. Phys.*, 1996, **104**, 1040–1046.
- A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098–3100.

- 20 C. T. Lee, W. T. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.
- 21 M. Dolg, U. Wedig, H. Stoll and H. Preuss, *J. Chem. Phys.*, 1987, **86**, 866–872.
- 22 D. Andrae, U. Häußermann, M. Dolg, H. Stoll and H. Preuß, *Theor. Chim. Acta*, 1990, **77**, 123–141.
- 23 S. F. Sousa, P. A. Fernandes and M. J. Ramos, *J. Am. Chem. Soc.*, 2007, **129**, 1378–1385.
- 24 N. M. F. S. A. Cerqueira, P. A. Fernandes, P. J. Gonzalez, J. J. G. Moura and M. J. Ramos, *Inorg. Chem.*, 2013, **52**, 10766–10772.
- 25 R. P. P. Neves, S. F. Sousa, P. A. Fernandes and M. J. Ramos, *J. Chem. Theory Comput.*, 2013, **9**, 2718–2732.
- 26 J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
- 27 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics Modell.*, 1996, **14**, 33–38.
- 28 S. R. Comeau, D. W. Gatchell, S. Vajda and C. J. Camacho, *Bioinformatics*, 2004, **20**, 45–50.
- 29 D. Kozakov, D. Beglov, T. Bohnuud, S. E. Mottarella, B. Xia, D. R. Hall and S. Vajda, *Proteins*, 2013, **81**, 2159–2166.
- 30 D. Kozakov, R. Brenke, S. R. Comeau and S. Vajda, *Proteins*, 2006, **65**, 392–406.
- 31 D. Kozakov, D. R. Hall, B. Xia, K. A. Porter, D. Padhorny, C. Yueh, D. Beglov and S. Vajda, *Nat. Protoc.*, 2017, **12**, 255–278.
- 32 J. V. Ribeiro, J. A. Tamames, N. M. Cerqueira, P. A. Fernandes and M. J. Ramos, *Chem. Biol. Drug Des.*, 2013, **82**, 743–755.
- 33 M. Mahro, N. F. Bras, N. M. Cerqueira, C. Teutloff, C. Coelho, M. J. Romao and S. Leimkuhler, *PLoS One*, 2013, **8**, e82285.