

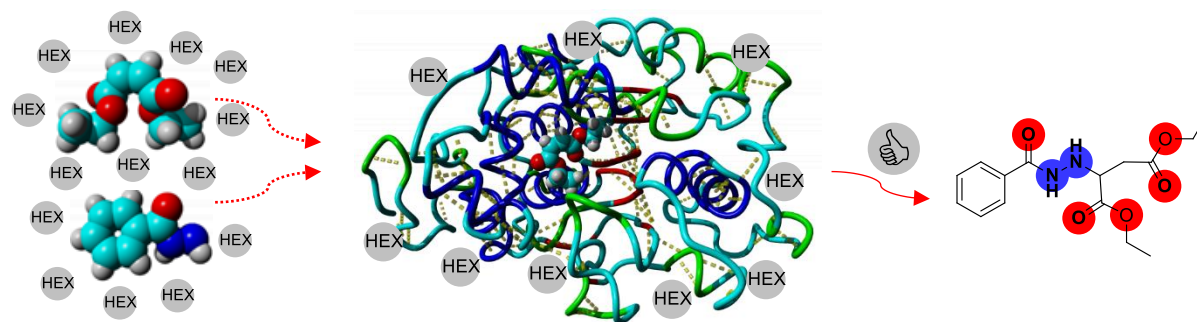
Solvent Dependent Activity of *Candida Antarctica* lipase B and its Correlation with a Regioselective Mono Aza-Michael Addition- Experimental and Molecular Dynamics Simulation Studies

Zohreh Nazarian^{a,b*}, Seyed Shahriar Arab^{a*}

^aFaculty of Biological Sciences, Tarbiat Modares University, Jalal Highway, Tehran 14115-154, Iran

^bDepartment of Chemistry and Petroleum Sciences, Shahid Beheshti University, Evin, Tehran 1983963113, Iran

Authors to whom correspondence should be addressed: sh.arab@modares.ac.ir, nazarianzohreh@yahoo.com.au (z_nazariyan@sbu.ac.ir)



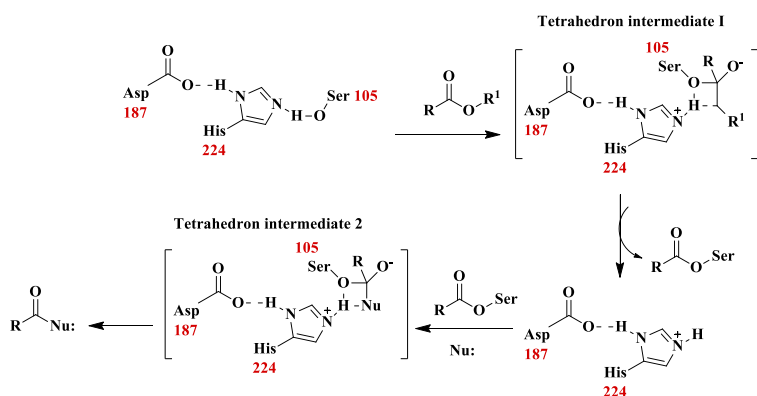
Abstract: With the aim of gaining understanding of the molecular basis of *Candida antarctica* lipase B (CALB) catalyzed regioselective mono aza-Michael addition of Benzhydrazide to Diethyl maleate (DEM) we decided to carry out molecular dynamics (MD) simulation studies in parallel with our experimental study. We found a correlation between the activity of CALB and the choice of solvent. Our study showed that solvent affects the performance of the enzyme due to the binding of solvent molecules to the enzyme active site region, and the solvation energy of substrates in the different solvents. We also found that CALB is only active in nonpolar solvent (*i.e.* Hexane), and therefore we investigated the influence of Hexane on the catalytic activity of CALB for the reaction. The results of this study and related experimental validation from our studies have been discussed here.

Keywords: aza-Michael addition, regioselective synthesis, CALB, biocatalysis, computational simulation, molecular dynamics simulation

Introduction

Lipases (triacylglycerol hydrolases E.C. 3.1. 1.3) are the most widely used versatile enzymes in organic synthesis, mainly because of its selectivity, commercial availability, broad specificity, and tolerance towards organic solvents, extreme temperature and pH.^[1] The concept of “promiscuity” has been perceived as a useful phenomenon, which can enhance the utility of lipase as a biocatalyst,^[2] leading to catalytic abilities in C-C bond formation, C-heteroatom (*e.g.* C-N, C-O, C-S coupling) bond formation, oxidative processes, and novel hydrolytic reactions.^[3] Among all the hydrolases used in promiscuously catalysis, *Candida antarctica* lipase B known as CALB is one of the most widely used biocatalysts in industry. CALB is a serine hydrolase belonging to the folding family of α/β hydrolases with the catalytic triad consists of Ser105, His224 and Asp187, and a resolved structure by X-ray diffraction.^[4,5] CALB immobilized on polyacrylic resin defined as Novozyme 435 is particularly useful enzyme, as it demonstrates exceptionally high stability and sound activity in organic solvents^[6] at high temperatures. Several reactions have been reported to use CALB, including the Michael addition,^[7,8] which is an unexpected reaction for hydrolytic enzymes. The Michael addition reaction (or 1, 4-addition) plays an important role in organic synthesis, as one of the most crucial methods to form C–C and C–X (X=heteroatom) bonds, owing to its atom-economy and efficiency.^[9] From the mechanism point of view Michael addition is a nucleophilic addition of a nucleophile to α , β -unsaturated carbonyl compound. When the nucleophile is nitrogen the reaction is called an aza-Michael addition. The aza-Michael additions products are β -amino carbonyl compounds, which are of significance both biologically and synthetically.^[10] So far, most related research within this

area have focused on aza-Michael addition incorporating amines as the Michael donor.^[11–13] However no research has been devoted to chemo-enzymatic aza-Michael addition involving amide as the Michael donor. Hydrazides are important amide intermediates^[14] specifically due to their contribution to the synthesis of compounds showing biological properties, including antituberculous agent (Isoniazid), HIV inhibitors, inhibitors of myeloperoxidase, glycogen phosphorylase, and pesticides.^[15] Acyl hydrazides have been known as novel inhibitors of mammalian cathepsin B and cathepsin H.^[16] Up until now, literatures provide mechanistic explanations of lipase catalyzed transesterification or amidation that proceeds through acylation of Ser105 (within the active site of CALB) by appropriate carbonyl compound, giving the acylated intermediate, followed by a nucleophilic attack (an alcohol, amine or water) at the Serine bound carbonyl group, and consequent liberation of the acyl group, giving the desired product (**Scheme 1**). Both steps proceed via an intermediate structure that has a Serine bound carbon, having a tetrahedral geometry. In seminal papers by Warshel^[17,18] and co-workers on the mechanism of protease and lipase, it was proposed enzyme catalyzed the reaction through stabilizing the negative charge of oxyanion hole (consists of Thr40 and Gln 106) and by the electrostatic interaction between Asp187 and the ionized His224.^[18]

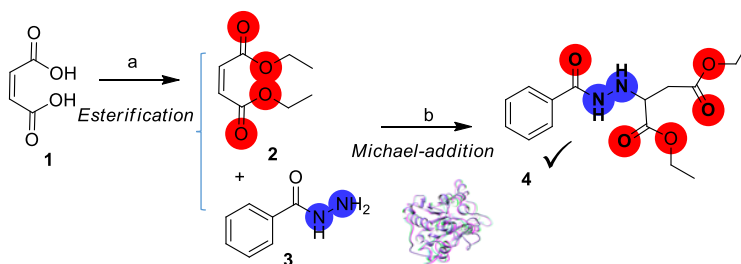


Scheme 1. The mechanism for the acylation/deacylation reaction catalyzed by CALB

In this research, chemo-enzymatic synthesis of diethyl 2-(2-benzoylhydrazinyl)succinate promoted us to undertake further studies, including computational simulation studies to gain insight into the solvent effect .

Results and Discussion

To investigate role of CALB as a biocatalyst in the reaction, initially noncommercial Benzhydrazide was treated with an excess amount of *Cis*-diethyl maleat (DEM) in a suspension of immobilized CALB (Novozyme 435) in hexanes (HEX), under argon at 65-68 °C. The reaction was allowed to stir sluggishly for 48 h, during which the reaction progress was monitored by TLC (**Scheme 2**).



Scheme 2. Reagents and conditions: a) EtOH (10.0 equiv, reflux, O/N), b) Novozyme 435 (60 mg), Hexanes (dry), rpm 170, 65- 68 °C, Argon atm, 2 Days

Since a portion of starting substrates remained after 24 h, therefore, the reaction time was extended to 48 h. After 48 h reaction showed a trivial progress, but starting substrates still were not consumed entirely. Once the reaction was deemed complete, the ^1H NMR spectrum (**Figure 1**) of the purified compound displayed a diagnostic peak at 5.70 ppm, attributing to CH in vicinity of NH, which was a hint for structure elucidation of the product.

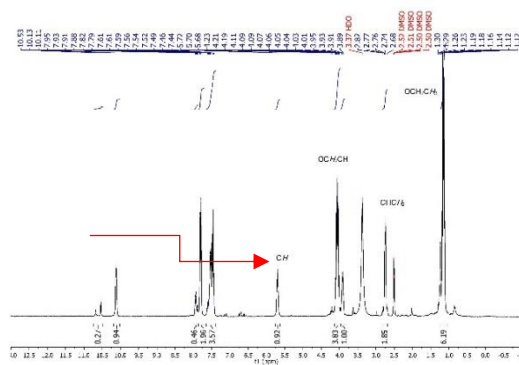


Figure 1. The ^1H NMR (mixture of two rotamers) spectrum showing diagnostic peak CHNH at 5.70 ppm

Surprisingly, the structure **4** (**Scheme 2**) turned out unknown, which was not reported previously. The compound was initially isolated as a transparent glassy solid, which under argon turned into an insoluble solid crystal. The isolated compound was unstable, and it was decomposed to the corresponding Benzhydrazide a few hours after isolation. According to this results reaction proceeded via regioselective mono adduct of Benzhydrazide and DEM, giving diethyl 2-(2-benzoylhydrazinyl) succinate as a major product. Considering the stereogenic center forged during the reaction optical rotation of the product was measured, which showed a racemic compound with no optical activity. The reaction condition is comparable to the research by Quirós *et al.*^[19] who treated DEM and/ or fumarate with amine in a suspension of CALB in Dioxane. However, under reported conditions no mono adduct product was isolated in either case. To further probe role of enzyme in the reaction, a control experiment was set up, which did not form the compound produced in the presence of enzyme. The possibility of undergoing an aza- Michael addition without using any catalyst or solvent was reported for amines previously.^[20,21] Noteworthy is failure of the reaction to generate the desired product with other solvents, including DCM and THF, which was in line with efficiency and activity of CALB in HEX as the nonpolar solvent compared with polar-solvents reported in literature.^[22] The choice of organic solvent is often crucial in the reaction outcome.^[23] Several factors have been proposed to affect enzyme activity in organic solvents, including enzyme flexibility, competitive inhibitory by solvent molecules, stabilization of the transition state, hydration level, and substrate solubility.^[22] Although these factors and parameters have provided some insight, still the complexity of the impact of organic solvent on the enzymatic reaction remains ambiguous. One reason is that only limited number of research attempted to conduct experiments and simulations in parallel to ensure that the reaction conditions are virtually comparable. For example the higher activity of CALB in nonpolar solvents has been proved for the esterification reaction in a parallel experimental theoretical studies.^[24] Specifically, CALB higher activity in HEX compared to polar solvents has been investigated in an esterification reaction,^[22] however its activity and stability at high temperature in a Michael addition has not been explored. To this end, we decided to combine experimental results with MD simulations to gain a molecular insight into organic solvent molecules effect on the catalytic activity of CALB in a Michael addition reaction.

Materials and Methods

The ^1H NMR spectra were recorded at 300 MHz, or at 500 MHz spectrometer and were reported in ppm relative to the solvent reference (CDCl_3 , δ 7.26 ppm; DMSO, δ 2.54 ppm). ^{13}C NMR spectrum was recorded at 75 MHz. Chemical shifts for ^{13}C NMR spectrum were reported in parts per million (ppm) downfield relative to the center line of the triplet

of DMSO at 39.5 ppm. Coupling constants (J) were quoted in Hertz (Hz). Unless otherwise noted, commercially available chemicals were used as received. Novozyme 435 was a generous gift from Novozymes (Denmark).

Experimental Results

Synthesis of *Cis*-diethyl maleat (DEM) (**2**): To a flask containing EtOH (10 equiv, 86 mmol, 4.0 mL) maleic acid (1.0 equiv, 8.6 mmol, 1.0 g) was added, followed by catalytic amount of H_2SO_4 . Reaction was allowed to reflux for 13-15 hrs. After reaction was deemed complete via TLC, reaction mixture was concentrated under reduced pressure, diluted with EtOAc, and dried over Na_2SO_4 . Concentration of the solvent in *vacuo* yielded a crude which was purified via flash chromatography (25% EtOAc: Hexanes) to produce the *title* compound as a colorless oil in 80% (1.1 g) yield. The data were in agreement with literature values.^[25] 1H NMR (500 MHz, $CDCl_3$, δ ppm) 6.25 (s, 2 H), 4.24 (q, J_{cis} = 7.1 Hz, 4 H), 1.3 (t, J_{cis} = 7.1 Hz, 6 H)

Synthesis of Diethyl 2-(2-benzoylhydrazinyl)succinate (**4**): To a vessel pre-charged with Argon Benzhydrazide (0.5 mmol, 0.68 g), DEM (0.75 mmol, 0.13 g), and Novozyme 435 (60 mg) were added. To this, dry hexane was added, and the reaction was allowed to stir at 65- 68 °C for two days. Upon completion of the reaction, the enzyme was filtered through a short plug of Celite, washed with EtOAc and the solvent was evaporated under reduced pressure. Then, crude was purified using neutralized column chromatography eluting with a solvent gradient of 20 % EtOAc/Hexanes to 40 % EtOAc/ Hexanes. The *title* compound was obtained as a white glassy solid in 60 % (92 mg) yield. 1H NMR (300 MHz, DMSO, δ ppm) (br d, 10.12, 1 H), (2 br d, 7.94 & 7.80, 5 H), 5.69 (t, 1 H), 4.06 (m, 4 H), 3.91 (m, 1 H), 2.5 (br s, 1 H), 1.24 (br d, 1 H), 1.15 (m, 6 H), ^{13}C NMR (75 MHz, DMSO, δ ppm) 171.4, 170.8, 166.3, 133.2, 132.0, 128.8, 127.9, 127.6, 61.1, 60.6, 59.2, 14.4, 14.3. ESI-MS Calcd. for $[C_{15}H_{20}N_2O_5]$ m/z = 308.14 found 308.3

Computational Method

The computational part of this study involves following sections: a) optimization and structural analyses of the docked poses via molecular docking, and b) MD simulation studies to verify the reliability of the results. The AutoDock 4.2^[26] software was used for molecular docking studies, and MGLTOOLS^[27] software for visualization and preparation of input files and output files analysis. The Chimera^[28] and LigPlus^[29] were used for 3D and 2D structure visualization, respectively. Details of system set up for the docking studies are explained in Supplementary Information associated with this manuscript.

Molecular Dynamics Simulation Studies

In order to understand the effect of HEX on the CALB activity in mediating the reaction, we performed a series of molecular dynamics simulations of CALB in HEX and WAT. The latter simulation was used as a reference to highlight structural changes in CALB when is exposed to organic medium. In the current study, classical MD simulations were employed, without considering the transition state, as this would require quantum mechanical calculations. Our analyses were directed towards the geometrical properties of CALB, and interactions of substrate/ solvent/water molecules within the CALB active site region. In the following sections, we initially analyze the position of each ligand within the active site over the course of simulations, and overall structure and flexibility of CALB in HEX and WAT. Then, interaction between CALB and solvent molecules, and interaction between substrate and solvent molecules are investigated. Lastly, based on distance analyses data together with level of stability of each ligand throughout simulation we study whether it makes a difference if ligand binding order to the active site is swapped?

Structure

The CALB crystal structure was taken from the Protein Data Bank (PDB ID: 1TCA; resolution 1.55Å). The enzyme is a monomer and consists of 317 amino acids. The active site consists of a catalytic triad formed by Ser105, His224, and Asp187.

Simulations

The MD simulations were performed by GROMACS 2021.2 software^[30] using ff99SB+ILDN force field.^[31] The systems were neutralized by adding the corresponding number of counterions (Na^+ and Cl^-). The protein was placed within TIP3P^[32] water molecules and Hexanes sphere of 10 Å diameter cubic box using gmx solvate software. Then the systems were energy minimized for 5000 cycles, using the steepest descent^[33] algorithm together with the conjugate gradient method to remove any bad interaction between atoms. The minimized systems were first gradually heated from 0 to 338 K over 200 ps using a Langevin thermostat at constant condition (NVT), and then systems were equilibrated for 200 ps at constant pressure (1.0 atm) (NPT). The MD simulations were done employing periodic boundary conditions adopting a 10 Å cut-off for non-bonded interactions, while non-bonded electrostatics interactions were carried out at a distance of 10 Å, adhering to the particle mesh Ewald (PME) method.^[34] The SHAKE algorithm^[35] was used to fix all covalent bonds involving hydrogen atoms. Finally, simulations were conducted at 338 K over 100 ns, during which data were saved after every 8.0 ps. The *Antechamber*^[36] software available in *AmberTools*^[37] package was utilized for parametrization of ligands, including DEM, Benzhydrazide and SEH. Generated parameters were then converted to GROMACS format via ACPYPE^[38] script.

Distance Analyses of Ligands with Respect to Ser105 and His224

The minimum distance changes and/ or variation between center of mass of ligands (*i.e.* DEM and Benzhydrazide) and center of mass Ser105 and His224 over the course of simulations (*i.e.* 100 ns) are shown in **Figure 2A** and **2B**, respectively. As can be seen in **Figure 2A**, DEM in HEX (red) demonstrated stability within the active site pocket, maintaining a distance between 0.2 to 0.4 nm to Ser105 throughout simulation. Comparably, DEM minimum distance to Ser105 in WAT (black) fluctuated, reaching 8.0 nm after 12 ns. The 8.0 nm distance between DEM and Ser105 remained unchanged until 22 ns, then the distance between these two conspicuously increased, reaching 1.6 nm at 24 ns. The distance fluctuation continued until DEM showed a distance around 4.5 nm to Ser105 at 96 ns, indicating complete detachment of the ligand from the enzyme. In contrast, minimum distance changes of Benzhydrazide with respect to Ser105 in aqueous medium, *i.e.* WAT (blue) demonstrates immediate exist of ligand from the active site pocket, which after 14 ns stayed at distance of 2.0 nm, implying that ligand completely detached from the enzyme. Probing Benzhydrazide status when simulation was conducted in HEX (green) again shows that ligand was pushed out of the active site pocket, staying at a distance of 1.4 nm with respect to Ser105 only after 0.1 ns, which indicated the ligand high instability. In an analogous fashion the minimum distance changes and/ or variation between center of mass of ligands and His224 was investigated, which showed similar behavior to those of observed for the minimum distance changes between ligands and Ser105. Given the position of the His224 at the surface of the active site, and close distance between Benzhydrazide and His224 in HEX (**Figure 2B**) compared with distance between Benzhydrazide

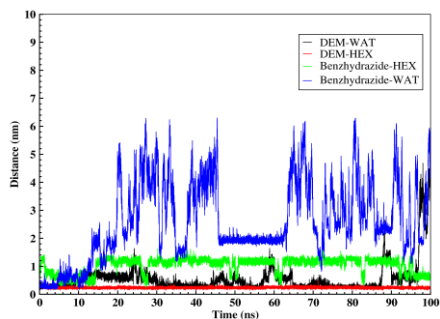


Figure 2 (A): The minimum distance changes between the ligands and Ser105

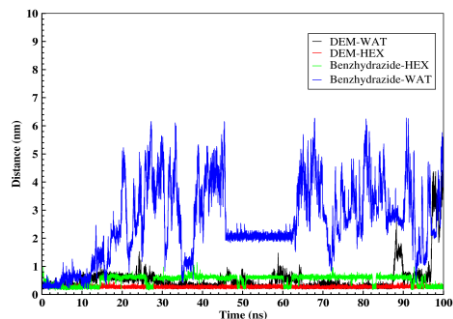


Figure 2 (B): The minimum distance changes between the ligands and His224

and Ser105, we could conclude that Benzhydrazide was pushed out of the active site in HEX, positioning at a distance close to His224. According to the **Figure 2A** and **2B** the Benzhydrazide detached from the CALB after 12 ns in WAT, while DEM is still in the active site after 51 ns. The status of Benzhydrazide in HEX is shown in **Figure 3**, where Benzhydrazide was engaged with hydrogen bonding with Ser105 and His224 in HEX before the simulation, yet it was not stable and has left the active site. Since Benzhydrazide completely detached from the CALB when it was in WAT, therefore no figure has been brought for this status. The status of DEM in WAT can be seen in **Figure 4A**, in which Ser105 was engaged with hydrogen bonding with DEM before the simulation, but has left the active site after 51 ns. To further elaborate on this, **Figure 4B** displays the situation of DEM with respect to Ser105 at 7.5 and 51 ns.

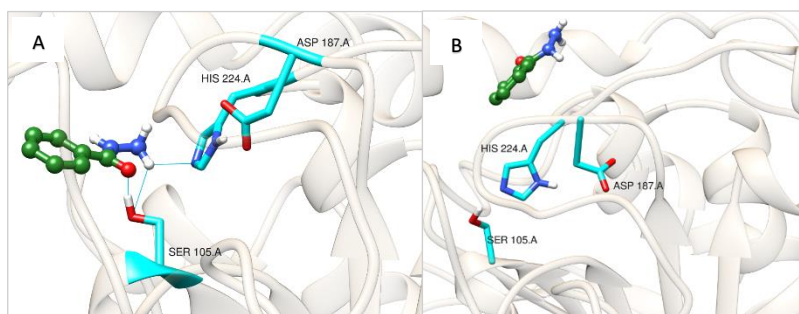


Figure 3. The status of Benzhydrazide in the active site pocket before (A) and after (B) the simulation in HEX

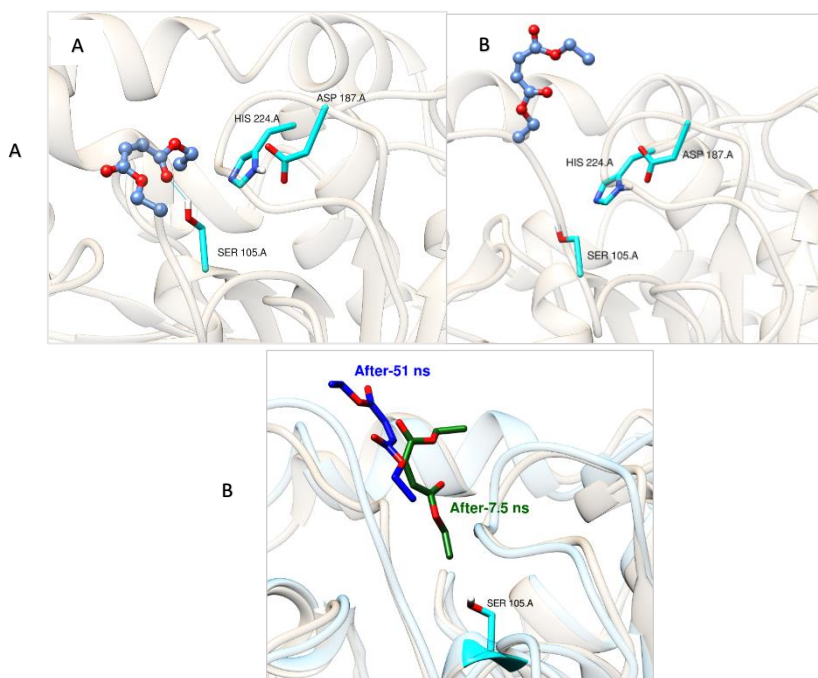


Figure 4. (A) The status of DEM in the active site before (A) and after (B) simulation in WAT; (B) The status of DEM in the active site at 7.5 and 51 ns in WAT

The position of DEM in active site before and at 51 ns after simulation in HEX is shown in **Figure 5**. As can be seen DEM was stable within the active site pocket, and was engaged with Ser105 and Gln106 through hydrogen bonding with the hydrogen bond length at 3.25 Å between oxygen of DEM and Ser105 at the end of simulation. Probing the interactions between DEM and amino acid residues through hydrogen bonding in HEX demonstrated that around 7.9 % of simulation time DEM was engaged with hydrogen bonding with Thr40. This calculation for other amino acid residues showed around 4.8 % simulation time Ser105 and around 15.9 % Gln106 were engaged with DEM through

hydrogen bonding.

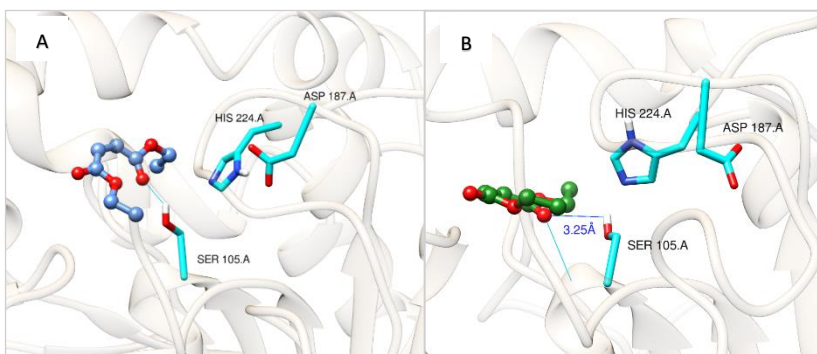


Figure 5. Position of DEM in the active site pocket before (A) and after (B) simulation in HEX

The HEX molecule position with respect to the CALB-DEM complex at the entrance of active site pocket is shown in **Figure 6A**, along with DEM position relative to Thr40, Ser105 and Gln106 in **Figure 6B**. Hydrogen bonding plays an important role in ligand stabilization within the active site pocket; the more hydrogen bonding interactions the more stable is the ligand in the active site. The hydrogen bonding interactions between DEM and amino acid residues of active site in HEX is shown in **Figure 6C**. According to this figure in the most majority of times one hydrogen bonding interaction was formed between DEM and CALB, which at times it reached three interactions. This figure also shows 0.317 on average interactions between DEM and the CALB.

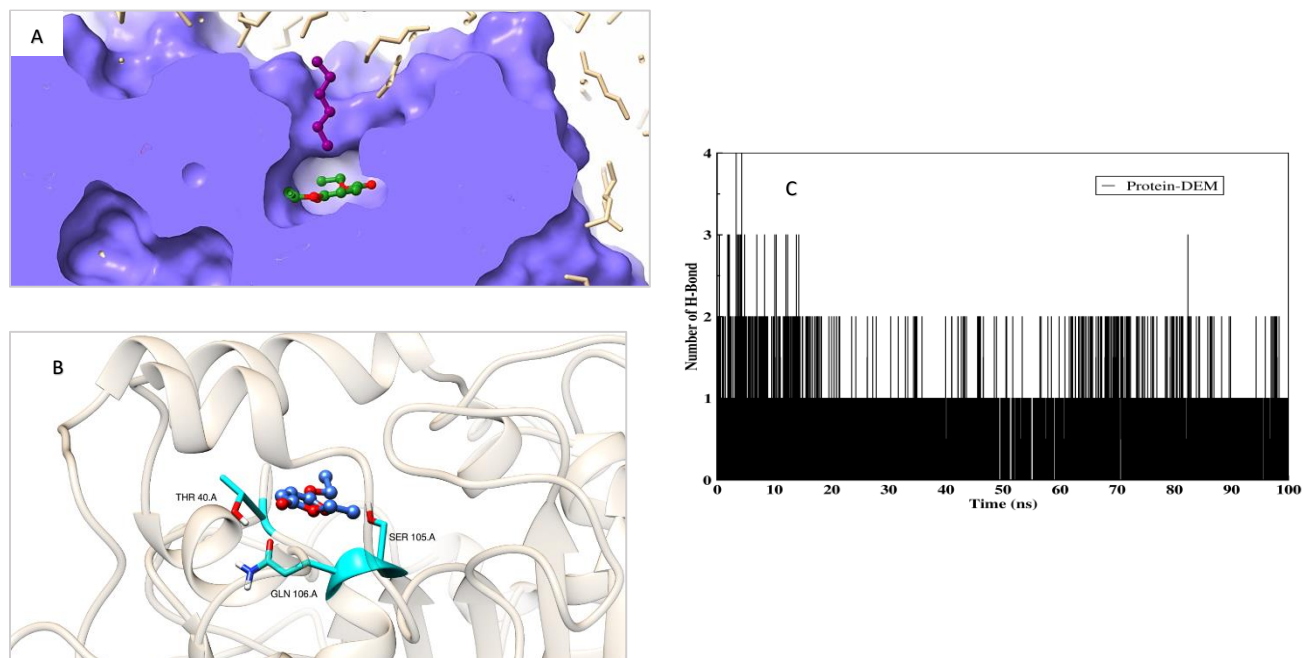
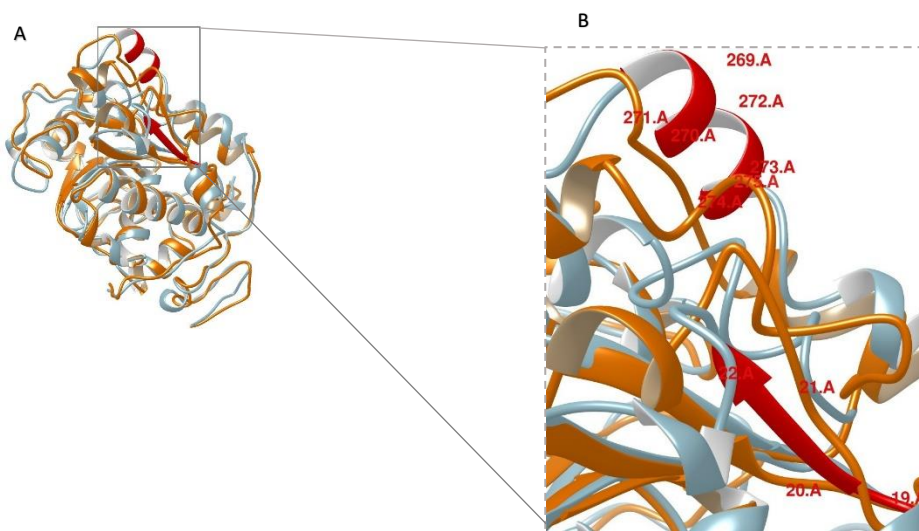


Figure 6. (A) DEM status in the active site pocket along with HEX at the entrance of the pocket; (B) DEM position relative to Thr40, Ser105 and Gln106 at the end of simulation in HEX; (C) Overall hydrogen bonding interaction changes between DEM and CALB in HEX throughout simulation.

Structure and Flexibility of CALB in Aqueous and Nonaqueous Medium

The superimposed backbone structures of CALB taken from the last frame of each simulation (*i.e.* 100 ns) in interaction with DEM both in HEX and WAT is shown in **Figure 7A**. Overall, it appears the backbone structure was maintained in the HEX and WAT throughout simulations, although it showed minor differences, which can be distinctively observed at regions 19-22 and 268-275 amino acid residues in WAT (light blue) (highlighted in red **Figure 7A** and zoomed in **Figure 7B**) that do not superimpose with corresponding regions in HEX (orange). Several geometrical properties were calculated to evaluate the stability and flexibility of CALB in HEX and WAT, including root-mean-square deviation (RMSD), radius of gyration (Rg), and solvent accessible surface area (SASA). The RMSD of the C α on the protein backbone with respect to the initial structure has been shown in **Figure 7C**. The higher RMSD value the more structural changes (*i.e.* the overall topology of the protein) with respect to the initial structure. Based on **Figure 7C** the CALB's RMSD the CALB reached stability in HEX (red) earlier than it did in WAT (black), but it underwent more structural changes relative to its initial structure in HEX, compared to WAT (which means it has a higher RMSD value in HEX). To investigate the flexibility of CALB in HEX and WAT one of the appropriate parameters to calculate is the root mean square fluctuation (RMSF). The RMSF of all amino acid residues of CALB in HEX and WAT in interaction with DEM throughout simulation has been shown in **Figure 7D**. The RMSF of the CALB in HEX in regions containing C α atoms for residues 20-30, 40-60 and 261-266 have shown considerable fluctuation compared to WAT. One of the main differences of RMSF in HEX and WAT was high fluctuation of the residues 138-150 (highlighted in red in **Figure 7E**) in WAT. The cited region (138-150) indeed is α 5 region located at the entrance of the pocket in CALB. This observation was closely similar to Skjøt *et al.*^[39] finding on the high flexibility of α 5 helix at region 138–152, and is in fact important region for CALB function as reported before.^[39] Additionally, amino acid residues from 168-187 and 184-207 corresponding to α 10 helix and loop L11 respectively, located in vicinity of α 5 region, did not show tangible fluctuation in **Figure 7D**. Overall, it can be deduced from **Figure 7D** that decrease in α 5 region fluctuation gave rise to stability of DEM in active site pocket.



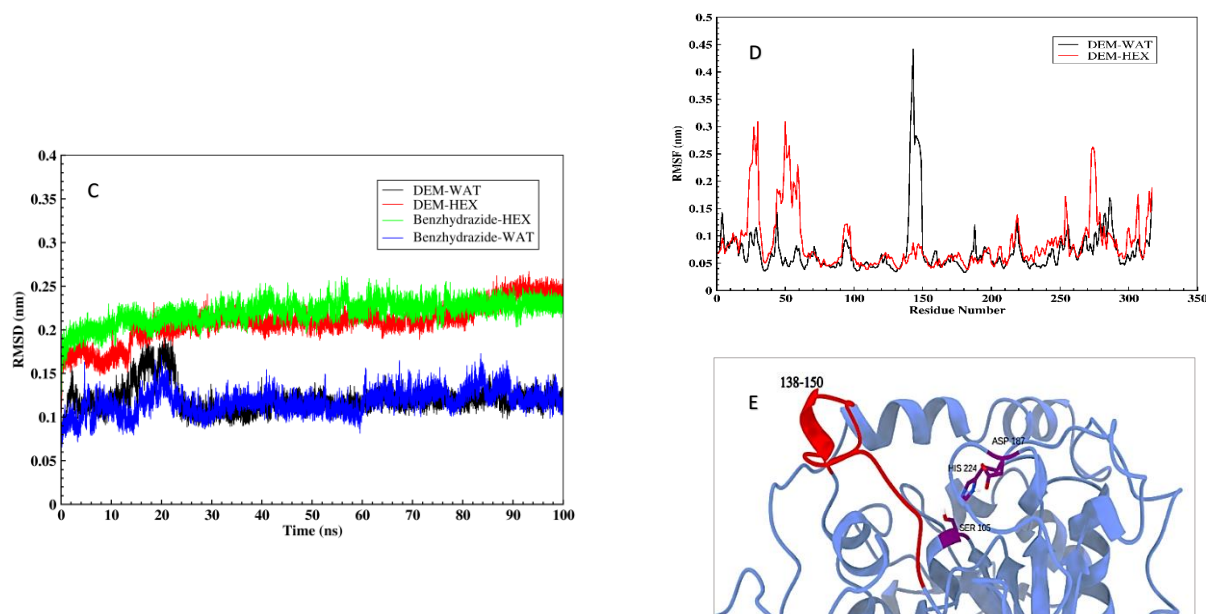
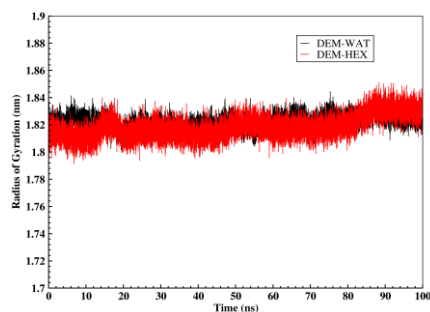


Figure 7. (A) Superimposition of the last frame of CALB in HEX (orange) with the last frame of CALB in WAT (light blue); (B) Regions of amino acid residues from 19 to 22 and 268 to 275 in WAT that do not superimpose with CALB in HEX is highlighted in red (C) The RMSD changes throughout simulation for all four systems; (D) The RMSF of all amino acid residues of CALB in interaction with DEM both in HEX and WAT (E) The high fluctuation region of the residues 138-150 (highlighted in red) in WAT

One of the indicators of protein structure compactness is the radius of gyration^[40] (R_g) that is affected by the medium in which the protein is placed. It represents the accessible surface of the protein as a result of interaction with solvent. On the basis of this definition, it appears according to **Figure 8** until 84 ns R_g in HEX had slightly lesser value compared to WAT, however after 84 ns it was increased. According to **Figure 8** R_g value in WAT throughout simulation was increased. In continuation solvent accessible surface area (SASA) of CALB both in HEX and WAT was explored. The SASA gradually increases with increasing solvent polarity as follows $\text{HEX} < \text{WAT}$ (**Figure 9A**). Water and polar solvent molecules interact with polar and charged residues of CALB, exposing them to the solvent. In comparison, in nonpolar solvents, exposed polar/charged residues are buried inside the enzyme to reduce the energy cost of exposing polar/charged residues to the hydrophobic media. As a result, the hydrophilic surface area decreases with increasing hydrophobicity of the solvent (**Figure 9B**), which was also reported by others.^[22,41]



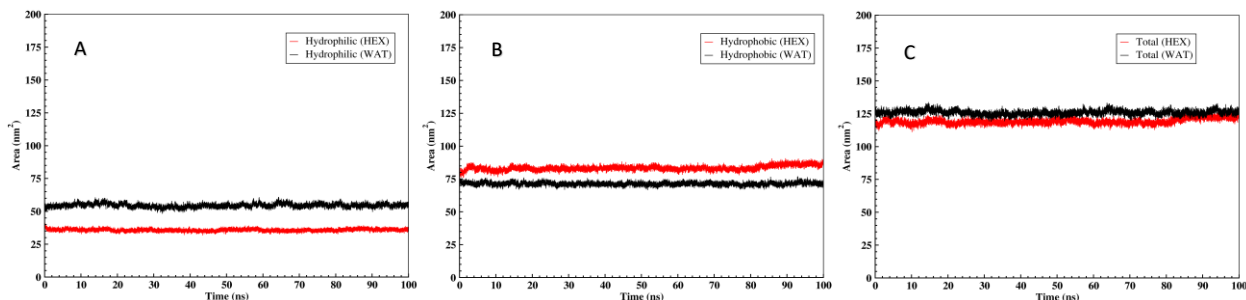


Figure 8. Radius of gyration (Rg) in WAT and HEX

Figure 9. (A) Hydrophilic solvent accessible surface area (SASA) of CALB in WAT and HEX; (B) Hydrophobic solvent accessible surface area (SASA) in HEX and WAT; (C) The total (Hydrophobic and Hydrophilic) solvent accessible surface area of CALB in WAT and HEX

Solvents Effect on the CALB Surface

The inhibitory effect of solvent molecules on the enzymatic reaction is obtained by the calculation of the binding free energy of substrate and solvent molecules to the active site pocket. Thus, ligand affinity to the protein was estimated by its binding free energy. We utilized *MMPBSA*^[42] (*g_mmpbsa*^[43] script for GROMACS) to estimate binding free energy of DEM to the active site pocket. To calculate free energy we opted for 400 frames from the stable regions of the simulation (*i.e.* 60-100 ns). As can be seen in **Table 1**. The Van der Waals energy (ΔG_{vdw}) was the main source for binding ligand to the active site pocket, showing -125.104 kJ compared with low electrostatic energy with only -9.759 kJ/mol.

Table 1. Binding free energies of DEM to the active site pocket

Unit: kJ/mol	ΔG_{Elec}	ΔG_{vdw}	ΔG_{polar}	ΔG_{apolar}	ΔG_{MMPBSA}
CALB-DEM	-9.795	-125.104	45.232	-12.058	-101.725

Also polar solvation energy (ΔG_{polar}) and non-polar solvation energy (ΔG_{apolar}) were calculated to be 45.232 kJ/mol and -12.058 kJ/mol, respectively. Accordingly, required energy for ligand solvation in HEX was significantly lower than it was required in WAT (positive ΔG indicates unfavorable binding). It is to say that the obtained binding free energy value (*i.e.* -101.725) for DEM via *MMPBSA*^[42] method indicates the relative free energy without considering entropy. Because Vander Waals energy plays a crucial role in binding the ligand to the protein, we therefore investigated the Vander Waals energy changes throughout simulation for Protein-DEM both in WAT and HEX (**Figure 10A**). Additionally electrostatic as well as Vander Waals energies have been calculated for interactions between CALB and HEX in the absence of ligand (**Figure 10B**).

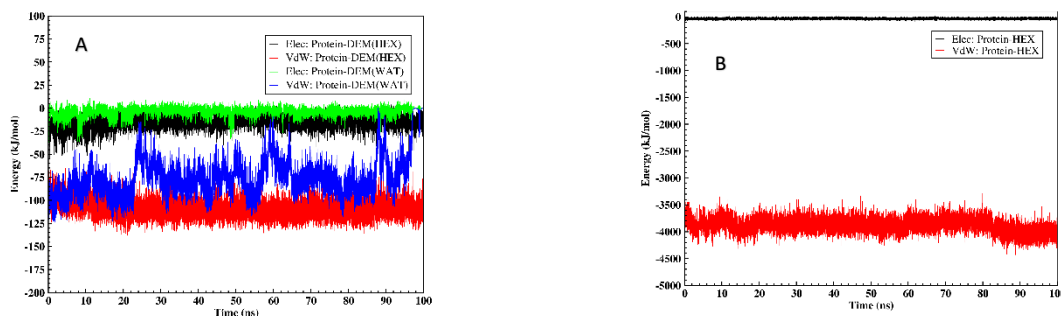


Figure 10. Vander Waals and electrostatic energy changes for (A): CALB -DEM and (B): CALB-HEX

As **Figure 10A** demonstrates the electrostatic energy between the CALB and DEM (CALB-DEM) in WAT (green) was close to zero, however it was around -10 KJ/mol in HEX. In comparison Vander Waals energy for CALB-DEM in WAT largely fluctuated (blue), indicating that DEM dissociates from the protein, whilst it was stable in HEX (red) after 15 ns. The electrostatic and Vander Waals energies for interaction between CALB and HEX throughout simulation is shown in **Figure 10B**. According to this graph, electrostatic energy was close to zero, while it was around -4000 KJ/mol for Vander Waals energy at the end of simulation.

Does It Make a Difference if “Ligand Binding Order” to the Active Site Is Swapped?

To probe whether DEM could be still stable if it enters and binds to the active site pocket as the second substrate (rather than the first substrate), Benzhydrazide was covalently bound to the Ser105, and the built complex (*i.e.* Ser105-Benzhydrazide) was labeled as SEH. The reason is that Benzhydrazide was not stable within the active site pocket through non-bonded interactions, and therefore it detached from the CALB soon at the beginning of the simulation (as it was shown in **Figure 2A** and **2B**). However, the made-up covalent bond between Benzhydrazide and CALB active site residue hindered Benzhydrazide dissociation from the enzyme. Then, DEM binding to the Ser105-benzhydrazide complex was analyzed by docking studies. Afterwards, the complex was used for MD simulation in HEX at 338 k. As it was shown previously DEM maintained a fixed distance towards Ser105, when it was bound to the active site as the first substrate (**Figure 11**, red), however it was not stable in the active site when it was bound as the second substrate, and it distanced from SEH, placing at a distance of 1.4 nm after 12 ns (orange), which indicated its exit from inside the active site.

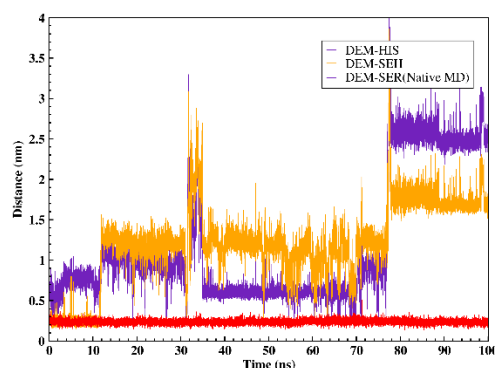


Figure 11: The minimum distance changes between center of mass DEM and His, as well as DEM and SEH

To prove exit of DEM from the active site we probed the distance between DEM and His224, since His224 is located at the surface of the active site compared to Ser105, which is located inside the active site pocket. Thus, according to the DEM-His distance analysis (purple) the ligand was pushed out of the active site, showing 1.2 nm distance from the His224 after 12 ns, and it was completely detached from the active site region after 52 ns. In order to envision this status **Figure 12** shows DEM close distance to the SEH at the beginning of simulation, and its further distance from the complex (Ser105-Benzhydrazide) over the course of simulation.

This observed instability for DEM-SEH rationalizes the priority of DEM binding to the active site over Benzhydrazide binding to the active site, as Benzhydrazide detached from the active site pocket at the beginning of the simulation both in WAT and HEX.

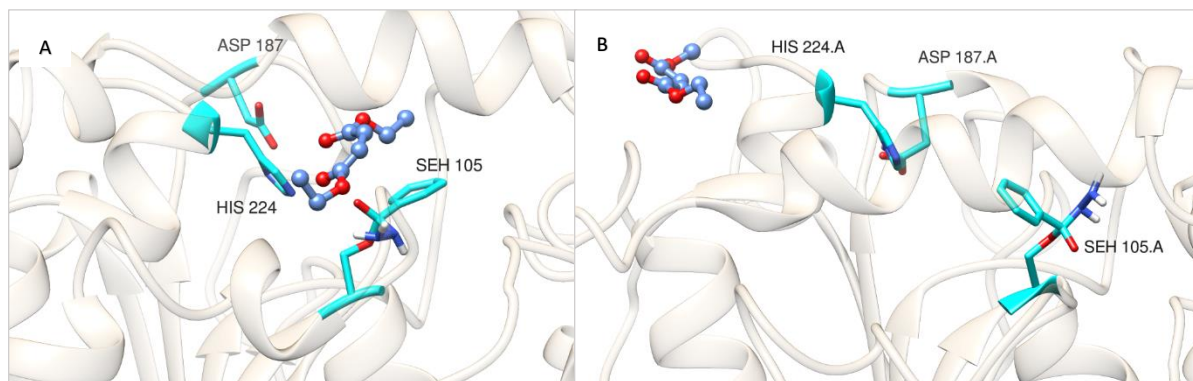


Figure 12. DEM position with respect to His224 and SEH (*i.e.* modified Ser105) before (A) simulation and after 50 ns (B)

Concluding remarks

To evaluate the influence of Hexane as the organic solvent on the activity of CALB at the molecular level we carried out a series of MD simulations of CALB in Hexane. We also included a simulation of CALB in water as a reference to be able to compare the structural changes in CALB when is exposed to organic medium. From geometrical analyses we deduced that the overall conformation of CALB was stable, and did not show major difference in WAT and HEX. Sound correlations exist between the binding free energies of the solvent molecules to CALB, and the solvation energy of the substrate molecules (*i.e.* DEM) with the catalytic activity of CALB. The formation of desired molecule in HEX despite instability of Benzhydrazide throughout simulation could be interpreted by the formation of DEM-CALB complex, as once DEM was engaged with the CALB, therefore the Benzhydrazide could easily interact with DEM-CALB complex to furnish desired target molecule.

Acknowledgement

Zohreh Nazarian acknowledges funding receipt from the bmn@international (Iran's elites foundation), and Ms Maryam Ghasemi for her undying support throughout this program.

Conflict of interest

The authors declare that there is no known conflict of interest.

Authors Contribution

Z.N: Conceived and proposed both experimental and computational studies, carried out and analysed the experimental part, performed and analysed most part of the computational studies, and wrote the manuscript
S. S. A: Supervised computational studies, revised the manuscript (computational work).

References

- [1] G. Angajala, P. Pavan, R. Subashini, *Biocatal. Agric* **2016**, 7, 257–270.
- [2] B. P. Dwivedee, S. Soni, M. Sharma, J. Bhaumik, J. K. Laha, U. C. Banerjee, *ChemistrySelect* **2018**, 3, 2441–2466.
- [3] E. Busto, V. Gotor-Fernández, V. Gotor, *Chem. Soc. Rev.* **2010**, 39, 4504–4523.
- [4] J. Uppenberg, M. T. Hansen, S. Patkar, T. A. Jones, *Structure* **1994**, 2, 293–308.
- [5] J. Uppenberg, N. Oehrner, M. Norin, K. Hult, G. J. Kleywegt, S. Patkar, V. Waagen, T. Anthonsen, T. A. Jones, *Biochemistry* **1995**, 34, 16838–16851.

- [6] C. Ortiz, M. L. Ferreira, O. Barbosa, J. C. S. dos Santos, R. C. Rodrigues, Á. Berenguer-Murcia, L. E. Briand, R. Fernandez-Lafuente, *Catal. Sci. Technol.* **2019**, *9*, 2380–2420.
- [7] R. O. M. A. de Souza, L. M. C. Matos, K. M. Gonçalves, I. C. R. Costa, I. Babics, S. G. F. Leite, E. G. Oestreicher, O. A. C. Antunes, *Tetrahedron Lett.* **2009**, *50*, 2017–2018.
- [8] Y. Yuan, L. Yang, S. Liu, J. Yang, H. Zhang, J. Yan, X. Hu, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2017**, *176*, 183–188.
- [9] A. Michael, *J. prakt. Chem.* **1887**, *35*, 349–356.
- [10] P. A. Magriotis, *Angew. Chem. Int. Ed.* **2001**, *40*, 4377–4379.
- [11] K. P. Dhake, P. J. Tambade, R. S. Singhal, B. M. Bhanage, *Tetrahedron Lett.* **2010**, *51*, 4455–4458.
- [12] S. Dutt, V. Goel, N. Garg, D. Choudhury, D. Mallick, V. Tyagi, *Adv. Synth. Catal.* **2020**, *362*, 858–866.
- [13] P. Steunenbergh, M. Sijm, H. Zuilhof, J. P. M. Sanders, E. L. Scott, M. C. R. Franssen, *J. Org. Chem.* **2013**, *78*, 3802–3813.
- [14] P. Majumdar, A. Pati, M. Patra, R. K. Behera, A. K. Behera, *Chem. Rev.* **2014**, *114*, 2942–2977.
- [15] A. J. Kettle, C. A. Gedye, M. B. Hampton, C. C. Winterbourn, *Biochem. J.* **1995**, *308*, 559–563.
- [16] N. Raghav, M. Singh, *Eur. J. Med. Chem.* **2014**, *77*, 231–242.
- [17] P. Schopf, A. Warshel, *Proteins*. **2014**, *82*, 1387–1399.
- [18] A. Warshel, G. Naray-Szabo, F. Sussman, J. K. Hwang, *Biochemistry* **1989**, *28*, 3629–3637.
- [19] M. Quirós, C. Astorga, F. Rebolledo, V. Gotor, *Tetrahedron* **1995**, *51*, 7715–7720.
- [20] B. C. Ranu, S. S. Dey, A. Hajra, *ARKIVOC* **2002**, *2002*, 76–81.
- [21] K. Kodolitsch, F. Gobec, C. Slugovc, *Eur. J. Org. Chem.* **2020**, *2020*, 2973–2978.
- [22] S. Dutta Banik, M. Nordblad, J. M. Woodley, G. H. Peters, *ACS Catal.* **2016**, *6*, 6350–6361.
- [23] A. J. RUSSELL, A. M. KLIBANOV, *Biochem. Soc. Trans.* **1989**, *17*, 1145.
- [24] A. Kumar, R. A. Gross, *Biomacromolecules* **2000**, *1*, 133–138.
- [25] N. Kaplaneris, A. Bisticha, G. N. Papadopoulos, D. Limnios, C. G. Kokotos, *Green Chem.* **2017**, *19*, 4451–4456.
- [26] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- [27] S. Forli, R. Huey, M. E. Pique, M. F. Sanner, D. S. Goodsell, A. J. Olson, *Nat Protoc* **2016**, *11*, 905–919.
- [28] T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, *Protein Sci.* **2018**, *27*, 14–25.
- [29] R. A. Laskowski, M. B. Swindells, *J. Chem. Inf. Model.* **2011**, *51*, 2778–2786.
- [30] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* **2005**, *26*, 1701–1718.
- [31] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, *Proteins* **2010**, *78*, 1950–1958.
- [32] P. Mark, L. Nilsson, *J. Phys. Chem. A* **2001**, *105*, 9954–9960.
- [33] J. Fliege, B. F. Svaiter, *Mathematical Methods of OR* **2000**, *51*, 479–494.
- [34] M. K. Gilson, K. A. Sharp, B. H. Honig, *J. Comput. Chem.* **1988**, *9*, 327–335.
- [35] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327–341.
- [36] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Am. Chem. Soc.* **2001**, *222*, U403.
- [37] T. J. Macke, W. A. Svrcek-Seiler, R. A. Brown, I. Kolossváry, Y. J. Bomble, D. A. Case, *AmberTools Users' Manual* **2010**.
- [38] A. W. Sousa da Silva, W. F. Vranken, *BMC Res. Notes* **2012**, *5*, 367.
- [39] M. Skjöt, L. De Maria, R. Chatterjee, A. Svendsen, S. A. Patkar, P. R. Østergaard, J. Brask, *ChemBioChem* **2009**, *10*, 520–527.
- [40] M. Y. Lobanov, N. S. Bogatyreva, O. V. Galzitskaya, *Mol Biol* **2008**, *42*, 623–628.
- [41] C. Li, T. Tan, H. Zhang, W. Feng, *J. Biol. Chem.* **2010**, *285*, 28434–28441.
- [42] N. Homeyer, H. Gohlke, *Mol. Inform* **2012**, *31*, 114–122.
- [43] R. Kumari, R. Kumar, A. Lynn, *J. Chem. Inf. Model.* **2014**, *54*, 1951–1962.