

Human Carboxylesterase 2 in Cocaine Metabolism

Pedro R. Figueiredo ^{a,b}, Ricardo D. González ^{a,b}, Alexandra T.P. Carvalho ^{a,c,*}

^a CNC – Center for Neuroscience and Cell Biology, Institute for Interdisciplinary Research (IIUC), University of Coimbra, 3004-504 Coimbra, Portugal

^b PhD Programme in Experimental Biology and Biomedicine, Institute for Interdisciplinary Research (IIUC), University of Coimbra, Casa Costa Alemão, 3030-789 Coimbra, Portugal

^c Almac Sciences, Department of Biocatalysis and Isotope Chemistry, Almac House, 20 Seagoe Industrial Estate, Craigavon BT63 5QD, Northern Ireland, United Kingdom

ARTICLE INFO

Keywords:

human carboxylesterase 2
enzymes
quantum mechanics/molecular mechanics md simulations
reaction mechanisms
cocaine hydrolysis
drug metabolism

ABSTRACT

Increased hydrolysis of cocaine to nontoxic compounds is a promising way to prevent cocaine-induced toxicity. However, the short half-life of cocaine in the blood and the rapid conversion in the body to the hydrolysis-resistant metabolite benzoylecgonine, limits the therapeutic potential of serum proteins. Therefore, hydrolysis by tissue-specific hydrolases that do not generate benzoylecgonine deserves further investigation. Here, we report for the first time the mechanism of cocaine hydrolysis by the human carboxylesterase 2. We have combined conventional and accelerated Molecular Dynamics, which allowed us to identify the structural motions of the α 1 and α 10' helices that act as a putative lid. Quantum Mechanics/Molecular Mechanics calculations on the full cycle showed that the rate-limiting step is the formation of benzoic acid (deacylation step) with a ΔG of 19.5 kcal mol⁻¹ when calculated with B3LYP-D3/6-311++G(d,p). This ΔG value is in close conformity with the experimental value of 19.7 kcal mol⁻¹.

1. Introduction

Carboxylesterases (CEs, E.C. 3.1.1.1) are members of the family of serine hydrolases and are mainly localized within the lumen of the endoplasmic reticulum in different tissues, such as the liver, small intestine, and kidneys. [1] Several CE families (CE1-CE5) have been identified to date, but only three are currently characterized in humans (hCE1-hCE3), with most studies focusing on hCE1. [2] hCEs are known to be involved in drug metabolism due to their role in therapeutic drug clearance, but also by their ability to activate prodrugs such as anti-cancer agents. [3] In general, they promiscuously catalyze several hydrolytic reactions with multiple endogenous and exogenous substrates (e.g. esters, thioesters, carbamates, and amides) into the corresponding carboxylic acid and alcohol products. [4] These products are usually more polar than the reagent, which results in a water solubility increase. This mechanism aids the elimination of ingested exogenous esters that might be potentially hazardous to the organism. [5] hCE1 and 2 display broad substrate specificity; however, hCE1 is more active towards substrates with a small alcohol group and a large acyl group, although sometimes it can accept structurally distinct alcohol moieties. [6] Conversely, the hCE2 enzyme recognizes substrates with large alcohol groups and small acyl groups, translating into higher substrate

specificity. [7]

Cocaine is the most studied CEs substrate, and its metabolism has been extensively analyzed in animal and human cells, both *in vitro* and *in vivo*. The hCE1 and 2 isoenzymes have different ester sites for cocaine hydrolysis. hCE1 catalyzes the hydrolysis of the methyl ester, resulting in benzoylecgonine (BzE) and methanol, while the hCE2 catalyzes the benzoyl hydrolysis to obtain ecgonine methyl ester (EME) and benzoic acid. Although EME and benzoic acid metabolites have a pharmacologically inactive profile, [8] the metabolite BzE, converted by hCE1, is a stronger vasoconstrictor than cocaine-like drugs (for example norcocaine and norepinephrine). [9] In addition, BzE is a zwitterion extremely difficult to hydrolyze in the body. The largest retention time of cocaine is responsible for its induced long-term toxicity. [10]

In the same line of hCE1 and 2 activity, the blood plasma butyrylcholinesterase was found to also catalyze the *in vivo* hydrolysis of cocaine in EME and benzoic acid. Nonetheless, the wild-type metabolic activity is considerably higher towards the biologically inactive enantiomer (+)-cocaine when compared to the naturally occurring (-)-cocaine. The (+)-enantiomer is quickly cleared from plasma (within seconds) before reaching the central nervous system (CNS), [11] while the active (-)-form remains in circulation for more than 45 minutes. This time-span is sufficient for the latter to reach the CNS, which has a

* Corresponding author.

E-mail addresses: pmrfigueiredo@cnc.uc.pt (P.R. Figueiredo), rgcunha@cnc.uc.pt (R.D. González), atpcarvalho@uc.pt (A.T.P. Carvalho).

maximum cocaine response profile within a minutes time frame. [12] Recently, this enzyme was been reengineered to perform (-)-cocaine detoxification, and its catalytic efficiency was increased more than 2000-fold. [13,14] Nonetheless, the short half-life of cocaine in the blood and the fast conversion to the hydrolysis-resistant metabolite (BzE) in the body limit the therapeutic applications of these enzymes. Therefore, hCE2 has the most desirable metabolic pathway considering that its products are less toxic and more easily excreted.

Mechanistically, the CEs hydrolytic reactions follow the classical ping-pong bi-bi reaction mechanism of serine hydrolases, via the formation of two tetrahedral intermediates and an acyl-serine enzyme intermediate. The catalytic triad is usually composed by the Ser-His-Asp/Glu residues, although other variations are known to occur. [7] The first hCE1 mechanism was proposed [15] as a single-step process, where the reaction moves from the reagents to the acyl monomer with a single first transition-state (TS), and the system resolved into the products with a second TS. Later, a second mechanism was reported [16] as a double-step mechanism with the classical ping-pong bi-bi reaction, where the stationary points are connected by four TSs. This mechanistic divergence is due to the different methodology used by these teams. Despite these variations, the first TS was pointed out as the rate-limiting step in both works, with free energy barriers close to the experimental results of $21.5 \text{ kcal mol}^{-1}$.

The hCE2 has been gathering interest for clinical settings applications on account of its ability to selectively activate and eliminate (pro) drugs; [17] for instance, a prodrug of gemcitabine (LY2334737) is only activated by hCE2. [18] Yet, the heterogeneity of the hCE2 N-glycan structures hinders its structural determination, so that only the hCE1 was structurally characterized by X-ray crystallography studies. [19,20] Understanding the detailed enzymatic mechanism of hCE2 can provide important insights for the design of novel pharmaceutical enzymes and their drug analogs activation (like the LY2334737).

Here, we performed conventional and accelerated MD simulations to identify the structural movements of the $\alpha 1$ and $\alpha 10'$ helices that act as a putative lid and calculated, using Quantum Mechanics/Molecular Mechanics methods, the one- (1D) and two- (2D) dimensional profiles for the complete catalytic mechanism of cocaine hydrolysis.

1. Methods

1.1. Modeling

A homology model of hCE2 was built using Modeller v9.24. [21] The crystal structure of hCE1 (PDB code 5A7F, 1.86 Å resolution monomer, UniprotKB code P23141) [22] was used as template and aligned with the "EST2_HUMAN" protein sequence (UniprotKB code O00748) [23] with the Clustal Omega server. [24] The hCE2 model protein sequence is presented in Fig. S1, presenting 46.8% sequence identity (71.1% similarity) with the hCE1. The built model was structurally compared with the predicted one by AlphaFold [25]. Both models present a $C\alpha$ root-mean-square deviation (RMSD) of 0.3 Å (excluding the N- and C-terminal ends, Fig. S1). The model was then subjected to 100 ns of conventional Molecular Dynamics (cMD) before any analysis, with the settings described in the MD section below.

1.2. Initial setup

The protonation state residues of the hCE2 model were assigned with MolProbity, considering an aqueous solution of pH=7.4 (values at which the kinetic parameters were measured) [26]. The INT-1 and INT-2 residues were geometry optimized in Gaussian09 [27] using B3LYP [28] with the 6-31G(d) basis set and a Polarizable Continuum Model solvent description. [29] The atomic partial charges were calculated from the optimized structure resorting to the Restrained Electrostatic Potential (RESP) method [30] from the HF/6-31G(d,p) single-point energy calculations.

1.3. Molecular docking

The initial position of the INT-1 was obtained by the flexible side-chain covalent docking method [31] and performed with the AutoDock4.2 suite of programs using the Lamarckian genetic algorithm (LGA). [32] A total of 150 runs were carried out. The population was set to 300, the maximum number of generations 27000, and the maximum number of energy evaluations 2500000.

1.4. Molecular Dynamics

The cMD simulations were performed using the Amber molecular dynamics program (AMBER18) [33] with the parm99SB [34] and GAFF [35] force fields for the protein and ligand, respectively. The structures were placed within an octahedral box of ≈ 13500 TIP3P waters [33] and 6 sodium ions were added to neutralize the system. The systems were then subjected to two initial energy minimizations and 500 ps of equilibration in an NVT ensemble using Langevin dynamics with small restraints on the protein of $10.0 \text{ kcal mol}^{-1}$ to heat the systems. Production simulations of 5 ns were carried out at 310.15 K in the NPT ensemble using Langevin dynamics with a collision frequency of 1 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. [36] The Particle Mesh Ewald (PME) method [37] was used to calculate electrostatic interactions with a cut-off distance of 10 Å. The accelerated Molecular Dynamics (aMD) was conducted for 50 ns and the parameters selected based upon the system size, average dihedral, and potential energy: $E_{\text{dih}}=7561.8 \text{ kcal mol}^{-1}$, $\alpha_{\text{dih}}=373.1 \text{ kcal mol}^{-1}$, $E_{\text{tot}}=-118, 535.1 \text{ kcal mol}^{-1}$ and $\alpha_{\text{tot}}=9476.0 \text{ kcal mol}^{-1}$. [38]

1.5. Principal components analysis and free energy landscape

Principal components analysis (PCA) was used to analyze the data from the aMD simulations, using the Bio3D package [39] (R statistical program). The protocol used was described in reference [40]. The analysis was performed to the $C\alpha$ of the total 8287 system atoms. The PDB and trajectory structures were fitted and superposed to remove translations and rotations. The variance of the data in each principal component was evaluated by calculating eigenvectors and the corresponding eigenvalues. To evaluate on average how much each residue moves during the trajectory, we calculated the root mean square fluctuations (RMSF). Free energy landscape (FEL) calculations were performed on the CPPTRAJ module from AMBER18, [33] followed by elastic-network driven/Brownian Dynamics Importance Sampling (eBDIMS) to evaluate conformation transition pathways, calculated in the eBDIMS server (default parameters).

1.6. Quantum Mechanics/Molecular Mechanics calculations

The Quantum Mechanics/Molecular Mechanics (QM/MM) calculations [41,42] were performed using the internal semi-empirical hybrid QM/MM functionality implemented in AMBER18 [33] with periodic boundary conditions. The last MD structure of each complex was used as the initial structure for the QM/MM calculations. The PM6 semi-empirical method [43,44] was employed to describe the QM region while the MM region was described by the parm99SB force field. [34] The reactions were conducted at 310.15 K. [26] Electrostatic embedding [45] was employed and the boundary treated via the link atom approach. Long-range electrostatic interactions were described by an adapted implementation of the PME method [37] for QM/MM. [46] For the 1D and 2D-PMFs the same set of reaction coordinates was used (Fig. 1): i) the distance between the H448 proton He and the S201 oxygen atom (O_{Ser}, θ) and ii) the distance between the H448 proton He and the substrate oxygen atom (O_{ee}, ϕ).

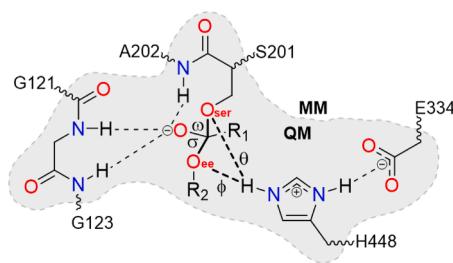


Fig. 1. Division of the QM/MM system: all atoms inside the shaded shape belong to the QM subsystem in the QM/MM calculations. All other atoms belong to the MM subsystem. The link atoms are represented with a wavy bond. INT-1 $R_1=Ph$, $R_2=EME$, INT-2 $R_1=Ph$, $R_2=H$.

In the 1D-PMF, one reaction coordinate was incrementally restrained in steps of 0.1 Å using the umbrella sampling method, except near the transition states where smaller steps of 0.02 Å were employed. Then, and in accordance with our previous work, [47] the structures from the 1D profiles were used as starting points for the 2D-PMF scans. To keep the reaction coordinates at the requested distances and to ensure enough overlap between windows, an umbrella constraint force of 200 kJ. mol⁻¹.Å⁻² was used along the reaction path and incrementally increased to 1000 kJ.mol⁻¹.Å⁻² as the reaction coordinates deflect from the minimum energy path. For every window, a total of 15 ps was simulated with a time step of 1 fs. The distances were scanned from 1.4 to 2.5 Å (ω and σ , excepting the third step, where the scanned reaction vary from 0.9 to 2.5 Å) and from 1.0 to 2.0 Å (θ and ϕ) comprising a total of 132 simulations for first, second and fourth steps, and 144 simulations for the third step. The PMFs were computed resorting to the Weighted Histogram Analysis Method (WHAM) with the Monte Carlo bootstrap error analysis [48,49] and the minimum energy path traced resorting to the MEPSA v1.4. [50]

1.7. High-level layer corrections

For the high-level (HL) layer corrections of the 2D-PMFs, we resorted to the method described in detail in reference [51], which is based on the work of Truhlar and co-workers. [52,53] Multiple structures were retrieved from the 2D-PMFs and subjected to single-point calculations using the Gaussian09 [27] program, first with the semi-empirical PM6 [43,44] and then with B3LYP with 6–31G(d,p), 6–311++G(d,p) [28] and Grimme D3 dispersion [54]. The corrected energies term (E) was interpolated from those structures employing the following equation:

$$E = E_{\text{MM}}^{HL} + S(\Delta E_{LL}^{HL})$$

Where the term ΔE_{LL}^{HL} corresponds to the difference between the free energies for the HL layer set, calculated with B3LYP-D3/6–31G(d,p) or 6–311++G(d,p) [28,54] and PM6 [43,44] (low level (LL) theory), while S represents the cubic spline function of the difference between the HL and LL theory representing the QM region. The PMFs corrected with the 6–31G(d,p) [28] basis set are presented in Supporting Information (Fig. S2 and S3), while the PMFs corrected with the 6–311++G(d,p) basis set [54] are featured herein. Unless otherwise stated, all mentioned energetic barriers concern the ones calculated with B3LYP-D3/6–311++G(d,p). [28,54]

2. Results and discussion

2.1. Structure of the hCE2 enzyme

The hCE2 has the typical α/β hydrolase-fold of the serine hydrolase superfamily. The active site is located in the catalytic domain and surrounded by the α/β and regulatory domains (Fig. 2), containing the classical Ser-His-Glu catalytic triad. The serine residue (S201) acts as the

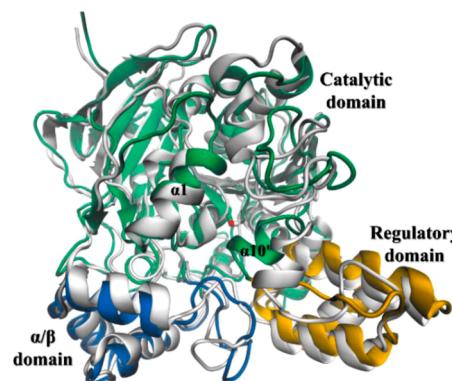


Fig. 2. Superimposition of the cMD (after 100 ns, colored) and aMD (active conformation cluster, silver) structures of hCE2. The catalytic domain is shown in green, the regulatory domain in yellow, and the α/β domain in blue. The active site, where the sidechain of the catalytic serine is represented, is capped by two mobile α -helices ($\alpha 1$ and $\alpha 10'$) that act as a putative lid (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

nucleophile in the acylation step, while the histidine (H448) function as an acid/base (responsible for proton transfer between the catalytic serine and the substrate) and is stabilized by a nearby glutamate residue (E334). The oxyanion hole region is composed by the backbone amides of G122, G123, and A202. [55] This spatial arrangement of hydrogen bond donors is responsible for the stabilization of a negative charge that is developed during the catalytic mechanism on the oxygen atom of the tetrahedral intermediates (INT-1 and INT-2). Similar to the isoenzyme 1, the hCE2 displays two main pockets (acyl and alcohol) that accommodate the substrate in the active site. [56] The main residues that line these two pockets are illustrated in Fig. S4. Resembling to what has been determined for the hCE1, [57] the catalytic triad and oxyanion hole residues of the hCE2 subdivides the active site into two regions with different polarities. The acyl pocket is mainly surrounded by hydrophobic residues (most of them Leucines) that may stabilize the acyl unit of the substrate by Van der Walls (VdW) interactions. On the other hand, the alcohol pocket is less hydrophobic, presenting more aromatic and basic residues.

The active site is capped with two α -helices ($\alpha 1$ and $\alpha 10'$, Fig. 2) that mediate the substrate access to the active site by switching between open/closed conformations. [58] We have performed aMD simulations to obtain the open conformation. In this kind of simulation, a potential energy term is added that allows overcoming the free energy barriers in the protein landscape and accessing rare events. [59] During the 50 ns aMD simulation performed, we observed that $\alpha 1$ and $\alpha 10'$ helices move considerably (Fig. 2 and S2). The largest movement was in the $\alpha 10'$ helix, with a value of up to 4.4 Å, while the $\alpha 1$ helix only varies up to 2.4 Å (Fig. S5).

The 50 ns aMD trajectory was clustered into open and closed conformations (Fig. S6). The open conformation, which allows the substrate to enter the active site, was selected for docking of the INT-1 structure.

To evaluate the main movements throughout the simulations we performed PCA analysis. It is based on a mathematical covariance-matrix-based technique capable of capturing the divergence of coordinates during the simulation, taking into account a reference structure. The three main principal components identified (PC1, PC2, and PC3) account for around 74% of the enzyme's motions (Fig. S7). The analysis was performed to the C α of the enzyme, primarily scanning the whole protein landscape, which resulted in low overall representation. The modeled nature of the protein could be inducing the N- and C-termini to influence the results. PCA was then adjusted to the residues comprising both $\alpha 1$ and $\alpha 10'$ regions alone, corresponding to the above-mentioned value. We considered that the $\alpha 10'$ movement would lead to

movement of the regulatory domain; however, the helices seemed to be the main contributors for the protein movement, as the inclusion of the regulatory domain in the analysis decreased PCA to $\approx 60\%$. PCA output was then used to calculate the free energy landscape (FEL, Fig. S8).

To further evaluate the nature of the motions, we analyzed the transition pathways in the webserver eBDIMS, which relies on the essential dynamics-refined elastic network model force field. The platform iteratively performs model evaluations and defines a starting structure for the projection by overlapping the two best-score normal modes (NMs) with the modeled transition. For our system, the NMs represent $\approx 61\%$ of the movements involved, with an initial RMSD of 3.40 Å and a final RMSD of 0.34 and 0.32 Å for forward and reverse transitions, respectively (Fig. S9).

With an open conformation established, we subjected the system to the flexible side-chain covalent docking method. This covalent method differs from the “free” counterpart by anchoring the backbone of the functionalized serine residue. The remaining bonds that are capable to rotate are set free to identify the best position. This placed the benzoyl group in the acyl pocket and the EME group in the alcohol pocket with a binding energy of $-7.86 \text{ kcal mol}^{-1}$. The system was then minimized through cMD, in order to achieve the local minima. The final structure was then selected to perform QM/MM calculations to study the catalytic mechanism of (-)-cocaine hydrolysis.

2.2. First reaction step: release of ecgonine methyl ester

The first half part of the (-)-cocaine hydrolysis catalytic cycle (also known as the acylation step) concerns the release of the EME to generate the enzyme-activated monomer (EAM, Fig. 3).

The reaction mechanism begins with the entry of a (-)-cocaine molecule to the enzyme’s active site pocket. The ecgonine ring and the methoxy group of cocaine make VdW interactions inside the pocket that accommodate this portion of the molecule: hydrophobic residues L77, F81, and L452 interact with the ring; W337 and P441 with the methoxy group (Fig. S10). In the RC structure (2.9 Å, Fig. 4A), a hydrogen bond is formed between the A202 amide hydrogen of the oxyanion hole and the carbonyl oxygen of the substrate (2.1 Å distance). Proton transfer to the H448 residue (1.8 Å and 155.1°, Fig. 4A) occurs simultaneously with the nucleophilic attack by the S201 oxygen atom (O_{Ser}) on the benzoyl carbonyl carbon atom of cocaine. The first transition state (TS_1) has a free energy barrier (ΔG^\ddagger) of $9.7 \text{ kcal mol}^{-1}$ and is resolved into the INT-1 structure, which is $3.4 \text{ kcal mol}^{-1}$ below TS_1 (Fig. 3B and S3). In the INT-1 structure (Fig. 4C), the histidine residue is positively charged and stabilized by the glutamate residue (E334). The oxyanion hole backbones amides of G122, G123, and A202 are now stabilizing the developed negative charge on the tetrahedral INT-1 oxygen atom (the hydrogen bond distance is 1.8 Å with the glycine residues and 1.9 Å for the alanine residue).

The reaction proceeds with the release of the EME (Fig. 3) to form the EAM structure. For that, the positively charged histidine needs to transfer its proton to the alcohol oxygen atom of the EME group (O_{ee} , Fig. 4C). The TS_2 is $8.7 \text{ kcal mol}^{-1}$ above the INT-1, and the EAM is $9.2 \text{ kcal mol}^{-1}$ below TS_2 (Fig. 4D and S3).

The overall $TS_2 \Delta G^*$ of $15.0 \text{ kcal mol}^{-1}$ reflects the higher distance of

the HE2 of H448 towards the ecgonine oxygen (O_{ee}) when compared to the serine’s oxygen (O_{Ser}). Also, the angles are far from ideal to move the reaction forward ($\beta=124.9^\circ$ and $\gamma=164.7^\circ$, Fig. 4C) and the reverse reaction is favored with a lower free energy barrier ($3.4 \text{ kcal mol}^{-1}$, Fig. 4B and S3).

2.3. Second reaction step: release of benzoic acid

The second half part of the (-)-cocaine hydrolysis cycle (also known as the deacylation step) concerns the release of benzoic acid to generate the product complex (PC, Fig. 5).

The active site pocket with the EAM structure contains a large number of water molecules (Fig. 4E). The nucleophilic attack of a water molecule to the covalently bonded substrate carbonyl carbon (2.4 Å, Fig. 4E) generates the INT-2 structure. The energy barrier for this reaction (TS_3 , Fig. 4F and S3) amounts to only $3.8 \text{ kcal mol}^{-1}$.

Alike the INT-1 structure, the oxyanion hole oxygen atom of the INT-2 structure makes three hydrogen bonds with the backbone amide groups of G122, G123, and A202 (the hydrogen bond distances are 1.6, 1.9, and 1.7 Å, respectively). These interactions stabilize the negative charge developed in this oxygen atom. The H448 residue (Fig. 4G) is at a similar distance to both oxygen atoms that move the reaction forward ($O_{\text{Ser}}, 2.2 \text{ Å}$) or backward ($O_{\text{WAT}}, 1.9 \text{ Å}$). The INT-2 (Fig. 4G) is exergonic in relation to the EAM with a ΔG of $-13.1 \text{ kcal mol}^{-1}$. Then, the system evolves towards the release of the benzoic acid molecule (PC, Fig. 5). The final transition state (TS_4 , Fig. 4H) has a ΔG^\ddagger of $19.5 \text{ kcal mol}^{-1}$ (Fig. 4H and S3). The PC is $12.7 \text{ kcal mol}^{-1}$ above the INT-2 (Fig. 3D and S3) and the enzyme ready for another turnover of cocaine hydrolysis.

The reported turnover number (k_{cat}) by Potter and co-workers [26] is 0.0889 s^{-1} for the hydrolysis of (-)-cocaine to afford EME and benzoic acid. According to Eyring’s equation, [60] this corresponds to a free energy of about $19.7 \text{ kcal mol}^{-1}$ at 37°C . Our simulations suggest that the rate-limiting step is the formation of benzoic acid (TS_4) with a ΔG^\ddagger of $19.5 \text{ kcal mol}^{-1}$, which is in excellent agreement with this value.

3. Conclusion

In this work, we have studied the helices movement that cap the active site by performing conventional and accelerated Molecular Dynamics simulations. Our structural analysis indicates that the $\alpha 1$ and $\alpha 10'$ helices motion warrants an open conformation of the enzyme (allowing the substrate to enter into the active site). We observe a higher displacement of the $\alpha 10'$ helix when compared to the $\alpha 1$, which is highlighted in the PCA and FEL results, and indicative of the role of both helices in the overall protein movements.

The active site functionalized serine intermediate was placed in the open conformation structure resorting to the flexible side-chain covalent docking method. The benzoyl and the ecgonine groups were placed in the acyl and alcohol pockets, respectively, with $-7.86 \text{ kcal mol}^{-1}$ of binding energy. We used this structure for the cMD and then to the QM/MM studies. Our calculated 2D-QM/MM profiles on the full catalytic cycle showed that the acylation step requires $15.0 \text{ kcal mol}^{-1}$ (TS_2) to move the reaction forward (the TS_1 is $9.7 \text{ kcal mol}^{-1}$ above the RC). This may be related to proton transfer that occurs concomitantly with CO

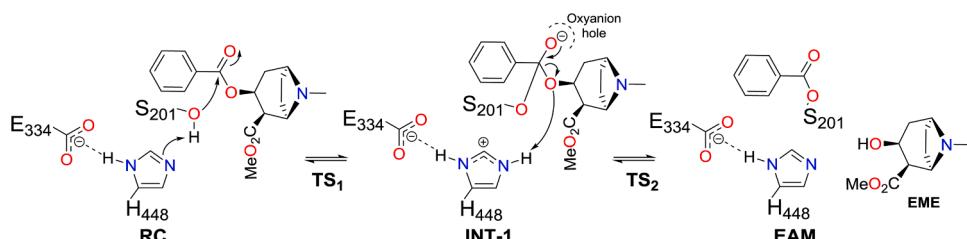


Fig. 3. First half part of the reaction mechanism for the hCE2 (-)-cocaine hydrolysis: release of EME.

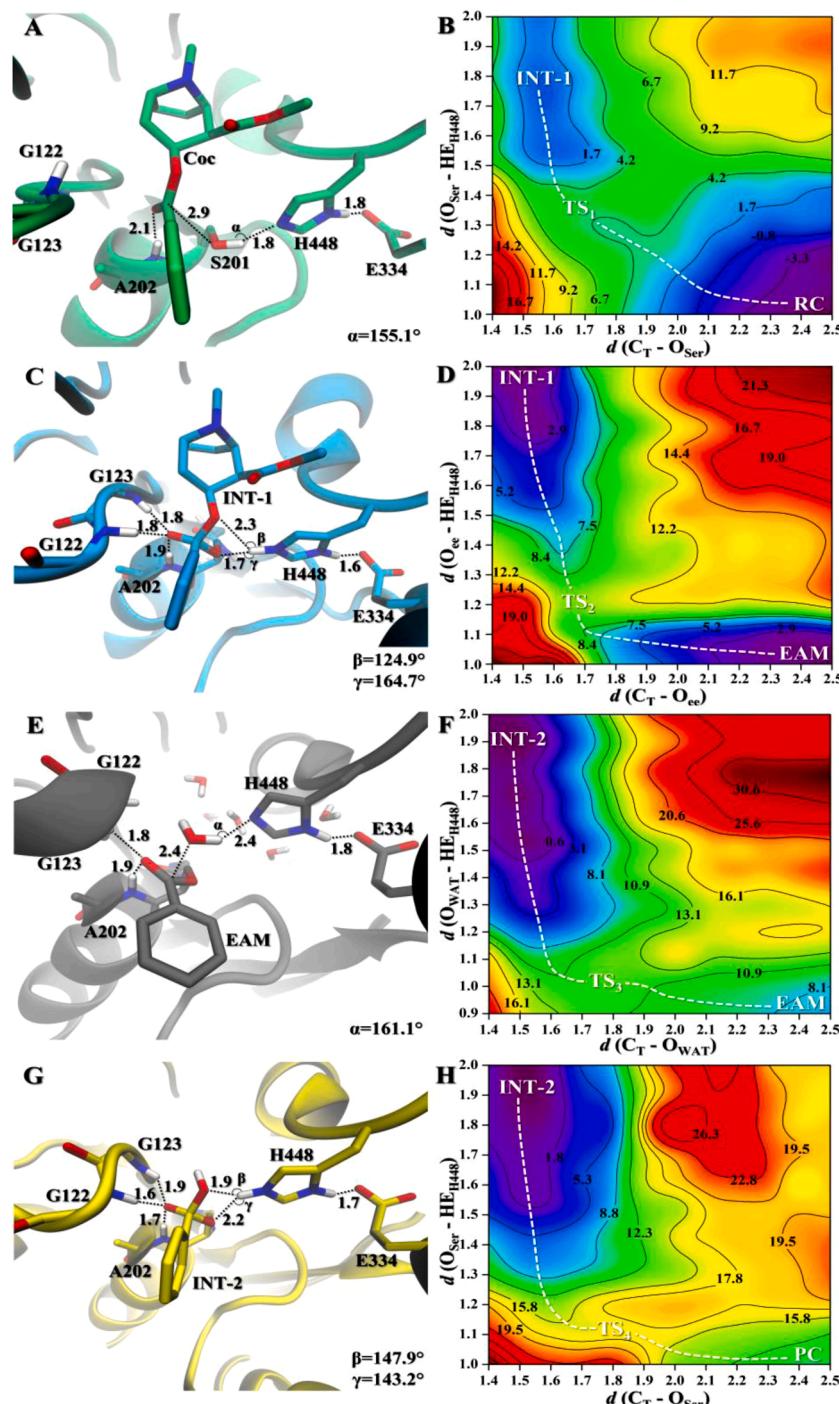


Fig. 4. Active site pocket reference structures of the lowest energy stationary points RC, INT-1, EAM and INT-2 optimized with PM6/parm99SB (A, C, E and G, respectively). 2D-PMFs for the acylation (B and D) and deacylation (F and H) steps calculated with B3LYP-D3/6-311++G(d,p) [28,54]. Key distances are given in Å and the relative energy values in kcal mol⁻¹.

bond formation/breaking: a combination of the higher distance of the HE2 atom of H448 towards the ecgonine oxygen (O_{ee}) in comparison to the serine's oxygen (O_{Ser}), expressing better positioned angles that favor the reverse reaction (TS₁). In the deacylation step, the TS₄ (formation of benzoic acid) was identified as the rate-limiting step of the (-)-cocaine hydrolysis by the hCE2. This step requires 19.5 kcal mol⁻¹ to be achieved, being the higher TS in the whole cycle and in agreement with the reported experimental barrier.

Comparing the PMFs corrected with either 6-31G(d,p) or 6-311++G(d,p) basis sets, an increase in the energetic barriers is observed in the

most complete basis set. In the acylation step, the TS₁ and INT-1 structures increased 2.9 kcal mol⁻¹, whereas the TS₂ presents the biggest increase of 5.1 kcal mol⁻¹. The acylation step has now a free energy barrier of 15.0 kcal mol⁻¹ associated to the TS₂. While the deacylation step stationary points INT-2 and PC have the same energy in both basis sets, the TS₃ and TS₄ increased their energetic barriers 1.7 and 1.2 kcal mol⁻¹, respectively. This placed the TS₄ in a closer energy value with the experimental barrier (19.7 kcal mol⁻¹) when compared to the PMFs corrected with the 6-31G(d,p) basis set (18.3 vs 19.5 kcal mol⁻¹).

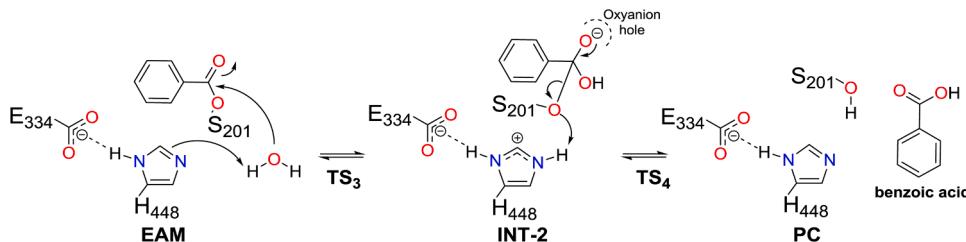


Fig. 5. Second half part of the reaction mechanism for the hCE2 (-)-cocaine hydrolysis: release of benzoic acid.

Since the hCE1 yields BzE (that is extremely difficult to hydrolyze) and the WT blood plasma butyrylcholinesterase is more active towards the biologically inactive cocaine enantiomer, the hCE2 has the most desirable metabolic pathway. Its metabolites are less toxic and more easily excreted. These novel mechanistic insights about the hydrolysis of (-)-cocaine by the hCE2 are expected to benefit the discovery of new enhanced enzyme variants that are capable to better hydrolyze (-)-cocaine.

CRediT authorship contribution statement

Pedro R. Figueiredo: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing. **Ricardo D. González:** Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Alexandra T.P. Carvalho:** Conceptualization, Validation, Writing – original draft, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare no competing financial interests.

Acknowledgements

This work was sponsored through grants SFRH/BD/144303/2019, 2020.10114.BD and IF/01272/2015, and project UIDB/04539/2020 financed by Portuguese National funds via FCT – Fundação para a Ciência e a Tecnologia and Regional Operational Program Centro (CENTRO2020) under the Portuguese Partnership Agreement 2020 by the European Regional Development Fund (ERDF). The authors acknowledge the support of the National Distributed Computing Infrastructure (INCD) funded by FCT and FEDER with computational resources, under the projects 01/SAICT/2016 n° 022153 and Advanced Computing Project CPC/A2/4568/2020.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.mcat.2021.111938](https://doi.org/10.1016/j.mcat.2021.111938).

References

- [1] T. Satoh, M. Hosokawa, Structure, function and regulation of carboxylesterases, *Chem. Biol. Interact.* 162 (2006) 195–211, <https://doi.org/10.1016/j.cbi.2006.07.001>.
- [2] T. Satoh, P. Taylor, W.F. Bosron, S.P. Sanghani, M. Hosokawa, B.N.L. Du, Current Progress on Esterases: from Molecular Structure to Function, *Drug. Metab. Dispos.* 30 (2002) 488–493, <https://doi.org/10.1124/dmd.30.5.488>.
- [3] M.K. Ross, J.A. Crow, Human carboxylesterases and their role in xenobiotic and endobiotic metabolism, *J. Biochem. Mol. Toxicol.* 21 (2007) 187–196, <https://doi.org/10.1002/jbt.20178>.
- [4] S.P. Sonal, S.C. Paresh, S.A. Marissa, B.F. William, Human Carboxylesterases: an Update on CES1, CES2 and CES3, *Protein Pept. Lett.* 16 (2009) 1207–1214.
- [5] S. Casey Laizure, V. Herring, Z. Hu, K. Witbrodt, R.B. Parker, The Role of Human Carboxylesterases in Drug Metabolism: have We Overlooked Their Importance? *Pharmacotherapy* 33 (2013) 210–222, <https://doi.org/10.1002/phar.1194>.
- [6] M. Hosokawa, Structure and Catalytic Properties of Carboxylesterase Isozymes Involved in Metabolic Activation of Prodrugs, *Molecules* 13 (2008) 412–431, <https://doi.org/10.3390/molecules13020412>.
- [7] D. Wang, L. Zou, Q. Jin, J. Hou, G. Ge, L. Yang, Human carboxylesterases: a comprehensive review, *Acta Pharm. Sin. B* 8 (2018) 699–712, <https://doi.org/10.1016/j.apsb.2018.05.005>.
- [8] E.T. Shimomura, G.F. Jackson, B.D. Paul, Cocaine, Crack Cocaine, and Ethanol, in: critical Issues in Alcohol and Drugs of Abuse Testing, Elsevier (2019) 215–224, <https://doi.org/10.1016/B978-0-12-815607-0.00017-4>.
- [9] N.J. Connors, R.S. Hoffman, Experimental Treatments for Cocaine Toxicity: a Difficult Transition to the Bedside, *J. Pharmacol. Exp. Ther.* 347 (2013) 251–257, <https://doi.org/10.1124/jpet.113.206383>.
- [10] R. Pal, M. Megharaj, K.P. Kirkbride, R. Naidu, Illicit drugs and the environment — A review, *Sci. Total Environ.* 463–464 (2013) 1079–1092, <https://doi.org/10.1016/j.scitotenv.2012.05.086>.
- [11] S.J. Gatley, R.R. MacGregor, J.S. Fowler, A.P. Wolf, S.L. Dewey, D.J. Schlyer, Rapid Stereoselective Hydrolysis of (+)-Cocaine in Baboon Plasma Prevents Its Uptake in the Brain: implications for Behavioral Studies, *J. Neurochem.* 54 (1990) 720–733, <https://doi.org/10.1111/j.1471-4159.1990.tb01933.x>.
- [12] S.J. Gatley, Activities of the enantiomers of cocaine and some related compounds as substrates and inhibitors of plasma butyrylcholinesterase, *Biochem. Pharmacol.* 41 (1991) 1249–1254, [https://doi.org/10.1016/0006-2952\(91\)90665-R](https://doi.org/10.1016/0006-2952(91)90665-R).
- [13] F. Zheng, L. Xue, S. Hou, J. Liu, M. Zhan, W. Yang, C.-G. Zhan, A highly efficient cocaine-detoxifying enzyme obtained by computational design, *Nat. Commun.* 5 (2014) 3457, <https://doi.org/10.1038/ncomms4457>.
- [14] K.E. Larimore, I.C. Kazan, L. Kannan, R.P. Kendle, T. Jamal, M. Marcus, A. Bolia, S. Brimijoin, C.-G. Zhan, S.B. Ozkan, T.S. Mor, Plant-expressed cocaine hydrolase variants of butyrylcholinesterase exhibit altered allosteric effects of cholinesterase activity and increased inhibitor sensitivity, *Sci. Rep.* 7 (2017) 10419, <https://doi.org/10.1038/s41598-017-10571-z>.
- [15] J. Yao, X. Chen, F. Zheng, C.-G. Zhan, Catalytic Reaction Mechanism for Drug Metabolism in Human Carboxylesterase-1: cocaine Hydrolysis Pathway, *Mol. Pharmaceutics*. 15 (2018) 3871–3880, <https://doi.org/10.1021/acs.molpharmaceut.8b00354>.
- [16] M. Yan, Z. Zhang, Z. Liu, C. Zhang, J. Zhang, S. Fan, Z. Yang, Catalytic Hydrolysis Mechanism of Cocaine by Human Carboxylesterase 1: an Orthoester Intermediate Slows Down the Reaction, *Molecules* 24 (2019) 4057, <https://doi.org/10.3390/molecules24224057>.
- [17] J. Lamego, B. Cunha, C. Peixoto, M.F. Sousa, P.M. Alves, A.L. Simplício, A. S. Coradinha, Carboxylesterase 2 production and characterization in human cells: new insights into enzyme oligomerization and activity, *Appl. Microbiol. Biotechnol.* 97 (2013) 1161–1173, <https://doi.org/10.1007/s0253-012-3994-3>.
- [18] S.E. Pratt, S. Durland-Busbice, R.L. Shepard, K. Heinz-Taheny, P.W. Iversen, A. H. Dantzig, Human Carboxylesterase 2 Hydrolyzes the Prodrug of Gemcitabine (LY2334737) and Confers Prodrug Sensitivity to Cancer Cells, *Clin. Cancer. Res.* 19 (2013) 1159–1168, <https://doi.org/10.1158/1078-0432.CCR-12-1184>.
- [19] S. Bencharit, C.L. Morton, J.L. Hyatt, P. Kuhn, M.K. Danks, P.M. Potter, M. R. Redinbo, Crystal Structure of Human Carboxylesterase 1 Complexed with the Alzheimer's Drug Tacrine, *Chem. Biol.* 10 (2003) 341–349, [https://doi.org/10.1016/S1074-5521\(03\)00071-1](https://doi.org/10.1016/S1074-5521(03)00071-1).
- [20] C.D. Fleming, C.C. Edwards, S.D. Kirby, D.M. Maxwell, P.M. Potter, D.M. Cerasoli, M.R. Redinbo, Crystal Structures of Human Carboxylesterase 1 in Covalent Complexes with the Chemical Warfare Agents Soman and Tabun, *Biochemistry* 46 (2007) 5063–5071, <https://doi.org/10.1021/bi700246n>.
- [21] B. Webb, A. Sali, Comparative Protein Structure Modeling Using MODELLER, *Curr. Protoc. Bioinformatics*. 54 (2016) 5.6.1–5.6.37, <https://doi.org/10.1002/cpb.3>.
- [22] V.A. de Souza, D.J. Scott, J.E. Nettleship, N. Rahaman, M.H. Charlton, M.A. Walsh, R.J. Owens, Comparison of the structure and activity of Glycosylated and Aglycosylated Human Carboxylesterase 1, *PLoS One* 10 (2015), e0143919, <https://doi.org/10.1371/journal.pone.0143919>.
- [23] E.V. Pindel, N.Y. Kedishvili, T.L. Abraham, M.R. Brzezinski, J. Zhang, R.A. Dean, W.F. Bosron, Purification and Cloning of a Broad Substrate Specificity Human Liver Carboxylesterase That Catalyzes the Hydrolysis of Cocaine and Heroin, *J. Biol. Chem.* 272 (1997) 14769–14775, <https://doi.org/10.1074/jbc.272.23.14769>.
- [24] F. Madeira, Y. mi Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R.N. Tivey, S.C. Potter, R.D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019, *Nucleic Acids Res* 47 (2019) W636–W641, <https://doi.org/10.1093/nar/gkz268>.
- [25] J. Juniper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler,

- T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, *Nature* (2021), <https://doi.org/10.1038/s41586-021-03819-2>.
- [26] M. Hatfield, L. Tsurkan, J. Hyatt, X. Yu, C. Edwards, L. Hicks, R. Wadkins, P. Potter, Biochemical and molecular analysis of carboxylesterase-mediated hydrolysis of cocaine and heroin: carboxylesterase hydrolysis of cocaine and heroin, *Br. J. Pharmacol.* 160 (2010) 1916–1928, <https://doi.org/10.1111/j.1476-5381.2010.00700.x>.
- [27] M.J. Frisch, G.W. Trucks, J.R. Cheeseman, G. Scalfani, M. Caricato, H. P. Hratchian, X. Li, V. Barone, J. Bloino, G. Zheng, T. Vreven, J.A. Montgomery, G. A. Petersson, G.E. Scuseria, H.B. Schlegel, H. Nakatsuji, A.F. Izmaylov, R.L. Martin, J.L. Sonnenberg, J.E. Peralta, J.J. Heyd, E. Brothers, F. Ogliaro, M. Bearpark, M. A. Robb, B. Mennucci, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, A. Rendell, R. Gomperts, V.G. Zakrzewski, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, Gaussian 09, Revision B.01, Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT, 2009.
- [28] C.S. Ashvar, F.J. Devlin, K.L. Bak, P.R. Taylor, P.J. Stephens, Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra: 6,8-Dioxabicyclo[3.2.1]octane, *J. Phys. Chem.* 100 (1996) 9262–9270, <https://doi.org/10.1021/jp953738p>.
- [29] J. Tomasi, B. Mennucci, R. Cammi, Quantum Mechanical Continuum Solvation Models, *Chem. Rev.* 105 (2005) 2999–3094, <https://doi.org/10.1021/cr9904009>.
- [30] C.I. Bayly, P. Cieplak, W. Cornell, P.A. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model, *J. Phys. Chem.* 97 (1993) 10269–10280, <https://doi.org/10.1021/j100142a004>.
- [31] G. Bianco, S. Forli, D.S. Goodsell, A.J. Olson, Covalent docking using autodock: two-point attractor and flexible side chain methods: covalent Docking with AutoDock, *Protein Sci.* 25 (2016) 295–301, <https://doi.org/10.1002/pro.2733>.
- [32] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A. J. Olson, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, *J. Comput. Chem.* 30 (2009) 2785–2791, <https://doi.org/10.1002/jcc.21256>.
- [33] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 3 (2013) 198–210, <https://doi.org/10.1002/wcms.1121>.
- [34] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins* 65 (2006) 712–725, <https://doi.org/10.1002/prot.21123>.
- [35] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, *J. Comput. Chem.* 25 (2004) 1157–1174, <https://doi.org/10.1002/jcc.20035>.
- [36] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.* 23 (1977) 327–341, [https://doi.org/10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5).
- [37] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems, *J. Chem. Phys.* 98 (1993) 10089–10092, <https://doi.org/10.1063/1.464397>.
- [38] J. Wereszczynski, J.A. McCammon, Accelerated Molecular Dynamics in Computational Drug Design, in: R. Baron (Ed.), Computational Drug Discovery and Design, Springer New York, New York, NY, 2012, pp. 515–524, https://doi.org/10.1007/978-1-61779-465-0_30.
- [39] B.J. Grant, L. Skjerven, X. Yao, The Bio3D packages for structural bioinformatics, *Protein Sci.* 30 (2021) 20–30, <https://doi.org/10.1002/pro.3923>.
- [40] P.C. Gedeon, J.R. Thomas, J.D. Madura, Accelerated Molecular Dynamics and Protein Conformational Change: a Theoretical and Practical Guide Using a Membrane Embedded Model Neurotransmitter Transporter, in: A. Kukol (Ed.), Molecular Modeling of Proteins, Springer New York, New York, NY, 2015, pp. 253–287, https://doi.org/10.1007/978-1-4939-1465-4_12.
- [41] A. Warshel, M. Levitt, Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, *J. Mol. Biol.* 103 (1976) 227–249, [https://doi.org/10.1016/0022-2836\(76\)90311-9](https://doi.org/10.1016/0022-2836(76)90311-9).
- [42] A.T.P. Carvalho, A. Barrozo, D. Doron, A.V. Kilshtain, D.T. Major, S.C.L. Kamerlin, Challenges in computational studies of enzyme structure, function and dynamics, *J. Mol. Graph. Model.* 54 (2014) 62–79, <https://doi.org/10.1016/j.jmgm.2014.09.003>.
- [43] G. Jindal, A. Warshel, Exploring the Dependence of QM/MM Calculations of Enzyme Catalysis on the Size of the QM Region, *J. Phys. Chem. B.* 120 (2016) 9913–9921, <https://doi.org/10.1021/acs.jpcb.6b07203>.
- [44] J.J.P. Stewart, Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements, *J. Mol. Model.* 13 (2007) 1173–1213, <https://doi.org/10.1007/s00894-007-0233-4>.
- [45] D. Bakowies, W. Thiel, Hybrid Models for Combined Quantum Mechanical and Molecular Mechanical Approaches, *J. Phys. Chem.* 100 (1996) 10580–10594, <https://doi.org/10.1021/jp9536514>.
- [46] K. Nam, J. Gao, D.M. York, An Efficient Linear-Scaling Ewald Method for Long-Range Electrostatic Interactions in Combined QM/MM Calculations, *J. Chem. Theory Comput.* 1 (2005) 2–13, <https://doi.org/10.1021/ct049941i>.
- [47] P.R. Figueiredo, B.C. Almeida, D.F.A.R. Dourado, A.F. Sousa, A.J.D. Silvestre, A.T. P. Carvalho, Enzymatic Synthesis of Poly(caprolactone): a QM/MM Study, *ChemCatChem* 12 (2020) 4845–4852, <https://doi.org/10.1002/cctc.202000780>.
- [48] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J. Comput. Chem.* 13 (1992) 1011–1021, <https://doi.org/10.1002/jcc.540130812>.
- [49] A. Grossfield, An Implementation of WHAM: The Weighted Histogram Analysis Method, v2.0.9 (2018), http://membrane.urmc.rochester.edu/wordpress/?page_id=126.
- [50] I. Marcos-Alcalde, J. Setoain, J.I. Mendieta-Moreno, J. Mendieta, P. Gómez-Puertas, MEPSA: minimum energy pathway analysis for energy landscapes, *Bioinformatics* (2015) btv453, <https://doi.org/10.1093/bioinformatics/btv453>.
- [51] J.J. Ruiz-Pernía, E. Silla, I. Tunón, S. Martí, V. Moliner, Hybrid QM/MM Potentials of Mean Force with Interpolated Corrections, *J. Phys. Chem. B.* 108 (2004) 8427–8433, <https://doi.org/10.1021/jp049633g>.
- [52] Y.-Y. Chuang, J.C. Corchado, D.G. Truhlar, Mapped Interpolation Scheme for Single-Point Energy Corrections in Reaction Rate Calculations and a Critical Evaluation of Dual-Level Reaction Path Dynamics Methods, *J. Phys. Chem. A.* 103 (1999) 1140–1149, <https://doi.org/10.1021/jp9842493>.
- [53] K.A. Nguyen, I. Rossi, D.G. Truhlar, A dual-level Shepard interpolation method for generating potential energy surfaces for dynamics calculations, *J. Chem. Phys.* 103 (1995) 5522–5530, <https://doi.org/10.1063/1.470536>.
- [54] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu, *J. Chem. Phys.* 132 (2010), 154104, <https://doi.org/10.1063/1.3382344>.
- [55] M. Alam, D.E. Vance, R. Lehner, Structure–Function Analysis of Human Triacylglycerol Hydrolase by Site-Directed Mutagenesis: identification of the Catalytic Triad and a Glycosylation Site, *Biochemistry* 41 (2002) 6679–6687, <https://doi.org/10.1021/bi0255625>.
- [56] G. Vistoli, A. Pedretti, A. Mazzolari, B. Testa, Homology modeling and metabolism prediction of human carboxylesterase-2 using docking analyses by GriDock: a parallelized tool based on AutoDock 4.0, *J. Comput. Aided Mol. Des.* 24 (2010) 771–787, <https://doi.org/10.1007/s10822-010-9373-1>.
- [57] G. Vistoli, A. Pedretti, A. Mazzolari, B. Testa, In silico prediction of human carboxylesterase-1 (hCES1) metabolism combining docking analyses and MD simulations, *Bioorg. Med. Chem.* 18 (2010) 320–329, <https://doi.org/10.1016/j.bmc.2009.10.052>.
- [58] J. Lian, R. Nelson, R. Lehner, Carboxylesterases in lipid metabolism: from mouse to human, *Protein Cell* 9 (2018) 178–195, <https://doi.org/10.1007/s13238-017-0437-z>.
- [59] L.C.T. Pierce, R. Salomon-Ferrer, C. Augusto F. de Oliveira, J.A. McCammon, R. C. Walker, Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics, *J. Chem. Theory Comput.* 8 (2012) 2997–3002, <https://doi.org/10.1021/ct300284c>.
- [60] M.G. Evans, M. Polanyi, Some applications of the transition state method to the calculation of reaction velocities, especially in solution, *Trans. Faraday Soc.* 31 (1935) 875, <https://doi.org/10.1039/tf9353100875>.