**QUEEN'S
UNIVERSITY
BELFAST**

Catalytic mechanism of Phenylacetone monooxygenases for non-native linear substrates: implications on rational engineering of BVMOs to expand the substrate specificity

Carvalho, A. T. P., Dourado, D. F. A. R., Skvortsov, T., de Abreu, M., Ferguson, L., Quinn, D. J., Moody, T. S., & Huang, M. (2017). Catalytic mechanism of Phenylacetone monooxygenases for non-native linear substrates: implications on rational engineering of BVMOs to expand the substrate specificity. *Physical Chemistry Chemical Physics*, 19(39), 26851-26861. <https://doi.org/10.1039/c7cp03640j>

Published in:
Physical Chemistry Chemical Physics

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

Copyright 2017 Royal Society of Chemistry. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Catalytic mechanism of Phenylacetone monooxygenases for non-native linear substrates: implications on rational engineering of BVMOs to expand the substrate specificity

Alexandra T.P. Carvalho,^{a,b} Daniel F.A.R. Dourado,^{a,b} Timofey Skvortsov,^{b,c} Miguel de Abreu,^b

Lyndsey J. Ferguson,^b Derek J. Quinn,^b Thomas S. Moody,^b Meilan Huang^{a,}*

a. School of Chemistry and Chemical Engineering, Queen's University, David Keir Building,
Stranmillis Road, Belfast BT9 5AG, Northern Ireland, UK

b. Almac Sciences, Department of Biocatalysis and Isotope Chemistry, Almac House, 20 Seagoe
Industrial Estate, Craigavon BT63 5QD, Northern Ireland, UK

c. School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97
Lisburn Road, Belfast BT9 7BL, Northern Ireland, UK

Corresponding Author

E-mail: *m.huang@qub.ac.uk

ABSTRACT

Phenylacetone monooxygenase (PAMO) is the most stable and thermo-tolerant member of the Baeyer-Villiger monooxygenases family, and therefore it is an ideal candidate for the synthesis of industrially relevant compounds. However, its limited substrate scope has largely limited its industrial applications. In the present work, we provide, for the first time, the catalytic mechanism of PAMO for the native substrate phenylacetone as well as for a linear non-native substrate 2-octanone, using molecular dynamics simulations, quantum mechanics and quantum mechanics/molecular mechanics calculations. We provide a theoretical basis for the preference of the enzyme for the native aromatic substrate over non-native linear substrates. Our study provides fundamental atomic-level insights that can be employed in the rational engineering of PAMO for wide applications in industrial biocatalysis, in particular, in the biotransformation of long-chain aliphatic oils into potential biodiesels.

1. Introduction

The design of tailored enzymes for industrial synthetic applications is highly desirable as a cost-effective and ecological friendly approach.¹ Diverse enzymes can be obtained from directed evolution through random mutagenesis methods such as error-prone polymerase chain reaction (epPCR),² DNA shuffling³ and saturation mutagenesis.⁴ Most of these methods do not require a previous knowledge of the protein structure and catalytic mechanism. However such approaches are time consuming, most mutations are neutral or even deleterious⁵⁻⁷ and the generated gene libraries (up to 10^{15} sequences) only represent a small fraction of the possible sequence space. Moreover, these methods usually require a large number of substitutions to achieve an enzyme

with the desired catalytic activity.⁸ This happens because in nature (and in random approaches) enzymes usually evolve through smooth weak trade-off routes where mutations accumulate keeping the robust native activity, while the promiscuous activities increase gradually.⁸ This type of route produces highly promiscuous generalist enzymes, which have low overall activities (including the main one) and hence require extensive modifications, in opposition to specialized enzymes that have higher catalytic proficiencies (defined as $k_{cat}/K_M/k_{uncat}$). Rational mutagenesis guided by understanding of the substrates binding and catalytic mechanisms offers the possibility of overcoming these limitations by restraining the changes to a few amino acid positions with functional significance.⁹

Baeyer-Villiger monooxygenases (BMVOs) are interesting targets for protein engineering. These enzymes are able to oxidize a variety of substrates such as cyclic or acyclic ketones into the corresponding lactones or esters.¹⁰⁻¹² The catalytic cycles of BVMOs require nicotinamide-adenine-dinucleotide phosphate 2'-monophosphoadenosine 5'-diphosphoribose (NADPH) and flavin-adenine dinucleotide (FAD) cofactors as well as molecular oxygen. Reduction of FAD by the NADPH cofactor and subsequent oxygenation leads to the formation of the highly reactive C4a-peroxyflavin intermediate (Fig. 1-Step A) that then attacks the carbonyl carbon of the substrate¹³ to yield a tetrahedral Criegee adduct (also called Criegee intermediate (Fig. 1-Step B))^{10, 14}. In the next step this intermediate rearranges into the product (an ester or a lactone, depending on the initial substrate) and a C4a-hydroxyflavin intermediate (Fig. 1-Step C) that is then regenerated back to the flavin cofactor by dehydration.¹³ A QM/MM study on the catalytic mechanism of cyclohexanone monooxygenase (CHMO) indicates the reaction involves the formation of stable C4a-peroxyflavin and tetrahedral Criegee intermediates, which are stabilized by a catalytic arginine and NADP^+ via strong hydrogen bonds.¹⁵

The prototype of the BVMO enzyme family is the *Thermobifida fusca* phenylacetone monooxygenase (PAMO).¹⁶ This enzyme is particularly thermo-stable, with an optimum temperature of 56 ° Celsius¹⁷ and a tolerance for different solvents.¹⁶ However, it has a limited substrate scope. In opposition, other BVMOs, such as CHMO, catalyze a broader range of substrates,^{10, 13, 18} but are not as thermo-stable as PAMO, with an optimum temperature of around 25-30 ° Celsius.¹⁹⁻²² Several studies attempted to introduce mutations on PAMO, aiming to increase the substrate scope by employing saturation mutagenesis approaches. Most of these studies were focused on the loop (residues 440-446) and remaining active site residues.^{23, 24} An exception was reported by Wu *et al.* who focused on distal residues.²⁵ In the initial studies the mutations were shown to only affect the PAMO substrate specificity for the bulky substrate 2-phenylcyclohexanone and its analogues. Deletion of S441 and A442²³ or randomization of positions 441 to 444 increased its conversion, but with low enantioselectivity.²⁶ Later, randomization of position 440 identified mutants with higher *E*-values for 2-phenylcyclohexanone, that can also convert bulky substituted cyclohexanones.¹⁷ In addition, iterative saturation mutagenesis (ISM) of five active site residues identified mutants able to convert para-substituted cyclohexanone derivatives, although with lower conversion than the wild type (WT) CHMO. More recently, starting from a triple mutant derived from combination of three previously identified mutations; F440P, Q93N and P94D,^{19, 23} mutants with higher cyclohexanone conversion and minimal trade-off in enzyme thermo-stability were generated by applying ISM. Additionally, in another study the M446G mutant was shown to increase the enantioselectivity for a range of ketone, amines and sulfides.²⁷ Some of these mutations severely decreased the conversion of the native substrate.²³ The aforementioned substrates are summarized in the supporting information (Fig. S1). Dudek *et al.* identified a PAMO mutant

with increased activity towards linear substrates by screening a random library composed of 1,500 mutants.²⁸ The best mutant, P253F/G254A/R258M/L443F, was shown to convert several substrates that could not be converted by the WT PAMO. The highest increase in activity was obtained for 2-octanone (19.2-fold increase in $k_{\text{cat}}/K_{\text{M}}$ in relation to the WT enzyme).²⁸ Recently, a large-scale biotransformation of ricinoleic acid into ester by engineered BVMO was reported, demonstrating the industrial application of BVMO-based whole-cell biocatalysis.²⁹ Ricinoleic acid is the precursor of castor oil, a main resource in biodiesel production. Thus the potential application of engineered BVMO to catalyze unsaturated linear substrates in production of biofuels can be envisioned.

In the present study we provide the rationale for PAMO's discrimination between linear (2-octanone) and aromatic (phenylacetone) substrates by performing molecular docking, classical molecular dynamics (MD) simulations, quantum mechanics (QM) calculations and quantum mechanics/molecular mechanics (QM/MM) simulations (Fig. S2). This is the first atomic-level study to describe the complete PAMO catalytic pathway including binding of the two distinct substrates, formation of the corresponding Criegee intermediates and decay into the corresponding esters.

Fig. 1 Simplified catalytic mechanism of PAMO for the native substrate phenylacetone. The mechanism starts with reduction of the FAD by the cofactor NADPH. It is followed by reaction of the reduced FADH with molecular oxygen forming the C4a-peroxoflavin intermediate (Step **A**). In the presence of the substrate phenylacetone, the peroxyflavin oxygen attacks the substrate ketone forming a tetrahedral intermediate (the Criegee intermediate) (Step **B**), which is then resolved into C4a-hydroxyflavin and a product (ester or lactone) (Step **C**). Oxidized FAD is then regenerated by elimination of water and the NADP⁺ is released allowing for binding a new NADPH cofactor.

2. Experimental

2.1 Computational methods

2.1.1 Modelling

The crystal structure of the flavin-peroxide intermediate of *T. fusca* PAMO (pdb code: 2YLT) was used as the initial model system.³⁰ This structure contains the FAD cofactor and the NADP⁺ cofactor (the 2-(n-morpholino)-ethanesulfonic acid inhibitor was removed). The crystal poses of the cofactors were kept. The peroxy form of FAD was generated with the peroxy group bound to the isoalloxazine ring of the FAD cofactor. The C4a-peroxyflavin was modeled in the anionic form due to stabilization by the surrounding catalytic residue Arg337 and anionic feature of the subsequently generated Criegee intermediate.

The peroxy FAD, substrate, NADP⁺ and Criegee intermediates geometries were optimized using the Gaussian09 program³¹ with the B3LYP exchange-correlation functional³²⁻³⁴ and the 6-31G(d) basis set in the condensed phase. Point charges were calculated resorting to the RESP method³⁵ from HF/6-31G(d) single point energy calculations (Table S2, Table S3). Remaining parameters were obtained from the parm99SB³⁶ and GAFF³⁷ force fields using the antechamber program in AMBER14.^{38, 39}

2.1.2 Molecular Docking

Molecular docking was performed using the AutoDock 4.2 suite of programs with the Lamarckian genetic algorithm (LGA).⁴⁰ A grid box was centered on the oxygen of the peroxy group. R337 was set to be flexible. A total of 100 LGA runs were carried out for each ligand:protein complex. The population was 300, the maximum number of generations was

27,000 and the maximum number of energy evaluations was 2,500,000.

2.1.3 Molecular dynamics

MD simulations were performed using the Amber molecular dynamics program (AMBER14)^{38, 39} with the parm99SB³⁶ and GAFF³⁷ force fields. The structures were placed within an octahedral box of TIP3P³⁸ waters and counter ions were added to make the entire system neutral. The systems were subjected to two initial energy minimizations and to 500 ps of equilibration in a NVT ensemble using Langevin dynamics with small restraints on the protein (10 kcal/mol) to heat the system from 0K to 300 K. Production simulations were carried out at 300 K in the NPT ensemble using Langevin dynamics with a collision frequency of 1.0 ps⁻¹. Constant pressure periodic boundary conditions were imposed with an average pressure of 1 atm. Isotropic position scaling was used to maintain pressure with a relaxation time of 2 ps. The time step was set to 2 fs. SHAKE constraints were applied to all bonds involving hydrogen atoms.⁴¹ The particle mesh Ewald (PME) method⁴² was used to calculate electrostatic interactions with a cutoff distance of 10 Å. For each reactant complex three 20-ns simulations with random initial velocities were run, while for each Criegee intermediate a 10-ns simulation was run.

2.1.4 QM cluster model calculations

We built models of the active centre that include the isoalloxazine ring and two carbon atoms of the ribitol of the FAD cofactor, the substrate and the R337 side-chain (Fig. S2).

The phenylacetone cluster model contains 73 atoms, while the 2-octanone model includes 78 atoms. For both the phenylacetone and 2-octanone systems linear scans of the reaction steps **B** and **C** (Fig. 1) were performed. The reaction mechanism was studied at two different theoretical

levels: semi-empirical PM3 and dispersion corrected DFT. In the DFT study the exchange-correlation functional B3LYP³²⁻³⁴ along with the Grimme D3 dispersion correction⁴³ was used with a 6-31G(d) basis set. The geometry optimizations were carried with the beta carbon (C β) of R337 and the C4a atom of the C4a-peroxyflavin kept fixed. In both cases, single point energies of the stationary states were recalculated with the conductor-like polarizable continuum model (CPCM)^{44 45} with a dielectric constant of 4. In the DFT study the basis set was increased to B3LYP/6-311++G(d), following a similar protocol applied in previous literature.⁴⁶ Frequency calculations were performed to confirm the nature of the transition state structures and intermediates as well as introduce zero-point and thermal corrections to the final energies. All QM calculations were performed using Gaussian09 program³¹.

2.1.5 QM/MM calculations

The QM/MM calculations were performed using the internal semi-empirical hybrid QM/MM functionality implemented in AMBER14⁴⁷ with periodic boundary conditions. The PM3 semi-empirical method was employed for the high level layer and the MM region was described by the Amber parm99SB force field. The high level layer includes the substrate, the side chain atoms of R337, the isoalloxazine ring of the peroxyflavin cofactor and the nicotinamide and ribose rings of the NADP⁺ cofactor (Fig. S2). The QM layer of the enzyme-phenylacetone complex contains a total of 123 atoms while that of the enzyme-2-octanone complex contains 128 atoms.

The boundary was treated via the link atom approach and long-range electrostatic interactions were described with an adapted implementation of the PME method for QM/MM.⁴⁷ Electrostatic embedding was employed.⁴⁸

The reaction coordinate for the formation of the Criegee intermediate (Fig. 1 - Step **B**) was defined as the distance between the peroxy oxygen of the peroxyflavin cofactor and the carbonyl carbon (C2) of the substrate ($d_{\text{OX-C2}}$). The reaction coordinate for the decay of the Criegee intermediate into benzyl acetate (Fig. 1 - Step **C**) was defined as the distance between the same oxygen and the C3 carbon ($d_{\text{OX-C3}}$). The distances were restrained to decrements of 0.1 Å using the umbrella sampling method, except that decrements of 0.02 Å were employed near the transition states for step **B**. Each sampling window was run for 200 ps. The potential of mean force (PMF) was calculated resorting to the WHAM method.⁴⁹

The main limitation of the PM3 Hamiltonian is the fact that it was parameterized to reproduce gas phase geometries and heats of formation for molecules in their ground states while the transition states were left out from the training set. However, the QM/MM with the QM region treated with PM3 has been shown to provide reasonable geometries in numerous enzyme studies.⁵⁰⁻⁵² Moreover, it is worth noting that the non-enzymatically catalyzed Baeyer-Villiger reaction was previously studied using PM3 and returned results in agreement with experimental data.^{53, 54} For each substrate studied we calculated the PMF with PM3 and then applied high level corrections to these energies, following the protocol reported in previous literature.⁵⁵ Provided that the interaction energy of the QM/MM does not change significantly at both levels of theory, the corrected free energy of the PMFs can be described by equation (1).

$$\Delta G_{\text{PMF,corr}}(\text{protein}) = \Delta G_{\text{PMF,PM3}}(\text{protein}) + [\Delta G_{\text{DFT}}(\text{model}) - \Delta G_{\text{PM3}}(\text{model})] \quad (1)$$

$\Delta G_{\text{DFT}}(\text{model}) - \Delta G_{\text{PM3}}(\text{model})$, is the difference in the free energies for the QM layer model calculated by PM3 and B3LYP/6-31G(d), respectively.

Assuming that the thermal and zero point energy corrections are small and thus negligible, the corrected free energy of the PMF can also be described by equation (2).

$$\Delta G_{PMF,corr}(protein) = \Delta G_{PMF,PM3}(protein) + [E_{DFT}(model) - E_{PM3}(model)] \quad (2)$$

$E_{DFT}(model) - E_{PM3}(model)$ corresponds to the difference in the energies of the stationary points that were optimized using PM3 and DFT (B3LYP using a 6-31G(d) basis set and B3LYP along with the Grimme D3 dispersion correction⁴³ using a TZVP basis set methods) . The correction was first performed for the reagents and products. For the rest of the potential energy surface, the correction was interpolated by incrementally adding to each point a correction factor, which is defined as the energy difference between the two stationary points divided by the number of points along the reaction coordinates.

2.2 Experimental Methods

2.2.1 Chemicals, Strain and Plasmid

Chemicals were purchased from Sigma Aldrich UK. The PAMO sequence of 1,638 bp (accession code: Q47PU3) was cloned into the pET-28a(+) plasmid utilising the NdeI and XhoI restriction sites.

2.2.2 Protein expression and purification

E.coli cultures were incubated overnight in 5mL LB media supplemented with kanamycin (50 mg/mL) in a shaking incubator at 37°C. Cultures were added to 1 L LB media supplemented with kanamycin and incubated at 37°C to OD₆₀₀ = 0.5. Cultures were induced by the addition of IPTG (1 mM final concentration) to the culture media and incubated at 25°C for 16h. After

incubation, the cell pellets were collected by centrifugation (5000 rpm at 4°C) and resuspended in 100mM Tris buffer pH 7.4. Cells were lysed by sonication on ice 10 x 10 s separated by a 1 min interval. Cell debris was removed by centrifugation (30 min at 5000rpm) at 4°C. The protein was purified using the HisPur™ Cobalt Resin protocol (Thermo Scientific™). Protein concentration was determined using Bradford protocol described by Sigma-Aldrich. Protein purification was conducted under mild conditions to prevent the loss of the flavin cofactor. The recombinant proteins displayed brightly yellow colour in the solution, indicating the presence of FAD prosthetic group. The concentration of FAD-bound PAMO was determined by measuring the absorbance of the protein at 441 nm. The extinction coefficient for FAD-bound PAMO ($\epsilon_{441} = 12.4 \text{ mM}^{-1} \text{ cm}^{-1}$) was previously obtained by Fraaije *et al.*¹⁶ The protein concentration of the Bradford assay and the concentration of the FAD-bound PAMO were compared and a molar ratio of 1:1 was obtained for the wild-type enzyme.

2.2.3 Characterization of enzymes

The kinetic parameters of the wild-type PAMO were determined using 96-well plate assay. The conditions used for kinetic analysis assays were as follows: 50 mM Tris·Cl (pH 7.4), 100 μM NADPH, 5% 1,4-dioxane (v/v), 200 μl final volume. All reactions were set up in triplicates and conducted at 25°C. The reactions were initiated by adding NADPH and monitored by continuously measuring the absorbance of the reaction plate samples at 340 nm ($\epsilon_{340} = 6.22 \text{ mM}^{-1} \text{ cm}^{-1}$) using Epoch Microplate Spectrophotometer (BioTek, USA). The substrate was in the range of 10 – 5000 μM , while the final concentration of enzymes present in the reaction was in the range 0.1–1.0 μM . The GraphPad Software v6 (GraphPad Software, USA) was used to fit

the data and to obtain the kinetic parameters (Fig. S6, Table 1). The program is based on the following formula: ⁵⁶

$$V=V_{\max}*[S]/(K_M + [S]*(1+[S]/K_i)),$$

where:

$V_{\max}$ is the maximum enzyme velocity,

$[S]$ is the substrate concentration,

V is the velocity,

K_M is the Michaelis-Menten constant,

K_i is the dissociation constant for the inhibitory ESS ternary complex.

3. Results and Discussion

3.1 Catalytic mechanism of WT PAMO towards phenylacetone and 2-octanone

3.1.1 Substrate binding

MD simulations of the complex enzyme:NADP⁺:FADOO⁻ with the native substrate phenylacetone shows that the substrate establishes few interactions with the enzyme and cofactors. The strongest interaction is a cation- π interaction between the phenyl group of substrate and the conserved R337 residue. R337 is in close proximity with the FAD cofactor as disclosed in the first crystal structure,⁵⁷ indicating that R337 might have an important role in PAMO's catalytic mechanism. Mutations of this residue (R337A and R337K) demonstrated loss of activity towards phenylacetone, which further proved the catalytic importance of this

residue.¹³ In the simulations the average distance between the centers of mass of the phenyl ring in the substrate and the side chain of R337 is 4.21 ± 0.42 Å. The carbonyl group in the ketone moiety is in close proximity of the isoalloxazine ring of the C4a-peroxyflavin cofactor, while the methyl group interacts via London dispersion forces with A442, M446, which are located on an active site loop (residues 440-446) previously shown to alter the conversion and enantioselectivity of either bulky cyclohexanone substituents or sulfides.^{23, 58} Throughout the simulations the distance between the attacking oxygen atom of the C4a-peroxyflavin and the substrate carbonyl carbon atom (dox1_c2) was retained around 3 Å, a favorable distance for the reaction to occur (Fig. 2, S4).

MD simulations were also conducted for the linear non-native substrate 2-octanone. The long aliphatic tail of 2-octanone interacts with the aliphatic part of the R337 side chain. In addition, it also interacts with P286, L338 and L340 via weak London dispersion forces. This represents a catalytically relevant pose where the peroxy-FAD oxygen is close to the carbonyl carbon of the substrate (replica 1& 3, Fig. 3A). It is worth mentioning that the hydrophobic interactions formed with P286, L338 and L340 do not necessarily always prompt the substrate to adopt a catalytically relevant pose. In one of the replicas (replica 2), the carbonyl moiety moves away from the C4a-peroxyflavin cofactor to establish an ion-dipole interaction with a second arginine R258 that is nested at the entrance of the active site pocket (Fig. 3B, S4).

Thus, it appears that the ability of PAMO to bind aromatic substrates in a catalytically relevant pose is attributed to the cation- π interaction with the conserved R337 in the relatively large substrate pocket. While this interaction is absent in the case of the linear substrate 2-octanone, it is compensated by the weak interactions formed with P286, L338 and L340. To further validate this hypothesis and to ascertain any other factors that contribute for the discrimination of

aromatic and linear substrates, we examined the chemical reaction path concerning the formation and decay of the Criegee intermediate using QM and QM/MM methods, respectively (Tables 2 and 3, Fig. 4 and 5).

Fig. 2 A) Crystal structure of the WT PAMO (pdb code: 2YLT) in complex with the modelled C4a-peroxyflavin intermediate and the native substrate phenylacetone docked into the active site; **B)** MD reference structure of the same complex. The MD reference structure corresponds to the lowest RMSD structure in relation to the average structure of the simulation. No significant changes were observed in the MD replicas. Relevant distances are shown (Å).

Fig. 3 **A)** Crystal structure of the WT PAMO (pdb code: 2YLT) in complex with the modeled C4a-peroxyflavin intermediate and the non-native substrate 2-octanone docked into the active site; **B)** MD reference structure of the replica 1 (replica 3 is similar to replica 1). **C)** MD reference structure of replica 2. The MD reference structure corresponds to the lowest RMSD structure in relation to the average structure of the simulation. Relevant distances are shown (Å).

3.1.2 Formation of the Criegee intermediate and decay into the ester product

The kinetic parameters associated with PAMO catalysis of phenylacetone were previously

measured.¹³ The k_{cat} is 1.9 s^{-1} at 25°C , which corresponds to a free energy barrier of 17.1 kcal/mol, and the K_{M} is 0.059 mM.¹⁶ The kinetic parameters for 2-octanone were not accurately determined due to substrate solubility issues in the previous assays,²⁸ where an approximate k_{cat} of $> 1 \text{ s}^{-1}$ at 37°C (corresponding to a free energy barrier of < 18.2 kcal/mol) and a K_{M} of > 2 mM were reported. In order to compare the PAMO-catalyzed conversion of the two substrates, we determined the steady-state kinetic parameters associated with the conversion of 2-octanone, and obtained a k_{cat} value of 0.22 s^{-1} at 25°C (corresponding to a free energy barrier of 18.4 kcal/mol) and a K_{M} of 3.2 mM (Table 1).

Table 1 Steady-state kinetic parameters of PAMO WT for the conversion of 2-octanone. The values obtained are averaged results of three measurements. Kinetic parameters were determined with GraphPad Prism 6; the kinetic curves were fitted using Michaelis-Menten equation for substrate inhibition (Fig. S6).

	WT PAMO
K_{M} [mM]	3.2
k_{cat} [s^{-1}]	0.22
$k_{\text{cat}}/K_{\text{M}}$ [$\text{M}^{-1} \text{ s}^{-1}$]	66

In the reaction catalyzed by CHMO, a covalent adduct between the substrate and the cofactor, the so-called Criegee intermediate, is obtained by nucleophilic attack of the substrate cyclohexanone by the C4a-peroxyflavin, which subsequently undergoes fragmentation to form the lactone product.¹⁵ The reaction mechanism for PAMO-catalyzed oxidation of phenylacetone may also occur via the formation of a Criegee intermediate that then decays into benzyl acetate (Fig. 1). To elucidate the catalytic mechanism of PAMO for native substrate phenylacetone and

non-native substrate 2-octanone, we conducted QM cluster calculations and QM/MM MD simulations to explore the respective reaction pathways for the two substrates.

3.1.2.1 QM Cluster Models

We built cluster models of the active centers for both the phenylacetone and 2-octanone substrates and calculated their respective reaction paths with DFT (B3LYP-D) method (Table 2 and S1). The optimized structures of the transition states are shown in Fig. 4.

For phenylacetone, the formation of the Criegee intermediate (Fig. 1 - Step B) is associated with an activation free energy ($\Delta G^{\ddagger}_{TS1}$) of 5.2 kcal/mol, while the reaction free energy (ΔG_r) is 6.4 kcal/mol (Fig. 5A, Table 2). The subsequent step, which corresponds to the decay into the ester (Fig. 1 - Step C), has a free energy barrier ($\Delta G^{\ddagger}_{TS2}$) of 10.3 kcal/mol and a ΔG_r of -68.9 kcal/mol (Fig. 5A, Table 2).

Fig. 4 QM cluster model calculations. DFT (B3LYP-D/6-31G(d)) optimized geometries of: (A)

and (B) TS1 and TS2 of phenylacetone complex; (C) and (D) TS1 and TS2 of 2-octanone complex. Relevant distances are reported in Table S1.

The overall energy barrier ($\Delta G^{\ddagger}$), which corresponds to the energy difference between the initial reactant complex (denoted as RC hereafter) and TS2, is 16.7 kcal/mol (Table 2). So our calculated $\Delta G^{\ddagger}$ is well in line with the experimentally determined kinetic constant which corresponds to an energy barrier of 17.1 kcal/mol at 25°C.¹⁶

For 2-octanone, step **B** has a $\Delta G^{\ddagger}_{TS1}$ of 7.0 kcal/mol and a ΔG_r of 5.5 kcal/mol, while Step **C** has a $\Delta G^{\ddagger}_{TS2}$ of 13.2 kcal/mol and a ΔG_r of -71.6 kcal/mol (Fig. 5B, Table 2). The overall energy barrier $\Delta G^{\ddagger}$ is 18.7 kcal/mol, which is line with the experimentally measured energy barrier of 18.4 kcal/mol. Thus, our *in silico* calculations and kinetic results for the two substrates showed similar reaction profiles, indicating that 2-octanone follows a similar reaction mechanism to the native substrate phenylacetone.

The predicted $\Delta G^{\ddagger}$ for the linear substrate 2-octanone is only 2.0 kcal/mol higher than that of the phenylacetone. This is not surprising since the experimental data indicates a difference of 1.6 kcal/mol and the catalytic reaction coordinates of TS2 derived from the Criegee intermediate for 2-octanone are similar to those of the corresponding transition state derived from the Criegee intermediate for phenylacetone.

Fig. 5 Gibbs free energies (in kcal/mol) of the optimized stationary points along the reaction path at 298.15 K for: A) phenylactone; B) 2-octanone. Optimized geometries were taken from both the PM3 and DFT (B3LYP-D/6-31G(d)) potential energy surfaces and frequencies calculated at the respective theory levels. The energies of the stationary points were further corrected using a larger basis set 6-311++G(d) in the DFT based reaction profiles. Energies of the stationary points were obtained in the condensed phase (PM3 (blue), B3LYP/6-311++G(d) (black)) with the self-

consistent reaction field CPCM and with a dielectric constant of 4. Zero point energy and thermal corrections were also added. INT stands for Criegee intermediate.

Table 2 Gibbs free energies (in kcal/mol) of the optimized stationary points along the reaction path, based on the cluster model calculations using the semi-empirical PM3 and DFT-D with a continuum description. Full geometry optimization were conducted with B3LYP-D/6-31G(d) followed by single point energy calculations using B3LYP-D/6-311++G(d), in the DFT-D calculations. The protein environment was considered by using the self-consistent reaction field CPCM with a dielectric constant of 4. All the energies reported are in relation to the initial reactant complex.

	<i>phenylacetone</i>		<i>2-octanone</i>	
	ΔG_{PM3}	$\Delta G_{\text{DFT-D}}$	ΔG_{PM3}	$\Delta G_{\text{DFT-D}}$
RC	0	0	0	0
TS1	12.9	5.2	16.1	7.0
INT	6.5	6.4	8.9	5.5
TS2	16.3	16.7	17.6	18.7
PC	-68.2	-62.5	-62.9	-66.1

We also conducted the same calculations using the PM3 semi-empirical to test its accuracy for these model systems. For both substrates the predicted $\Delta G^{\ddagger\dagger}$ barriers calculated using PM3 are similar to the ones obtained with DFT (Fig. 5, Table 2). For phenylacetone the difference calculated using the two methods is only 0.4 kcal/mol, while for 2-octanone it is 1.1 kcal/mol.

However, when analyzing the energy barriers of the individual reaction steps, we find that the energy barriers of Step **B** calculated using PM3 are substantially higher than the ones calculated with the DFT method (Fig. 5A, Table 2). The energy difference in $\Delta G^{\ddagger}_{TS1}$ estimated using the two methods is 7.7 kcal/mol for phenylacetone and 9.1 kcal/mol for 2-octanone, respectively (Table 2). For Step **C** of phenylacetone, the energy difference in $\Delta G^{\ddagger}_{TS2}$ calculated by the two methods is insignificant (0.5 kcal/mol), since the differences in energies of both TS2 and the Criegee intermediate are negligible. However, a notable energy difference of 4.5 kcal/mol is observed for 2-octanone due to the distinct energies of the Criegee intermediates obtained with PM3 and DFT (Fig. 5B, Table 2). Therefore, in the subsequent QM/MM simulations we applied high-level quantum chemical (DFT) corrections to the PMF-based free energy profile for each reaction step to account for any possible energy differences between the methods.

3.1.2.2 QM/MM model

Although the cluster model calculations provided important insights into the reaction mechanism and the obtained energy barriers are in line with the experimental data, only a limited number of residues of the protein were taken in account. In order to identify other residues that might also be involved in the reaction mechanism so as to guide in the rational engineering of the enzyme, we employed a hybrid QM/MM method to sample the energy profiles of the entire protein complexes. The QM layer was calculated with the PM3 Hamiltonian (Fig. S2). In addition, DFT corrections were applied to the energy profiles as described in previous literature.⁵⁵

For phenylacetone, the Step **B** that corresponds to the formation of the Criegee intermediate has a $\Delta G^{\ddagger}_{TS1}$ of 16.2 kcal/mol (Fig. 6A), while ΔG_r is 6.0 kcal/mol. Step **C**, which corresponds to the

decay into the ester, has a $\Delta G^\ddagger_{TS2}$ of 7.9 kcal/mol and a ΔG_r of -34.3 kcal/mol. The overall $\Delta G^\ddagger$ is 13.9 kcal/mol. With DFT (B3LYP/TZVP) corrections applied to the QM layer, the $\Delta G^\ddagger_{TS1}$ of step **B** becomes 15.7 kcal/mol and ΔG_r is 5.4 kcal/mol, while the $\Delta G^\ddagger_{TS2}$ is 9.9 kcal/mol (Table 3). The corrected $\Delta G^\ddagger$ becomes 15.3 kcal/mol, which is close to the experimental measured barrier (17.1 kcal/mol).¹⁶ The energy difference is within the error of the computational method. Correction at the B3LYP/6-31G(d) level was also applied to the QM layer (Table 3). It is worth noting that the corrected $\Delta G^\ddagger$ is close to $\Delta G^\ddagger_{TS1}$ with either B3LYP/TZVP (15.3 versus 15.7 kcal/mol) or B3LYP/6-31G(d) corrections (15.5 versus 15.4 kcal/mol) (Table 3).

Fig. 6 QM/MM Gibbs free energies profile for the WT PAMO catalyzed reaction to convert (A) phenylactone and (B) 2-octanone. Dashed line denotes the energies calculated using PM3/Amber_parm99SB and solid line denotes the corrected energies using B3LYP/TZVP. Step **B**: addition of the substrate to the C4a-peroxyflavin (RC). Step **C**: decay of the Criegee intermediates (INT) to the corresponding esters (PC).

Table 3 Gibbs free energies (kcal/mol) from QM/MM models. The energies were corrected by full geometry optimizations using B3LYP/6-31G(d) and B3LYP/TZVP with empirical dispersion.

B3LYP/6-31G(d)							
	Step B		Step C		$\Delta G^{\ddagger}$	Exp.	Error
	$\Delta G^{\ddagger}_{\text{TS1}}$	ΔG_r	$\Delta G^{\ddagger}_{\text{TS2}}$	ΔG_r			
phenylacetone	15.4	4.7	10.8	-25.4	15.5	17.1	-1.3
2-octanone	16.4	10.0	10.1	-27.4	20.1	18.4	1.7
B3LYP/TZVP							
	Step B		Step C		$\Delta G^{\ddagger}$	Exp.	Error

	$\Delta G^{\ddagger}_{TS1}$	ΔG_r	$\Delta G^{\ddagger}_{TS2}$	ΔG_r			
phenylacetone	15.7	5.4	9.9	-27.9	15.3	17.1	-1.8
2-octanone	14.5	7.4	9.8	-28.5	17.2	18.4	-1.2

The conversion of the linear substrate 2-octanone follows a similar catalytic mechanism, which is accordance to what was found in the QM cluster model study that indicates the formation of a Criegee intermediate. For step **B** we obtained a $\Delta G^{\ddagger}_{TS1}$ of 14.9 kcal/mol and a ΔG_r of 8.0 kcal/mol (Fig. 6B). For step **C** $\Delta G^{\ddagger}_{TS2}$ is 7.1 kcal/mol and ΔG_r is -39.7 kcal/mol. This corresponds to an overall $\Delta G^{\ddagger}$ of 15.1 kcal/mol. With corrections applied to the QM layer (B3LYP/TZVP), the $\Delta G^{\ddagger}_{TS1}$ of step B becomes 14.5 kcal/mol and ΔG_r is 7.4, while the $\Delta G^{\ddagger}_{TS2}$ is 9.8 kcal/mol (Table 3). The overall $\Delta G^{\ddagger}$ becomes 17.2 kcal/mol, which is, as expected, higher than the predicted energy barrier for phenylacetone and also in good agreement with experimental kinetic data that corresponds to an overall energy barrier of 18.4 kcal/mol. We also introduced corrections for the QM layer using the 6-31G(d) basis set. The results show that B3LYP/6-31G(d) correction gives reasonable $\Delta G^{\ddagger}$ energy barriers in relation to the experimental data (Table 3).

Based on the above QM cluster calculations and QM/MM simulations, it is evident that in the reaction catalyzed by PAMO, both the native substrate phenylacetone and the non-native aliphatic substrate 2-octanone are subjected to nucleophilic attack by the deprotonated C4a-peroxyflavin to form a Criegee intermediate, which then decays into their respective ester products. Taking together the mechanism of CHMO proposed by Polyak *et al.*,¹⁵ we demonstrate that the Criegee intermediates in PAMO needs to undergo similar fragmentation to yield the product.

3.1.3 Alkyl migration in decay of the Criegee intermediate of 2-octanone

Decay of the Criegee intermediate of 2-octanone via the aforementioned n-hexyl migration gives the normal product, whereas an abnormal product may also be yielded with the migration of the methyl group. It would be interesting to decide the migratory preference of the n-hexyl group in relation to the methyl group. To present a whole mechanistic picture, we also calculated the potential energy surface corresponding to the methyl migration using QM/MM (PM3/Amber_parm99SB) simulations, plotting the free energy of the enzyme system as a function of the distance between the distal oxygen atom of the C4a-peroxoflavin peroxy group and the carbon atom of the methyl group in 2-octanone (Fig. S5). We found that the methyl migration is associated with a much higher free energy barrier than the n-hexyl migration (27.4 kcal/mol versus 15.1 kcal/mol). The large difference in the energy barriers indicates the decay of the Criegee intermediate of 2-octanone preferably undergoes the n-hexyl migration, and therefore produces the normal product rather than the abnormal product.

2.1.4 MD simulations of the Criegee intermediates

We also performed MD simulations for the Criegee intermediates of phenylacetone and 2-octanone (Fig. 7). Similar to the substrate binding, the Criegee intermediate of the native substrate establishes a cation- π interaction with R337 (Fig. 7A), whereas for the intermediate of 2-octanone only weak London dispersion interactions are observed between its aliphatic tail and residues L289, L338 and L340 (Fig. 7B).

Thus the difference in catalytic proficiencies of the WT PAMO towards the two distinct substrates is attributed to the different binding modes of the substrates as well as the corresponding Criegee intermediates.

Fig. 7 MD reference structure of the Criegee intermediates for the WT PAMO that forms a tetrahedral covalent adduct with: A) phenylacetone; B) 2-octanone. The structures correspond to the ones with lowest RMSD compared to the respective average MD structures. No significant changes were observed in the MD replicas. The aliphatic tail of the 2-octanone adduct establishes interactions with L289, L338 and L340.

4. Conclusions

The high stability and thermo-tolerance of the flavoprotein PAMO makes it an ideal biocatalyst for the oxidation of ketones. However, its limited substrate scope precludes its broader application in industry. The catalytic mechanisms of PAMO calculated by DFT cluster models and QM/MM simulations demonstrate the effective energy barriers are in good accordance with

experimental data. This research provides atomic level insight on the catalytic mechanism for phenylacetone, and provides, for the first time, a description of the mechanism for a linear substrate 2-octanone.

The native substrate phenylacetone makes a cation- π interaction with the conserved R337, which in turn interacts with the peroxy moiety of the C4a-peroxyflavin cofactor. It is this mutual interaction that keeps the substrate in place in the wide active site pocket of the enzyme to facilitate the formation of the Criegee intermediate. In contrast, the missing cation- π interaction for linear substrates such as 2-octanone is compensated by the weak interactions formed between the aliphatic tail of the substrate and a hydrophobic region constituted by P286, L338 and L340. The weak interactions enable the carbonyl end of the substrate to move more freely in the binding pocket, as a consequence, it may move towards R258 to form an ion-dipole interaction.

We observed that in the enzyme catalyzed reactions both substrates have similar free energy barriers in TS1, which corresponds to the formation of the Criegee intermediate. The overall barrier $\Delta G^{\ddagger}$ is higher for 2-octanone than the native substrate phenolacetone, as a result of energy differences in the Criegee intermediates and TS2s.

Our study has shown that additional design efforts should be made in improving the binding of the linear substrates as well as of the corresponding Criegee intermediates. The calculations showed that it is possible to reshape the relatively large active site pocket of PAMO by introducing mutations that would result in preferential interactions with the aliphatic part of this substrate, while the substrate remains close to the C4a-peroxyflavin. These hydrophobic interactions would improve the binding of 2-octanone as well as allow the corresponding Criegee intermediate to be properly stabilized.

In summary we find the spatial requirement essential for improvement in the binding and conversion of long aliphatic substrates by BVMOs, which may provide significant insight in rationally engineering the enzymes for industrial production of biofuels such as castor oils.

Supporting Information: Supplementary results, tables and figures.

Conflicts of interest

There is no conflict of interests to declare.

Acknowledgements

The authors acknowledge the financial support from INVEST NI Research and Development Programme, part financed by the European Regional Development Fund under the Investment for Growth and Jobs programme 2014-2020. We are grateful for the computing resources from QUB high performance computing Centre.

References

- 1 U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore and K. Robins, *Nature*, 2012, **485**, 185-194.
- 2 R. C. Cadwell and G. F. Joyce, *Genome Res.*, 1994, **3**, S136-S140.
- 3 W. P. Stemmer, *Nature*, 1994, **370**, 389-391.
- 4 R. M. Myers, L. S. Lerman and T. Maniatis, *Science*, 1985, **229**, 242-247.
- 5 A. Crameri, S.-A. Raillard, E. Bermudez and W. P. C. Stemmer, *Nature*, 1998, **391**, 288-291.
- 6 D. L. Hartl, *Curr. Opin. Microbiol.*, 2014, **21**, 51-57.
- 7 P. A. Romero and F. H. Arnold, *Nat. Rev. Mol. Cell. Biol.*, 2009, **10**, 866-876.
- 8 O. Khersonsky and D. S. Tawfik, *Annu. Rev. Biochem.*, 2010, **79**, 471-505.
- 9 A. T. P. Carvalho, A. Barrozo, D. Doron, A. V. Kilshtain, D. T. Major and S. C. L. Kamerlin, *J. Mol. Graph. Model.*, 2014, **54**, 62-79.

- 10 G. J. ten Brink, I. W. C. E. Arends and R. A. Sheldon, *Chem. Rev.*, 2004, **104**, 4105-4124.
- 11 R. D. Schmid and V. Urlacher, *Modern Biooxidation: Enzymes, Reactions and Applications*, John Wiley & Sons, 2007.
- 12 D. E. Torres Pazmiño, H. M. Dudek and M. W. Fraaije, *Curr. Opin. Chem. Biol.*, 2010, **14**, 138-144.
- 13 D. E. Torres Pazmiño, B.-J. Baas, D. B. Janssen and M. W. Fraaije, *Biochemistry*, 2008, **47**, 4082-4093.
- 14 R. Criegee, *Liebigs Ann Chem*, 1948, **560**, 127-135.
- 15 I. Polyak, M. T. Reetz and W. Thiel, *J. Am. Chem. Soc.*, 2012, **134**, 2732-2741.
- 16 M. W. Fraaije, J. Wu, D. P. H. M. Heuts, E. W. v. Hellemond, J. H. L. Spelberg and D. B. Janssen, *Appl. Microbiol. Biotechnol.*, 2004, **66**, 393-400.
- 17 M. T. Reetz and S. Wu, *J. Am. Chem. Soc.*, 2009, **131**, 15424-15432.
- 18 M. M. Kayser, *Tetrahedron*, 2009, **65**, 947-974.
- 19 D. Sheng, D. P. Ballou and V. Massey, *Biochemistry*, 2001, **40**, 11156-11167.
- 20 F. Secundo, F. Zambianchi, G. Crippa, G. Carrea and G. Tedeschi, *J. Mol. Catal. B: Enzym.*, 2005, **34**, 1-6.
- 21 N. A. Donoghue, D. B. Norris and P. W. Trudgill, *Eur. J. Biochem.*, 1976, **63**, 175-192.
- 22 Y. Hasegawa, Y. Nakai, T. Tokuyama and H. Iwaki, *Biosci. Biotechnol. Biochem.*, 2000, **64**, 2696-2698.
- 23 M. Bocola, F. Schulz, F. Leca, A. Vogel, M. W. Fraaije and M. T. Reetz, *Adv. Synth. Catal.*, 2005, **347**, 979-986.
- 24 L. P. Parra, J. P. Acevedo and M. T. Reetz, *Biotechnol. Bioeng.*, 2015, **112**, 1354-1364.
- 25 S. Wu, J. P. Acevedo and M. T. Reetz, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 2775-2780.
- 26 M. T. Reetz and S. Wu, *Chem. Commun.*, 2008, 5499-5501.
- 27 D. E. T. Pazmiño, R. Snajdrova, D. V. Rial, M. D. Mihovilovic and M. W. Fraaije, *Adv. Synth. Catal.*, 2007, **349**, 1361-1368.
- 28 H. M. Dudek, M. J. Fink, A. V. Shivange, A. Dennig, M. D. Mihovilovic, U. Schwaneberg and M. W. Fraaije, *Appl. Microbiol. Biotechnol.*, 2014, **98**, 4009-4020.
- 29 J.-H. Seo, H.-H. Kim, E.-Y. Jeon, Y.-H. Song, C.-S. Shin and J.-B. Park, *Sci. Rep.*, 2016, **6**, 28223.
- 30 R. Orru, H. M. Dudek, C. Martinoli, D. E. T. Pazmiño, A. Royant, M. Weik, M. W. Fraaije and A. Mattevi, *J. Biol. Chem.*, 2011, **286**, 29284-29291.
- 31 M. J. T. Frisch, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.

- Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., *Gaussian 09, Revision A.2*, Gaussian, Inc., Wallingford CT edn., 2009.
- 32 C. T. Lee, W. T. Yang and R. G. Parr, *Phys. Rev. B.*, 1988, **37**, 785-789.
- 33 K. Raghavachari, *Theor. Chem. Acc.*, 2000, **103**, 361-363.
- 34 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623-11627.
- 35 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269-10280.
- 36 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins*, 2006, **65**, 712-725.
- 37 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, T. V. J. Wang and K. Words, *J. Compt. Chem*, 2004, **25**, 1157-74.
- 38 R. Salomon-Ferrer, D. A. Case and R. C. Walker, *Wires. Comput. Mol. Sci.*, 2013, **3**, 198-210.
- 39 D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, J. S. W. Smith, Walker, J. Wang, R. M. Wolf, X. Wu and P. A. Kollman, *AMBER 14*, University of California, San Francisco, 2014.
- 40 G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *J. Comput. Chem.*, 1998, **19**, 1639-1662.
- 41 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Compt. Phys.*, 1977, **23**, 327-341.
- 42 T. Darden, D. York and L. Pedersen, *J. Chem Phys.*, 1993, **98**, 10089-10092.
- 43 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem Phys.*, 2010, **132**, 154104.
- 44 T. Vreven, B. Mennucci, C. O. da Silva, K. Morokuma and J. Tomasi, *J. Chem Phys.*, 2001, **115**, 62-72.
- 45 S. J. Mo, T. Vreven, B. Mennucci, K. Morokuma and J. Tomasi, *Theor. Chem. Acc.*, 2004, **111**, 154-161.
- 46 D. F. A. R. Dourado, P. A. Fernandes, M. J. Ramos and B. Mannervik, *Biochemistry*, 2013, **52**, 8069-8078.
- 47 R. C. Walker, M. F. Crowley and D. A. Case, *J. Comput. Chem.*, 2008, **29**, 1019-1031.
- 48 D. Bakowies and W. Thiel, *J. Phys. Chem.*, 1996, **100**, 10580-10594.
- 49 A. Grossfield.
- 50 L. S. Devi-Kesavan and J. Gao, *J. Am. Chem. Soc.*, 2003, **125**, 1532-1540.
- 51 J. Zurek, A. L. Bowman, W. A. Sokalski and A. J. Mulholland, *Struct. Chem.*, 2004, **15**, 405-414.
- 52 D. Xu, Y. Wei, J. Wu, D. Dunaway-Mariano, H. Guo, Q. Cui and J. Gao, *J. Am. Chem. Soc.*, 2004, **126**, 13649-13658.
- 53 R. Cadenas, L. Reyes, J. Lagúñez-Otero and R. Cetina, *J. Mol. Struc. THEOCHEM*, 2000, **497**, 211-225.
- 54 M. A. M. A. Iglesias-Arteaga, G. A. Velázquez-Huerta, J. M. Méndez-Stivalet, A. Galano and J. R. Alvarez-Idaboy, *ARKIVOC*, 2005, 109-126.
- 55 A. L. Bowman, I. M. Grant and A. J. Mulholland, *Chem. Commun.*, 2008, **7**, 4425-4427.
- 56 R. A. Copeland, *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*, Wiley-VCH, New York, 2 edition edn., 2000.

- 57 E. Malito, A. Alfieri, M. W. Fraaije and A. Mattevi, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 13157-13162.
- 58 L. P. Parra, R. Agudo and M. T. Reetz, *ChemBioChem*, 2013, **14**, 2301-2309.