

In Silico characterization of acylamidase from *Actinoalloteichus hoggarensis*: A promising biocatalyst for the bioremediation of persistent organic pollutants

Sara Zergoun^{a,1}, Ikrame Abayahia^{a,1}, Mahfoud Bakli^{b,*}, Fatih Şaban Beriş^c, Ahmad Fadhlurrahman Ahmad Hidayat^d, Saharuddin Bin Mohamad^{d,e}, Muhammad Munir^{f,*}

^a Département de Biologie, Faculté des Sciences de la Nature et de la Vie et Sciences de la Terre, Université de Ghardaïa, B.P. 455, Ghardaïa 47000, Algeria

^b Laboratoire de Valorisation et Conservation des Écosystèmes Arides (LVCEA), Faculté des Sciences de la Nature et de la Vie et Sciences de la Terre, Université de Ghardaïa, B.P. 455, Ghardaïa 47000, Algeria

^c Recep Tayyip Erdogan University, Faculty of Arts & Sciences, Department of Biology, Rize, Türkiye

^d Institute of Biological Sciences, Faculty of Science, Universiti Malaya, Kuala Lumpur, Malaysia

^e Center of Research in Systems Biology, Structural Bioinformatics and Human Digital Imaging (CRYSTAL), Universiti Malaya, Kuala Lumpur, 50603, Malaysia

^f Date Palm Research Center of Excellence, King Faisal University, Al-Ahsa 31982, Saudi Arabia

ARTICLE INFO

Keywords:

Acylamidase
Actinoalloteichus hoggarensis
 Micropollutants
 Bioremediation
 Molecular docking
 Molecular dynamics simulations

ABSTRACT

Widespread contamination by persistent organic micropollutants, particularly polycyclic aromatic hydrocarbons (PAHs) from petrochemical activities, poses risks to ecosystems and human health, necessitating sustainable remediation strategies. Enzymatic bioremediation offers a promising alternative, with amidases emerging as versatile catalysts. This study investigates the catalytic potential of an acylamidase from *Actinoalloteichus hoggarensis*, a thermotolerant and halotolerant actinobacterium isolated from the Algerian Sahara Desert, whose extremophilic origin suggests inherent stability for harsh conditions. Using a comprehensive *in silico* approach, the enzyme's structural and functional properties were characterized, and its interactions with 34 micropollutants were evaluated via molecular docking. A high-quality three-dimensional model was generated and validated. Conserved domain analysis classified the enzyme within the amidase signature family. Molecular docking revealed strong binding affinities across diverse contaminants. Among PAHs, benzo[a]pyrene exhibited the strongest affinity, followed by chrysene and anthracene. The enzyme also showed notable interactions with the plastic additive triphenyl phosphate, highlighting its substrate versatility. Molecular dynamics simulations confirmed the stability of enzyme-ligand complexes (RMSD < 0.35 nm) and provided insights into key binding residues. These results indicate that this acylamidase possesses structural determinants accommodating both PAHs and other organic micropollutants, providing a computational foundation for its application in enzymatic bioremediation.

1. Introduction

Micropollutants have recently emerged as a significant class of environmental contaminants, typically detected at trace concentrations ranging from ng/L to µg/L [1]. Largely derived from industrial activities, they include pharmaceuticals, petroleum-derived hydrocarbons, and industrial chemicals, all of which have been widely detected in aquatic

environments. For instance, pharmaceutical micropollutants have been reported in approximately 65% of river systems across 72 countries [2], and up to 72% of wastewater samples have been found to contain them [3].

The biodegradability of many micropollutants is limited, and their concentrations often persist within the range of 1–100 µg/L [1]. Such persistence can promote bioaccumulation within ecosystems, adversely

* Corresponding authors.

E-mail addresses: zergoun.sara@univ-ghardaia.dz (S. Zergoun), abayahia.ikrame@univ-ghardaia.dz (I. Abayahia), bakli.mahfoud@univ-ghardaia.edu.dz (M. Bakli), fatih.beris@erdogan.edu.tr (F.Ş. Beriş), fadhlurrahmanhidayat@um.edu.my (A.F. Ahmad Hidayat), saharuddin@um.edu.my (S. Bin Mohamad), mmunir@kfu.edu.sa (M. Munir).

¹ These authors contributed equally

<https://doi.org/10.1016/j.dwt.2026.101834>

Received 8 April 2026; Received in revised form 12 June 2026; Accepted 14 June 2026

Available online 15 June 2026

1944-3986/© 2026 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

affecting their integrity. Regarding human health, drinking water contamination has been associated with neurological disorders, hormonal imbalances, and gastrointestinal infections [4]. These concerns underscore the urgent need of effective removal organic micropollutants from wastewater. Conventional treatment methods, including coagulation, filtration, adsorption, chemical precipitation, and advanced oxidation processes, are commonly employed. However, their efficiency is often limited by high operational costs, challenges associated with residual disposal, and potential formation of toxic byproducts that may be more hazardous than the original compounds. Consequently, biological approaches have attracted increasing attention as sustainable alternatives [5].

Microbial bioremediation has been increasingly recognized as a green and eco-friendly strategy for pollutant removal. *In situ* microbial populations exhibit remarkable metabolic diversity, allowing them to utilize a wide variety of petroleum-related compounds as sources of carbon and energy [6]. This metabolic versatility makes bioremediation a cost-effective and environmentally safe method for eliminating organic contaminants without generating harmful byproducts [5]. Among biological strategies, enzyme-driven processes are particularly notable for their high specificity and catalytic effectiveness, as they rely on enzymes produced by microbial communities, including bacteria and fungi, to degrade or transform environmental pollutants. Notably, amidases represent a key class of enzymes involved in the biotransformation of diverse organic compounds. Also known as acylamidases, they are widely distributed across bacteria, eukaryotes, and archaea. A significant proportion of microbial species can utilize amides as their sole sources of carbon and nitrogen for growth [7]. Acylamidases (EC 3.5.1.4) catalyze the hydrolysis of low-molecular-weight carboxylic acid amides to produce corresponding acids and ammonium ions. The activity of most well-characterized amidases is generally restricted to amides containing free amino groups. However, only a limited number of amidases have been reported to efficiently hydrolyze low-molecular-weight N-substituted amides [8]. Structurally, amidases are classified into two main families: the nitrilase superfamily and amidase signature family. Beyond their biological roles, they are increasingly recognized as versatile biocatalysts in the biosynthesis of valuable compounds and degradation of environmental contaminants, including micropollutants, with important applications in white and green chemistry [7].

Actinoalloteichus hoggarensis is a relatively rare actinobacterium isolated from the Algerian Saharan Desert [9]. Currently, limited information is available regarding its enzymatic profile, particularly concerning acylamidases. Nevertheless, this organism exhibits distinctive characteristics, such as thermotolerance, halotolerance, and the ability to survive under extreme environmental conditions, highlighting its potential for biotechnological applications [10].

Despite the growing interest in amidases, their interactions with diverse micropollutants, particularly polycyclic aromatic hydrocarbons (PAHs), remain insufficiently characterized, limiting their full exploitation in bioremediation strategies. Gaining a deeper understanding of these interactions through experimental approaches remains challenging because such studies are often time-consuming, labor-intensive, and costly [11]. Therefore, a computational approach was employed to gain preliminary insights into the structure and function of this enzyme.

In this study, *in silico* analyses were performed to characterize the structural and functional features of an acylamidase from *Actinoalloteichus hoggarensis*. In addition, molecular docking experiments were performed with 34 micropollutants to investigate enzyme-ligand interactions and evaluate the enzyme's potential application in bioremediation strategies targeting micropollutant removal.

2. Materials and methods

2.1. Protein sequence retrieval

To obtain the acylamidase sequence from *Actinoalloteichus hoggarensis*, a comprehensive search was performed using the UniProt Universal Protein Knowledgebase (<https://www.uniprot.org/>). The sequence of acylamidase with accession ID A0A221W1Z4 was retrieved in FASTA format and used as a query for further analysis.

2.2. Analysis of physicochemical properties

The physicochemical characteristics of the acylamidase sequence obtained from *Actinoalloteichus hoggarensis* were determined using the online tool ExPASy ProtParam (<https://web.expasy.org/protparam/protparam.html>) [12]. The tool provides relevant parameters for the acylamidase sequence, such as molecular weight, sequence length, number of amino acids, isoelectric point, amino acid composition, number of positively and negatively charged residues, elemental composition, chemical formula, total number of atoms, extinction coefficient, half-life, instability index, aliphatic index, and grand average of hydropathy (GRAVY) score. In addition, the hydrophilic and hydrophobic region structures of the protein were determined using the Kyte and Doolittle scales with the ProtScale tool (<https://web.expasy.org/protscale/>) on the ExPASy platform [13].

2.3. Prediction of signal peptides and determination of subcellular localization

The presence or absence of a signal peptide was predicted using SignalP 6.0 (<https://services.healthtech.dtu.dk/services/SignalP-6.0/>) [13]. Subcellular localization was predicted using the PSORTb version 3.0 server (<https://psort.org/psortb/index.html>) [14].

2.4. Secondary structure prediction

The secondary structure of the acylamidase sequence has been predicted by two reliable algorithms: SOPMA (Self Optimized Prediction Method with Alignment) (https://npsa.lyon.inserm.fr/cgi-bin/npsa_a_automat.pl?page=/NPSA/npsa_sopma.html) and PSIPRED 4.0 (PSI-BLAST based secondary structure PREDiction) (<https://bioinf.cs.ucl.ac.uk/psipred>). The first method, developed as a tool for protein structure prediction, provides different states of protein structure: α -helices, extended strands, and random coils [15]. In addition, PSIPRED, based on sequence profiles produced by PSI-BLAST, provides a more sensitive method of analysis, allowing for the determination of protein conformational states based on evolutionary data [16].

2.5. 3D protein structure modeling and evaluation

The three-dimensional structural conformation of the acylamidase target sequence was predicted using five different online tools, which include SWISS-MODEL [17], Phyre2.2 [18], AlphaFold [19], I-TASSER (Iterative Threading Assembly Refinement) [20], and trRosetta (transformed-restrained Rosetta) [21]. The parameters used in the prediction servers corresponded to the default settings.

Each of the initial three-dimensional models was then refined to improve structural accuracy and physicochemical properties by employing the GalaxyRefine server (<https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE>), which is a high-resolution server used for refining protein structures [22]. To evaluate the impact of refinement on the quality of the protein models, various tools available on the SAVES v6.1 server (<https://saves.mbi.ucla.edu/>) were employed. PROCHECK was used to assess both the initial and refined models by analyzing Ramachandran plots [23]. Furthermore, the generated models were validated using VERIFY3D, which evaluates the compatibility of

each amino acid residue with its three-dimensional structural environment, thereby ensuring structural reliability [24]. Additionally, ERRAT measures non-specific atomic interactions by comparing local structural statistics within sliding windows of nine residues with those obtained from highly refined experimental models. Structural images were generated using PyMOL (version 2.3).

2.6. Functional analysis

The conserved domains related to the structure and function of acylamidase were identified using the Conserved Domain Database (CDD) v3.21 provided by the National Center for Biological Information (NCBI) (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) [25]. Additionally, the acylamidase protein sequence from *A. hoggarensis* was analyzed using the MOTIF Finder online tool (<http://www.genome.jp/tools/motif/>) and the Pfam database to detect functional motifs and classify the protein within its respective family.

2.7. Molecular docking

Molecular docking was performed using the CB-Dock2 server (<https://cadd.labshare.cn/cb-dock2/>), which predicts protein–ligand interactions with an approximate success rate of 85% for binding pose prediction in benchmark tests [26]. This server first identifies the spatial centers and dimensions of potential binding pockets using a curvature-based cavity detection approach. Subsequently, docking was performed using the advanced molecular docking software AutoDock Vina [27]. This tool was used to investigate the interactions between the refined model of acylamidase from *A. hoggarensis* and the micropollutant ligands. The chemical structures of the ligands were retrieved primarily from the PubChem database (<https://pubchem.ncbi.nlm.nih.gov>) [28], and di(2-ethylhexyl) phthalate (DEHP) was retrieved from the ChemSpider database (<https://www.chemspider.com>). The ligand structures in the SDF format were subjected to energy minimization using Chem3D Ultra 8.0. Docking analysis identified five potential binding cavities and several enzyme–ligand interaction models. The complex with the lowest binding energy, as indicated by the minimal Vina score, was selected as the most favorable interaction. Ligand-binding interactions and amino acid residues were visualized using BIOVIA Discovery Studio Visualizer (v25.1.0).

Vina score values were compared among ligand–enzyme complexes across different micropollutant classes using a heatmap generated with RStudio (version 2024.12.1, Build 563).

2.8. Molecular dynamics simulations (MDS)

MDS were carried out using GROMACS (version 2025) to evaluate the stability and dynamic behavior of all proteins and complexes. Protein topology was generated using the CHARMM36 force field (2022 release) [29], while ligand parameters were prepared through the SwissParam website [30]. Each complex was placed in a cubic simulation box with a minimum distance of 1.0 nm from the box edges, solvated using the TIP3P water model, and neutralized with Na⁺ and Cl[−] ions. Energy minimization was performed using the steepest descent algorithm for up to 50,000 steps until the maximum force was below 1000 kJmol^{−1}·nm^{−1}. The systems were subsequently equilibrated under NVT conditions at 300 K for 100 ps using the V-rescale thermostat, followed by NPT equilibration at 1 bar for 100 ps using the Parriello–Rahman barostat [31]. Simulations employed the leap-frog integrator with a 2 fs time step, while hydrogen-containing bonds were constrained using the LINCS algorithm. Long-range electrostatic interactions were calculated using the Particle Mesh Ewald method, and periodic boundary conditions were applied in all directions [32]. Production simulations were conducted for 100 ns under NPT conditions at 300 K and 1 bar. The trajectories were recorded every 10 ps and analyzed for the total energy, root mean square deviation (RMSD), root

mean square fluctuation (RMSF), radius of gyration (Rg), and solvent-accessible surface area (SASA).

2.8.1. Molecular Mechanics/Generalized Born Surface Area (MM/GBSA)

Binding free energies were estimated using the MM/GBSA approach employed using gmx MMPBSA tools [33]. Snapshots from the MD trajectories were extracted at every 10-frame interval across the 100 ns simulation for MM/GBSA analysis. The binding free energy (ΔG_{bind}) was calculated from Eq. (1):

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S \quad (1)$$

where the enthalpy contribution (ΔH) is defined as in Eq. (2):

$$\Delta H = \Delta E_{\text{gas}} + \Delta G_{\text{solvation}} \quad (2)$$

The entropy contribution ($-T\Delta S$) was estimated using the interaction entropy method to account for dynamic fluctuations along the trajectory. All reported energies represent average values over the sampled trajectory frames, ensuring statistically reliable estimation of binding affinity and interaction profiles.

3. Results and discussion

3.1. Protein sequence retrieval

The amino acid sequence of the acylamidase of *Actinoalloteichus hoggarensis*, retrieved in FASTA format from UniProt, along with its primary features, is listed in Table 1.

3.2. Analysis of physicochemical properties

The physicochemical parameters of the acylamidase protein, as computed by ExPASy ProtParam (Table 2), showed that the acylamidase protein consisted of 477 amino acids. The molecular weight of the acylamidase protein was approximately 50,320.67 Da. This was calculated by summing the average isotopic mass values of all amino acids plus the mass of a water molecule [12]. Therefore, the function of a protein depends on its isoelectric point (pI), which is the pH at which the molecule has a net charge of zero. Proteins with a basic character have a pI of > 7, whereas acidic proteins have a pI of < 7 [34]. According to our results, the pI of the studied protein was 4.83, confirming its acidic nature. In this regard, negatively charged residues dominate over basic amino acids, which reduces the solubility of the protein as the pH approaches the isoelectric point and increases its solubility under alkaline conditions [35].

The reference method for determining the concentration of a protein is based on measuring its absorbance at 280 nm in an aqueous solution using its specific molar extinction coefficient at this wavelength (unlike cysteine, which does not absorb significantly beyond 260 nm) [12]. An extinction coefficient of 60,180 M^{−1}·cm^{−1} was calculated based on the assumption that all the cysteines were oxidized to cystines. Moreover, the stability of acylamidase was predicted based on the instability index, which was below 40 (33.44), indicating that the protein was stable, whereas values above 40 indicated protein instability [36]. The Grand Average of Hydropathicity (GRAVY) score characterizes the hydrophilic or hydrophobic profile of a protein. For acylamidase, a negative value indicates a hydrophilic nature, representing a strong affinity for water [37]. The aliphatic index (AI) indicates the proportion of amino acids with aliphatic side chains, such as alanine (A), leucine (L), isoleucine (I),

Table 1

Characteristics of the query sequences for the acylamidase enzyme from *Actinoalloteichus hoggarensis*.

Name	Accession number	Organism	Strain
Acylamidase	A0A221W1Z4	<i>Actinoalloteichus hoggarensis</i>	DSM 45943

Table 2

The physicochemical properties of the protein sequence were analyzed using the ProtParam tool of ExPASy.

Property	Value
Number of amino acids (AA)	477
Molecular weight (Da)	50320.67
Theoretical pI	4.83
Total number of positively charged residues (Arg + Lys)	43
Total number of negatively charged residues (Asp + Glu)	61
Extinction coefficients	60180
Estimated half-life	30 h
Instability index	33.44
Grand average of hydropathicity (GRAVY)	−0.093
Aliphatic index	78.66

and valine (V) [34]. The protein studied had an AI of 78.66, reflecting its high thermal stability [38]. This aligns with a previous study showing that the amidase from *Aspergillus fumigatus* exhibited significant thermal stability, with a $T_{1/2}$ (half-life) value of 13.37 h at 50°C [39].

3.3. Prediction of signal peptide and subcellular localization

The SignalP–6.0 server revealed that *A. hoggarensis* acylamidase lacks signal peptides at the N-terminus. Moreover, using the PSORTb subcellular localization server version 3.0, we found that acylamidase is a cytoplasmic protein, with a confidence score of 9.97.

3.4. Secondary structure analysis

Secondary structure predictions were further validated using the SOPMA server. The model predicted the following secondary structures: 40.25% alpha-helical regions, 7.97% extended strands, and 51.78% random coils. The acylamidase enzyme did not exhibit any signs of disordered protein-binding sites. This prediction was performed using PSIPRED to verify the results obtained from SOPMA, as shown in Fig. 1.

3.5. Homology modeling, refinement, and validation

No three-dimensional structure is available in the Protein Data Bank (PDB) for *A. hoggarensis* acylamidase. Accordingly, five structure prediction servers were used, including SWISS-MODEL, Phyre2.2, AlphaFold, I-TASSER and trRosetta. Each predicted model was then refined using the GalaxyRefine server, and the stereochemical quality of both the initial and refined models was validated using a Ramachandran plot with PROCHECK and VERIFY 3D. The Ramachandran plot statistics for each structural model are listed in Table 3.

Ramachandran plot analysis indicated a high number of residues in the most favored regions, whereas only a minimal number of residues were found in the disallowed regions for both models generated by the SWISS-MODEL and AlphaFold servers.

Table 3

Ramachandran plot assessments of acylamidase models produced using PROCHECK.

Model	Type of model	Favored region	Additional allowed region	Gnerously allowed region	Disallowed region
AlphaFold	Initial	91.0%	9.0%	0.0%	0.0%
	Refined	94.6%	5.1%	0.0%	0.3%
SWISS-MODEL	Initial	91.8%	8.2%	0.0%	0.0%
	Refined	94.6%	5.1%	0.0%	0.3%
Phyre2.2	Initial	84.6%	13.4%	1.8%	0.3%
	Refined	91.4%	8.1%	0.3%	0.3%
I-TASSER	Initial	77.0%	15.1%	5.6%	2.3%
	Refined	90.5%	6.4%	0.8%	2.3%
Tr-Rosetta	Initial	89.5%	10.2%	0.0%	0.3%
	Refined	94.1%	5.9%	0.0%	0.0%

Notably, the structure obtained using SwissModel was based on a template derived from AlphaFold, with 100% sequence identity and nearly complete coverage. This high degree of correspondence explains the consistency observed in Ramachandran plots.

Although the AlphaFold-generated model demonstrates a high level of accuracy, additional validation steps are necessary to confirm its reliability and suitability [40].

In this context, the model built using Swiss-Model was preferred due to its integration of advanced evaluation and structural refinement approaches, particularly aimed at improving the accuracy of local geometric features [41]. Furthermore, additional validation was performed using the Global Model Quality Estimate (GMQE) score, which reached 0.96, indicating a high-quality and reliable model. Therefore, the SWISS-MODEL-refined model was selected as the most suitable for further analysis.

A similar approach was applied in a study conducted by Chadha et al., where the structure of BAP31 was modeled using the SWISS-MODEL server, with a template derived from the AlphaFold database. This methodology illustrates the synergistic integration of predictive models generated using artificial intelligence and homology modeling techniques to improve the structural accuracy and increase the reliability of functional analysis [42].

In addition, the VERIFY3D prediction results suggest that the average 3D–1D scores indicate a high level of structural model quality. The ERRAT analysis score was 89.5075, indicating strong structural integrity and few conformational inaccuracies. Fig. 2.A shows the 3D structure of the acylamidase enzyme from *A. hoggarensis*, which was created using SWISS-MODEL and then refined with GalaxyRefine. The various conformations of the amino acid residues, as revealed by the Ramachandran plot, are represented in the three-dimensional protein structure depicted in Fig. 2.C. The results obtained from the validation tests using the VERIFY3D and ERRAT tools are shown in Fig. 2.B, D, respectively.

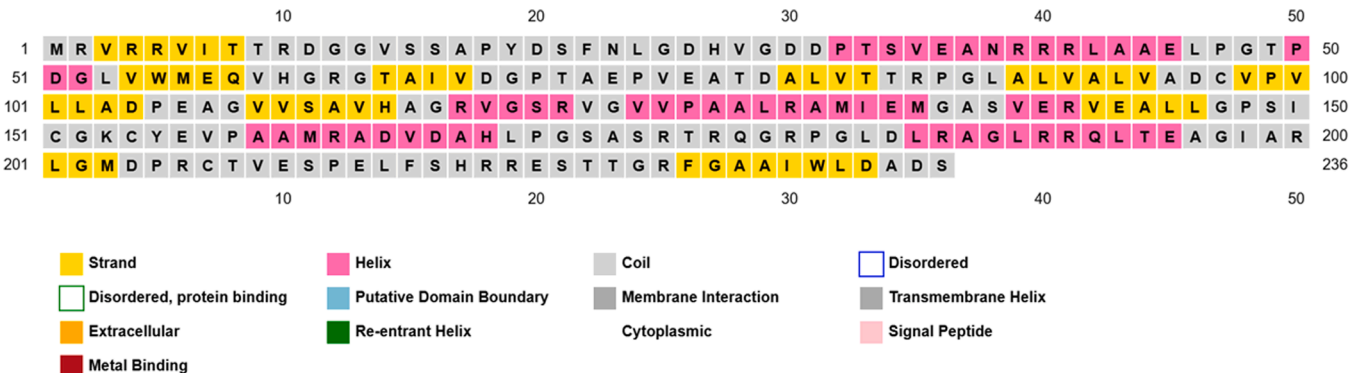


Fig. 1. Prediction of the secondary structure of *A. hoggarensis* acylamidase using the PSIPRED server.

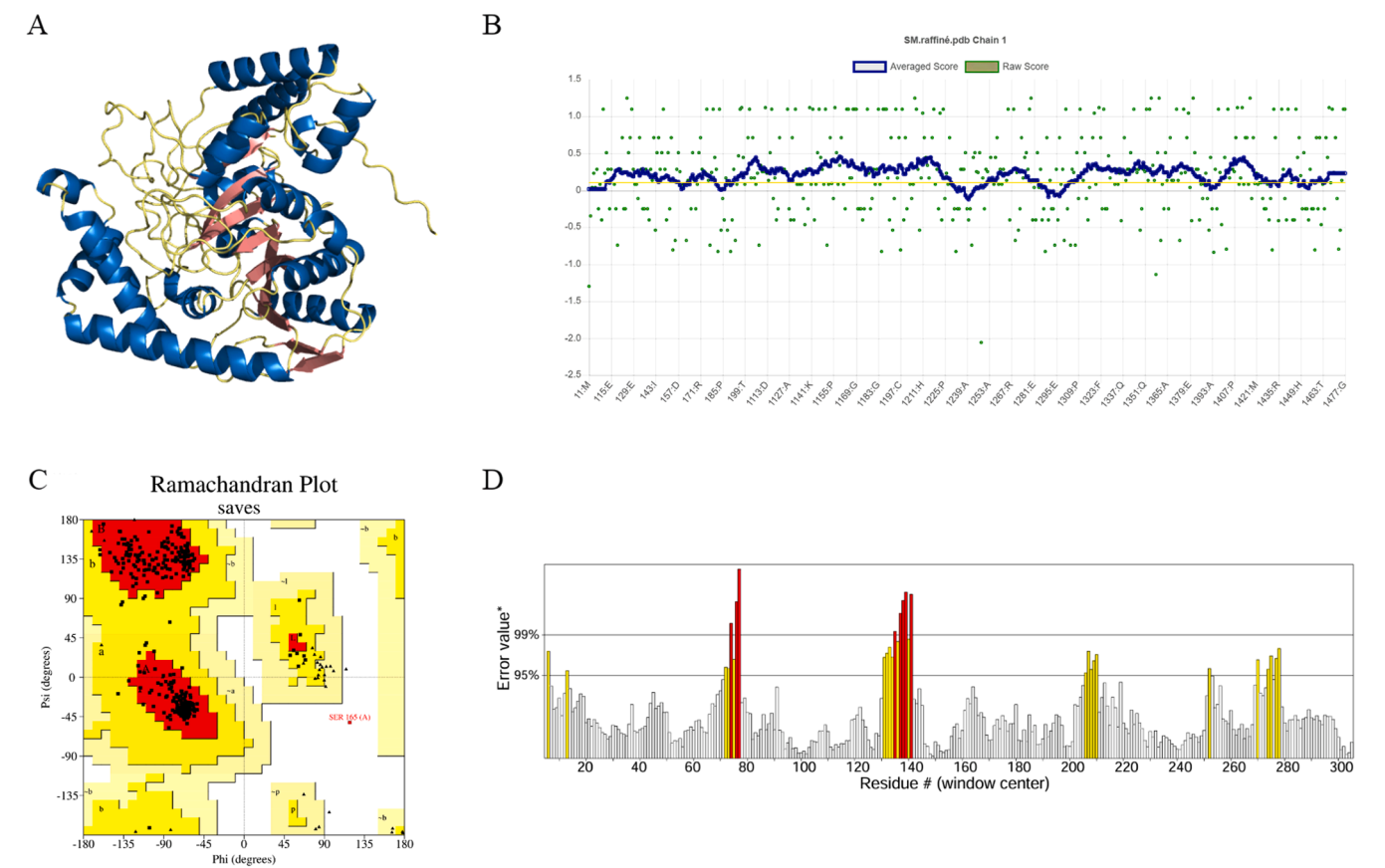


Fig. 2. Three-Dimensional Structural Model of Acylamidase Enzyme from *A. hoggarensis* and its Validation. a. The structural model of the acylamidase enzyme was developed using SWISS-MODEL, refined using GalaxyRefiner, and viewed using PyMOL software, illustrating the α -helix (sky blue), β -sheet (salmon), and loop regions (light yellow). b. Evaluation graph for the refined acylamidase structure using the VERIFY3D program. c. Ramachandran plot for the refined acylamidase structure generated using PROCHECK. The x-axis corresponds to the Phi (Φ) torsion angles, whereas the y-axis corresponds to the Psi (Ψ) torsion angles. The regions are defined as favored (A, B, L), additionally allowed (a, b, l, p), and generously allowed (~a, ~b, ~l, ~p); disallowed regions are depicted as red, yellow, beige, and white colors, respectively. Non-glycine or non-proline residues are depicted as black filled squares, while non-terminal glycine residues are depicted as black filled triangles. d. ERRAT quality plot of the refined acylamidase structure.

3.6. Functional analysis

Protein domains are evolutionarily conserved units that form three-dimensional structural modules and perform specific biological functions [43]. Analysis using the Conserved Domain Database (CDD), shown in Fig. 3, allowed the protein under study to be classified into the

amidase family. This family of enzymes catalyzes the hydrolysis of amide bonds in carboxylic acids, producing the corresponding acid and ammonium [8].

Using Pfam annotation performed with MotifFinder, two conserved functional motifs were identified in the acylamidase (Fig. 4). The amidase domain, corresponding to accession PF01425 (positions

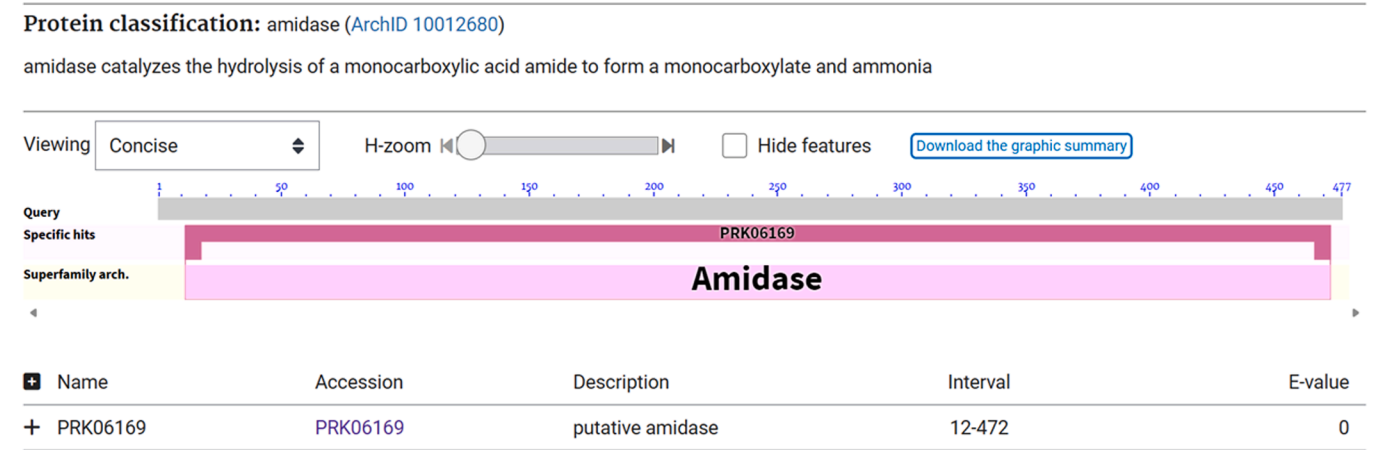
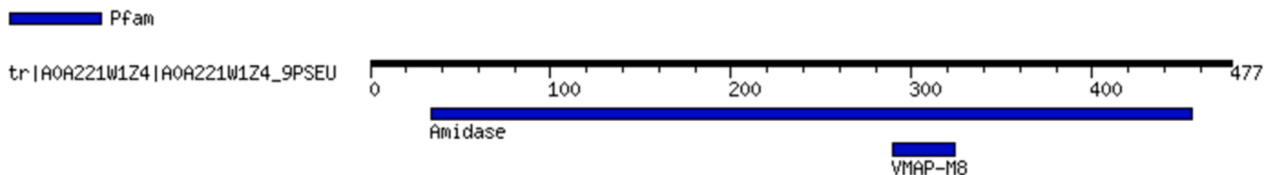


Fig. 3. Identification of acylamidase protein domains using the CDD. Specific hit corresponding to the amidase domain (PRK06169) and a superfamily belonging to the amidase family was identified. The domain is located between positions 12 and 472, with an E-value of 0 indicating high significance.

Result of MotifFinder

Number of found motifs: 2 



Pfam (2 motifs)

Pfam	Position(Independent E-value)		Description
Amidase	34..455(1.5e-103)	Detail	PF01425, Amidase
VMAP-M8	290..324(0.57)	Detail	PF19969, vWA-MoxR associated protein middle region (VMAP-M) 8

Fig. 4. Analysis of functional motifs in the sequence of acylamidase from *Actinoalloteichus hoggarensis* using MOTIF Finder. The two blue lines represent conserved domains identified using the Pfam database, along with their amino acid ranges within the protein sequence.

34–455), exhibited an extremely significant E-value (1.5×10^{-103}). This result is consistent with the annotations obtained from the CDD database. In addition, a VMAP-M8 domain (PF19969) was detected in the central region of the protein, which is associated with the vWA-MoxR system.

3.7. Docking analysis

Molecular docking analysis was performed to assess the interaction of various ligands with the acylamidase enzyme of *A. hoggarensis*. Each ligand, along with the refined enzyme, was subjected to docking analysis

Table 4

Physicochemical characteristics of the chosen micropollutant compounds for molecular docking study.

Ligand	Pubchem CID	Molecular formula	Molecular weight (g/mol)	Classification of Environmental Micropollutants
Diflubenzuron	37123	C ₁₄ H ₉ ClF ₂ N ₂ O ₂	310.68	Pesticides
Boscalid	213013	C ₁₈ H ₁₂ Cl ₂ N ₂ O	343.2	
Fluopyram	11158353	C ₁₆ H ₁₁ ClF ₆ N ₂ O	396.71	
2,6-dichlorobenzamide	16183	C ₇ H ₅ Cl ₂ NO	190.02	Pharmaceutical Micropollutants
Cyflufenamid	135515530	C ₂₀ H ₁₇ F ₅ N ₂ O ₂	412.4	
Carbendazim	25429	C ₉ H ₉ N ₃ O ₂	191.19	
Sulfanilamide	5333	C ₆ H ₈ N ₂ O ₂ S	172.21	
Sulfamethoxazole	5329	C ₁₀ H ₁₁ N ₃ O ₃ S	253.28	
Sulfadiazine	5215	C ₁₀ H ₁₀ N ₄ O ₂ S	250.28	
Sulfamethazine	5327	C ₁₂ H ₁₄ N ₄ O ₂ S	278.33	
N [4]-acetylsulfamethoxazole	65280	C ₁₂ H ₁₃ N ₃ O ₄ S	295.32	Industrial Micropollutants
2-Methylantraquinone	6773	C ₁₅ H ₁₀ O ₂	222.24	
2-Methylbenzenesulfonamide	6924	C ₇ H ₉ NO ₂ S	171.22	
Polyethylene glycol mono(octylphenyl) ether	94543	C ₁₆ H ₂₆ O ₂	250.38	
Galaxolide	91497	C ₁₈ H ₂₆ O	258.4	Plastic Additives and Plasticizers
Bisphenol AF	73864	C ₁₅ H ₁₀ F ₆ O ₂	336.23	
Bisphenol S	6626	C ₁₂ H ₁₀ O ₄ S	250.27	
Triphenyl phosphate	8289	C ₁₈ H ₁₅ O ₄ P	326.3	
4,4'-(4-Methylpentane-2,2-diyl)diphenol	81259	C ₁₈ H ₂₂ O ₂	270.4	Nitrogen-Containing Aromatic Compounds
Di(2-ethylhexyl) phthalate (DEHP)	8343	C ₂₄ H ₃₈ O ₄	390.6	
N-ethyl-o-toluenesulfonamide	14110	C ₉ H ₁₃ NO ₂ S	199.27	
N-butylbenzenesulfonamide	19241	C ₁₀ H ₁₅ NO ₂ S	213.30	
N-cyclohexyl-4-methylbenzenesulfonamide	6633	C ₁₃ H ₁₉ NO ₂ S	253.36	
p-Phenylene terephthalamide (PPTA)	21761075	C ₁₄ H ₁₀ N ₂ O ₂	238.24	
BDE-153	155166	C ₁₂ H ₄ Br ₆ O	643.6	
Carbazole	6854	C ₁₂ H ₉ N	167.21	
Indole	798	C ₈ H ₇ N	117.15	
Acridine	9215	C ₁₃ H ₉ N	179.22	
bipyridines	14338	C ₁₂ H ₁₂ N ₂	184.24	Polycyclic Aromatic Hydrocarbons (PAHs)
terpyridines	70848	C ₁₅ H ₁₁ N ₃	233.27	
4H-1,2,3,5-Benzotetrazepin-4-one, 3,5-dihydro-3,5-dimethyl-7-nitro-	154099	C ₉ H ₉ N ₅ O ₃	235.20	
Anthracene	8418	C ₁₄ H ₁₀	178.23	
Fluorene	6853	C ₁₃ H ₁₀	166.22	
Chrysene	9171	C ₁₈ H ₁₂	228.3	
Benzo[a]pyrene	2336	C ₂₀ H ₁₂	252.3	

using the CB-Dock2 server. The docking procedure was implemented to evaluate the binding affinities of the enzyme with 34 micropollutant ligands (Table 4).

After molecular docking, the previously selected ligands belonging to different application categories showed remarkable Vina scores (Fig. 5; Table 5). The molecular docking results for pesticide-type micropollutants were favorable, demonstrating a notable binding affinity for the enzyme. Among them, fluopyram exhibited the strongest affinity, with a Vina score of -7.7 kcal/mol. In addition, diflubenazuron displayed a comparable binding energy of -7.4 kcal/mol, suggesting a relatively strong interaction with the enzyme active site. These findings are consistent with previous studies on the enzyme AmiH52, which belongs to the amidohydrolase (amidase) family and has been investigated through both experimental and *in silico* approaches. This enzyme has demonstrated the ability to degrade amide pesticides, including diflubenazuron [44]. Whereas 2,6-dichlorobenzamide showed the lowest estimated value (-4.7 kcal/mol). Indeed, an experimental study identified an amidase (BbdA) with a low K_m value ($0.71 \mu\text{M}$) for 2,6-dichlorobenzamide, indicating a high affinity of the enzyme for this substrate. Furthermore, the authors reported that BbdA could efficiently degrade 2,6-dichlorobenzamide, even at low concentrations [45].

Therefore, the agreement between the docking results and experimental evidence supports the hypothesis that enzymes belonging to the amidase family play an important role in the bioremediation of environments contaminated with amide pesticides. In addition, a Vina score of -7.8 kcal/mol was obtained for sulfadiazine, representing the lowest binding energy among pharmaceutical micropollutants. n [4]-acetylsulfamethoxazole showed a Vina score of -7.5 kcal/mol, reflecting good binding affinity, consistent with the recent study of Marti *et al.*, who demonstrated the potential of microbial Amidase Signature (AS) enzymes to catalyze the biotransformation of acetylsulfamethoxazole [46]. Moreover, for industrial micropollutants, the docking scores obtained ranged from -7.3 to -5.1 kcal/mol, corresponding to 2-methylantraquinone and 2-methylbenzenesulfonamide, respectively. These results further support the ability of the enzyme to interact with these compounds.

Research on the degradation of plastic micropollutants by acylamidase has been relatively well explored but is still limited to certain types of compounds. In this study, we evaluated its potential on ten representative ligands of these micropollutants. Among them, triphenyl phosphate exhibited a favorable docking score (-8.0 kcal/mol), whereas the remaining nine ligands showed scores ranging from -7.6 to -5.7 kcal/mol, indicating that they can also interact with the enzyme's active site. These results suggest that acylamidase may effectively recognize and bind various plastic micropollutants. Furthermore, studies on Amidase Signature (AS) enzymes have shown that they can hydrolyze diverse aromatic micropollutants [46]. In this context, our findings indicate that acylamidase may also interact with nitrogen-containing aromatic compounds such as carbazole and acridine, underscoring the enzyme's versatility and its broad substrate recognition. Regarding Polycyclic Aromatic Hydrocarbons (PAHs), including benzo[a]pyrene, chrysene, fluorene, and anthracene, benzo[a]pyrene exhibited the most favorable molecular docking results, with the lowest binding energy (-8.8 kcal/mol), indicating a high binding affinity toward the enzyme. These organic molecules are characterized by the fusion of at least two benzene rings (aromatic cycles) [47]. They primarily originate from the incomplete combustion of organic matter but are also naturally present in petroleum products [48]. In the petroleum sector, polycyclic aromatic hydrocarbons (PAHs) are generated through the exploration, extraction, transportation, and refining activities of crude oil, as well as through the incomplete combustion of petroleum derivatives [47].

In this context, our results highlight that exploiting the transformation of organic pollutants, particularly using the capabilities of acylamidase, is a promising and well-established approach. The observed interactions between the enzyme and these compounds suggest a favorable binding affinity, reinforcing the hypothesis of its involvement in bioremediation processes.

To our knowledge, this is the first study reporting the ability of acylamidase to interact with polycyclic aromatic hydrocarbons, which highlights the significance of our findings.

The analysis of the results generated by CB-Dock2 (Figs. 6 and 7) enabled the identification of the most promising ligand based on the

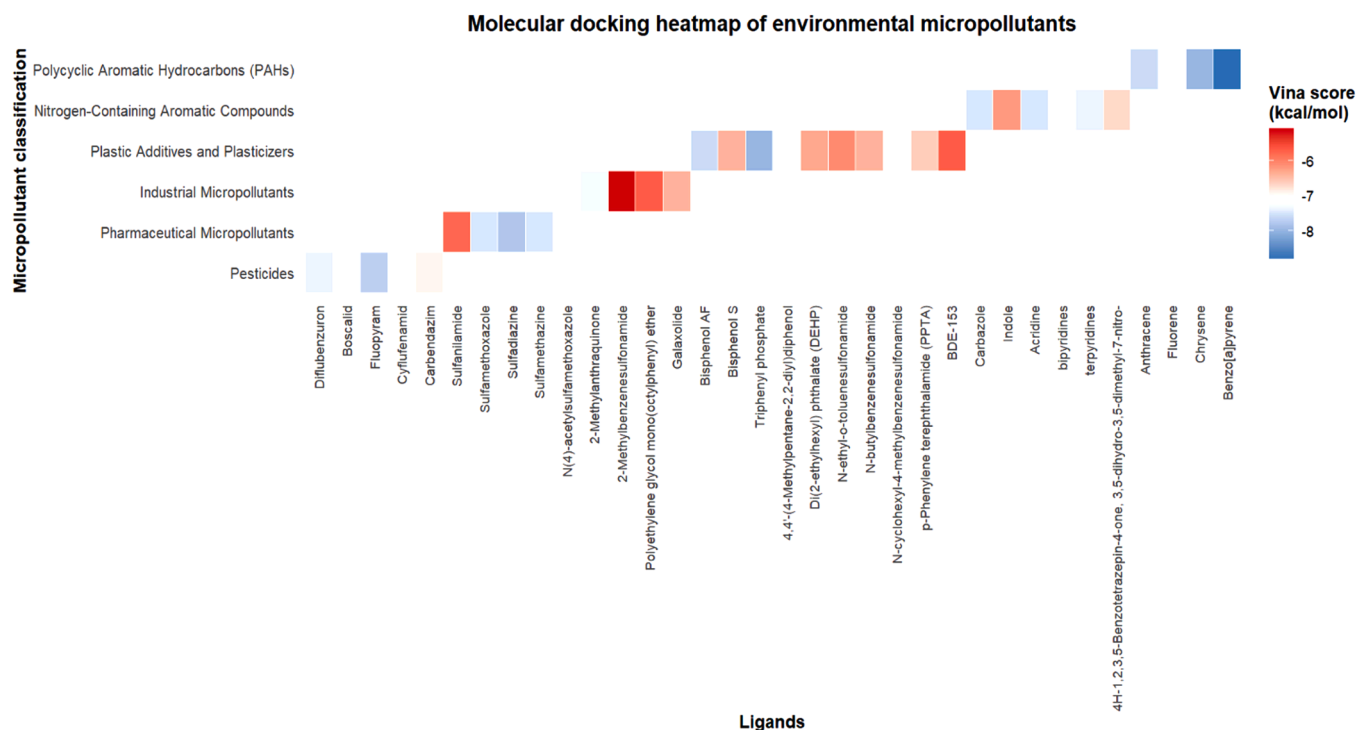


Fig. 5. Vina score heatmap illustrating enzyme–ligand interactions for different classes of micropollutants. The colors indicate the binding energy (kcal/mol), ranging from red (low affinity) to blue (high affinity).

Table 5

Minimum binding energy predicted by the CB-Dock web server using Vina scores.

Ligands name	Vina score (kcal/mol)	Classification of Environmental Micropollutants
Diflubenzuron	−7.4	Pesticides
Boscalid	−7.0	
Fluopyram	−7.7	
Cyflufenamid	−7.1	
2,6-dichlorobenzamide	−4.7	
Carbendazim	−6.9	
Sulfanilamide	−5.8	
Sulfamethoxazole	−7.5	
Sulfadiazine	−7.8	
Sulfamethazine	−7.5	
N [4]-acetylsulfamethoxazole	−7.0	Industrial Micropollutants
2-Methylanthraquinone	−7.3	
2-Methylbenzenesulfonamide	−5.1	
Polyethylene glycol mono(octylphenyl) ether	−5.7	
Galaxolide	−6.4	
Bisphenol AF	−7.6	Plastic Additives and Plasticizers
Bisphenol S	−6.4	
Triphenyl phosphate	−8.0	
4,4'-(4-Methylpentane-2,2-diyl) diphenol	−7.1	
Di(2-ethylhexyl) phthalate (DEHP)	−6.3	
N-ethyl-o-toluenesulfonamide	−6.1	Nitrogen-Containing Aromatic Compounds
N-butylbenzenesulfonamide	−6.4	
N-cyclohexyl-4-methylbenzenesulfonamide	−7.2	
p-Phenylene terephthalamide (PPTA)	−6.6	
BDE-153	−5.7	
Carbazole	−7.5	
Indole	−6.2	
Acridine	−7.5	
bipyridines	−7.1	
terpyridines	−7.4	
4H-1,2,3,5-Benzotetrazepin-4-one, 3,5-dihydro-3,5-dimethyl-7-nitro-	−6.7	Polycyclic Aromatic Hydrocarbons (PAHs)
Anthracene	−7.6	
Fluorene	−7.1	
Chrysene	−8.0	
Benzo[a]pyrene	−8.8	

lowest binding energy values (Vina scores). Benzo[a]pyrene interacts with key active-site residues, namely ARG123, PHE129, ALA64, and ALA67 (Fig. 7). In contrast, sulfadiazine exhibited a more extensive interaction network with acylamidase, including residues ALA64, ALA67, ARG78, TRP97, ARG123, PHE129, LEU122, and VAL119 (Fig. 6). In addition, the presence of conventional hydrogen bonds with the ligands contributes to the stabilization of the enzyme–ligand complex. However, members of the Amidase Signature enzyme family are characterized by a canonical Ser-cisSer-Lys catalytic triad [49]. However, our enzyme did not exhibit these residues in its interactions with the selected ligands.

3.8. Molecular dynamics simulations

MDS was performed to evaluate the structural stability, flexibility, compactness, and solvent accessibility of acylamidase (ACL) upon binding with benzo[a]pyrene (BaP), triphenyl phosphate (TPhP), and sulfadiazine (SDZ) [50]. According to Fig. 8 (A), the total energy of BaP–ACL, TPhP–ACL, and SDZ–ACL complexes remained constant throughout the 100 ns simulation. RMSD analysis was further employed to evaluate the conformational stability of the protein backbone during the simulation. As shown in Fig. 8 (B), BaP–ACL complex showed a slightly higher RMSD, while TPhP–ACL and SDZ–ACL complexes demonstrated lower RMSD values when compared to ACL alone. These results suggest that BaP binding induced small structural fluctuations

within the protein backbone, while TPhP and SDZ binding induce structural stabilization. Nevertheless, the RMSD values remained below 0.35 nm, indicating that the complex maintained an overall stable conformation [51]. The residue-level flexibility of a protein was investigated using RMSF analysis. Interestingly, the RMSF of ACL and all complexes show similar fluctuation patterns throughout the 100 ns simulation (Fig. 8 (C)). The low RMSF values indicate limited residue mobility and suggest structural stability of the protein during the simulation period [52]. The compactness of the protein structure was then assessed through Rg analysis. According to Fig. 8 (D), BaP–ACL complex displayed a slightly higher Rg value, while TPhP–ACL and SDZ–ACL complexes showed lower Rg values when compared to ACL alone. These results denote that upon BaP binding, the ligand induces minor loosening compactness of the protein, while TPhP and SDZ binding may contribute to enhanced conformational compactness of ACL [53]. Finally, SASA analysis was conducted to evaluate the changes in protein surface exposure to the solvent environment. From Fig. 8 (E), the SASA for all complexes was comparable to the ACL alone, indicating a minimal alteration in solvent exposure upon ligand. Overall, the low RMSD and RMSF values together with stable Rg and SASA profiles support the formation of stable ligand–protein complexes during the 100 ns simulation period (Table 6).

3.8.1. MM/GBSA analysis

The binding free energies of all complexes were further evaluated using the MM/GBSA approach. The calculated average of ΔH , $-T\Delta S$, and ΔG are presented in Table 7. According to Fig. 9 (A), the ΔH of all complexes remained negative value during 100 ns simulation which indicate favorable interactions between the ligands and ACL. The more negative ΔH suggests stronger non-covalent interactions, involving van der Waals forces, hydrogen bonds and electrostatic interactions [54]. The ΔH values followed the order of TPhP > SDZ > BaP toward ACL, with TPhP exhibiting the most negative ΔH value. Subsequently, the $-T\Delta S$ was analyzed where the positive value suggests changes in system disorder upon ligand binding, commonly associated with hydrophobic interactions [55]. From Fig. 9 (B), the $-T\Delta S$ of all complexes remained positive value during 100 ns simulation. The $-T\Delta S$ values followed the order of SDZ > BaP > TPhP toward ACL, with SDZ exhibiting the most positive $-T\Delta S$ value. The combination of negative ΔH and positive $-T\Delta S$ values indicates that the ligand–protein binding are mainly stabilized by hydrogen bonding, hydrophobic interactions, and van der Waals forces [56]. From Table 7, all ligand–ACL complexes exhibited negative ΔG values, suggesting thermodynamically favorable and spontaneous binding interactions [55]. However, in Fig. 9 (C), the ΔG trajectory of SDZ–ACL complex displayed fluctuations, suggesting a reduced binding interaction after 40 ns. Collectively, these findings suggest the ΔG trend of TPhP > BaP > SDZ toward ACL, with TPhP exhibiting the most negative ΔG value.

4. Conclusions

In this study, we combined homology modeling, molecular docking, and explicit-solvent molecular dynamics simulations to comprehensively characterize an acylamidase enzyme from *Actinoalloteichus hogarensis* and evaluate its potential for the bioremediation of environmental micropollutants for bioremediation applications. Using an integrated *in silico* approach, we analyzed the enzyme's physico-chemical properties, revealing its acidic nature, structural stability, and hydrophilic character. In addition, key structural features were predicted and its functional properties were further characterized. Molecular docking analyses suggested that several environmental contaminants may be susceptible to enzymatic interaction, as indicated by favorable binding affinities with acylamidase. Notably, polycyclic aromatic hydrocarbons (PAHs) exhibited strong interactions with the enzyme, with benzo[a]pyrene showing the highest binding affinity, with a Vina score of -8.8 kcal/mol.

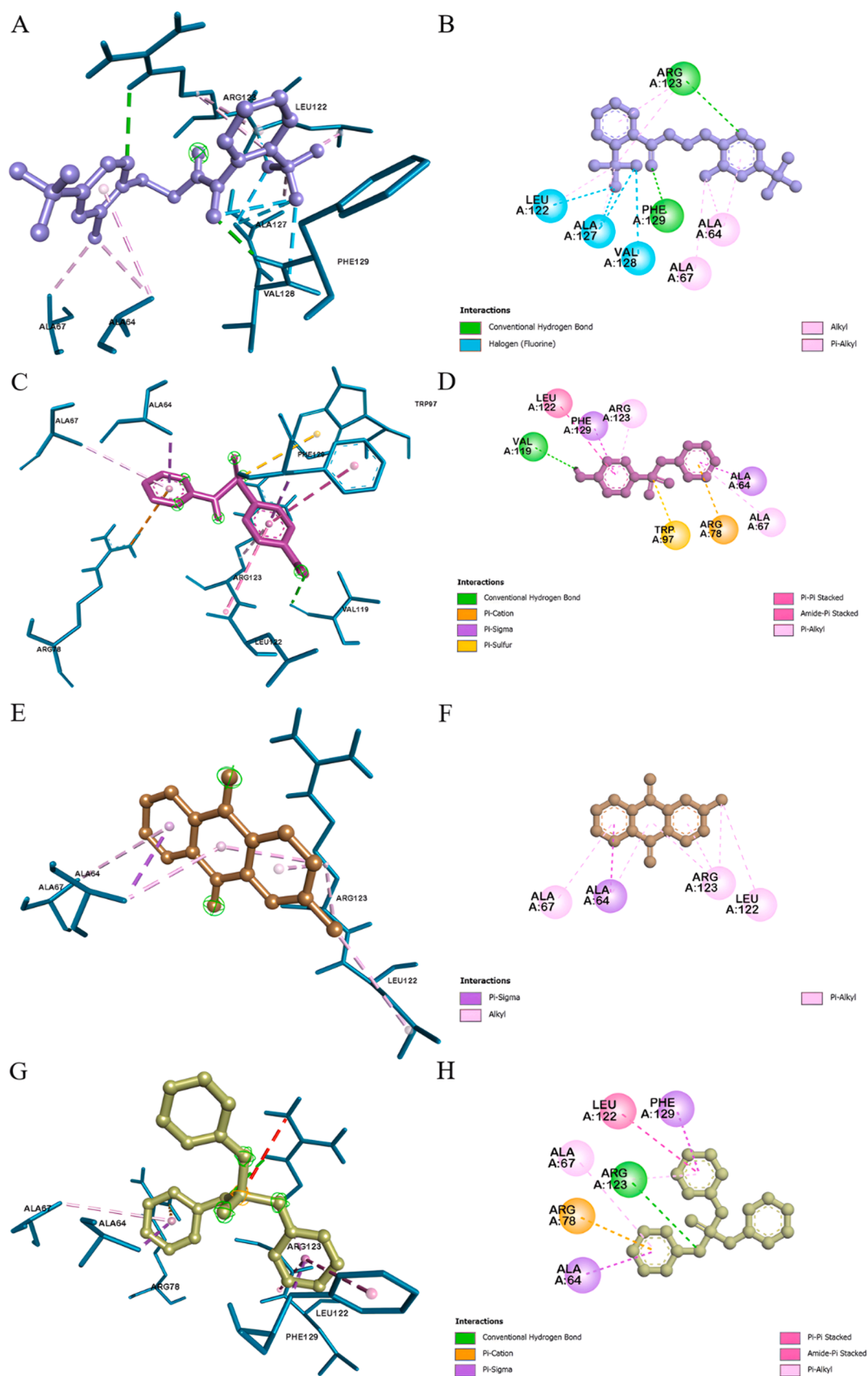


Fig. 6. Interactions between different micropollutant ligands and acylamidase enzymes. (A and B) refer to fluopyram, which is classified as a pesticide; (C and D) refer to sulfadiazine, classified as a pharmaceutical micropollutant; (E and F) refer to 2-Methylantraquinone, classified as an industrial micropollutant; and (G and H) refer to triphenyl phosphate, used in plastics as an additive and plasticizer. (A, C, E, G) 3D visualization of the ligand-receptor interacting residues, using BIOVIA Discovery Studio Visualizer; (B, D, F, H) 2D schematic diagram (BIOVIA Discovery Studio Visualizer).

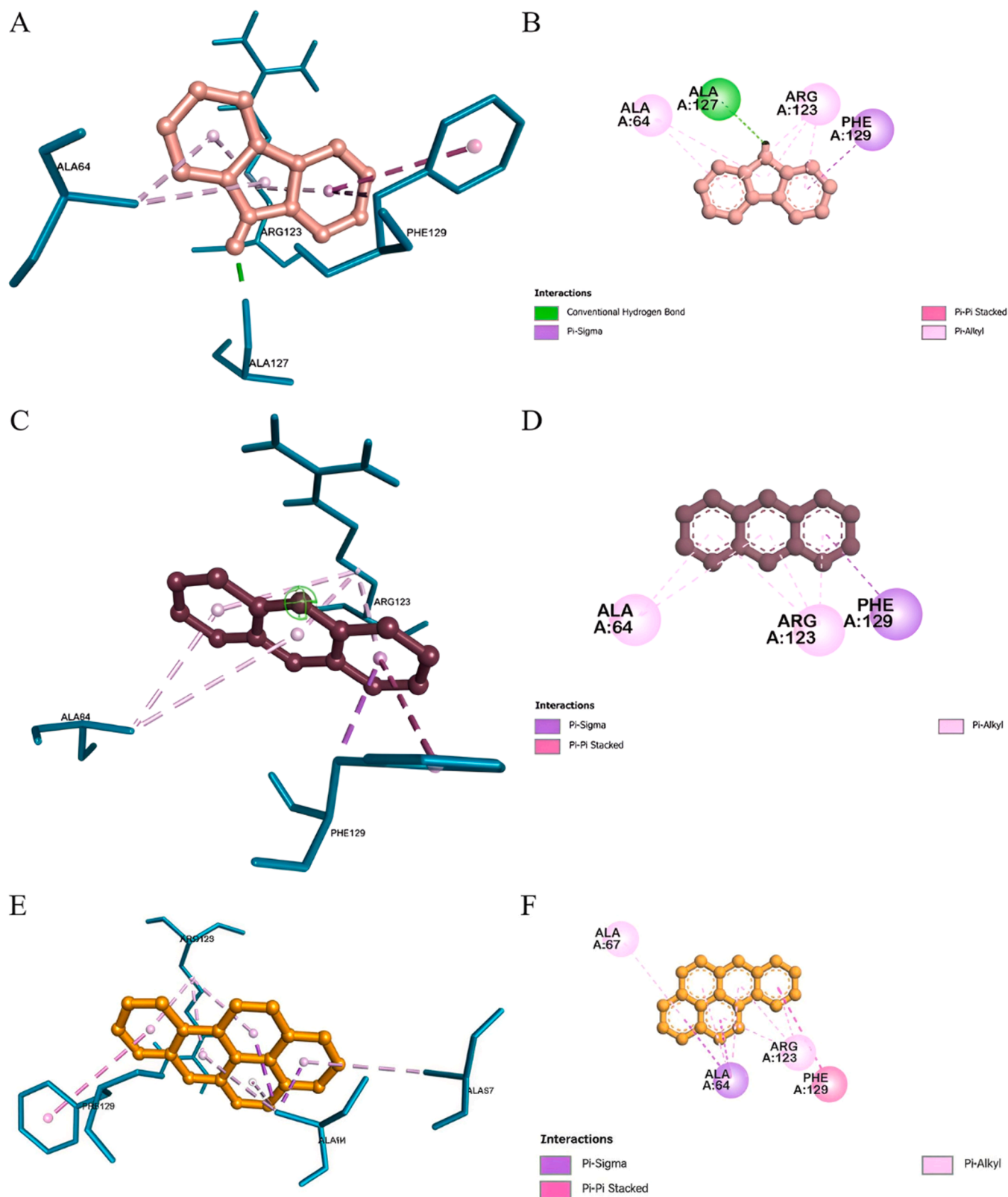


Fig. 7. Interactions between micropollutant ligands from the petroleum industry and acylamidase enzymes. (A and B) refer to carbazole, while (C and D) refer to acridine, both of which are nitrogen-containing aromatic compounds. (E and F) correspond to benzo[a]pyrene, which is classified as a polycyclic aromatic hydrocarbon (PAH); (A, C, E) 3D visualization of the ligand-receptor interacting residues, using BIOVIA Discovery Studio Visualizer. (B, D, F) 2D schematic diagram (BIOVIA Discovery Studio Visualizer).

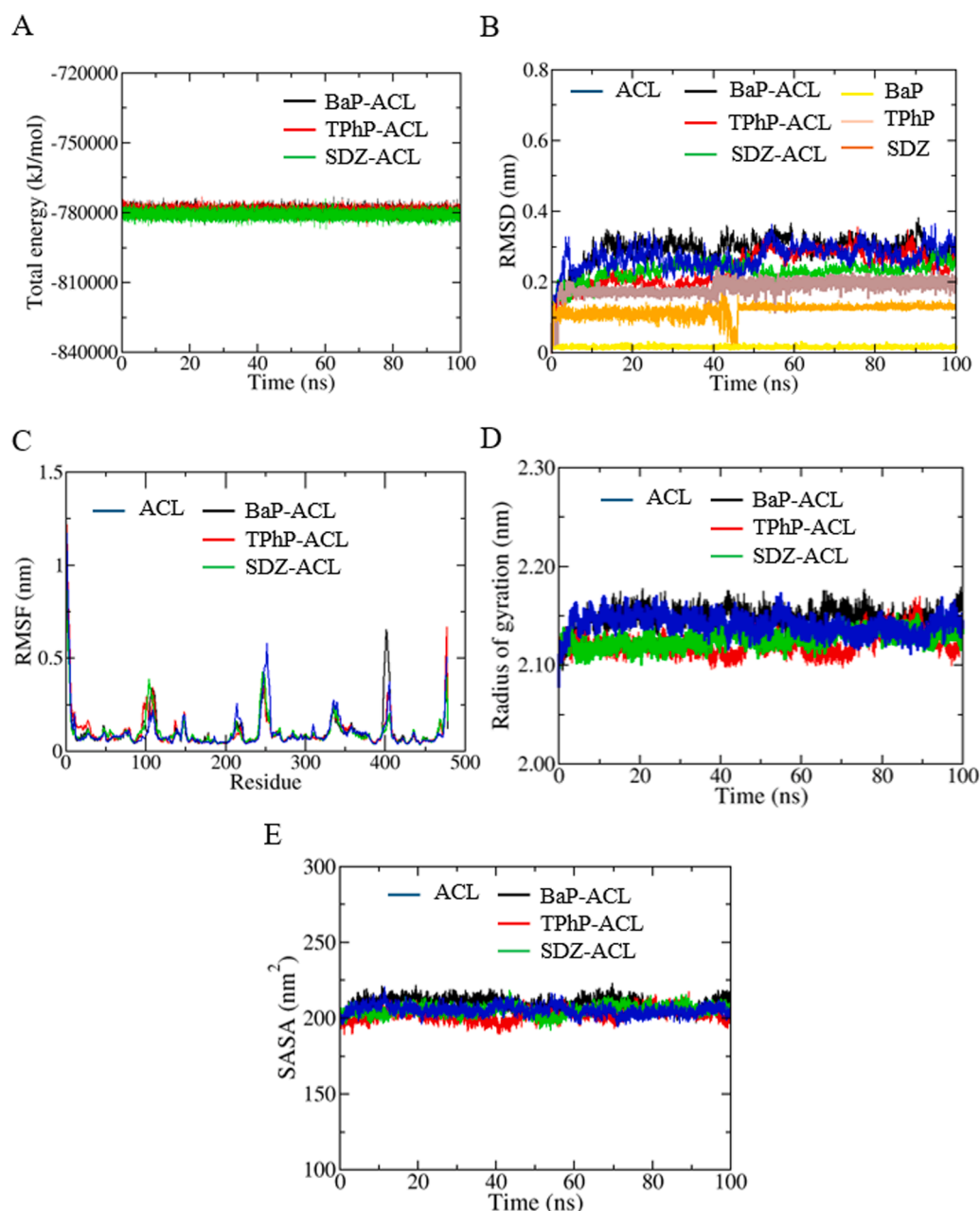


Fig. 8. MDS data obtained for ACL, BaP-ACL, TPhP-ACL and SDZ-ACL complexes over 100 ns. Graph (A) the total energy over time, (B) RMSD over time, (C) RMSF over residue, (D) Rg over time (E) SASA over time. RMSD, RMSF and Rg were calculated based on the α -carbon atoms, whereas SASA was determined using the entire protein structure.

Table 6

RMSD, RMSF, Rg and SASA for ACL, BaP-ACL, TPhP-ACL and SDZ-ACL complex obtained from 100 ns MDS. All data are expressed as mean $\pm$ standard error.

	RMSD (nm)	RMSF (nm)	Rg (nm)	SASA (nm ²)
ACL	0.270 $\pm$ 0.000	0.103 $\pm$ 0.005	2.141 $\pm$ 0.000	205.2 $\pm$ 0.0
BaP-ACL	0.291 $\pm$ 0.000	0.106 $\pm$ 0.005	2.148 $\pm$ 0.000	209.7 $\pm$ 0.0
TPhP-ACL	0.239 $\pm$ 0.000	0.107 $\pm$ 0.005	2.122 $\pm$ 0.000	202.0 $\pm$ 0.0
SDZ-ACL	0.230 $\pm$ 0.000	0.102 $\pm$ 0.004	2.127 $\pm$ 0.000	205.5 $\pm$ 0.0

Molecular dynamics simulations over 100 ns provided critical validation of these docking results. The enzyme–ligand complexes (with benzo[a]pyrene, triphenyl phosphate, and sulfadiazine) maintained stable conformations, as evidenced by low root-mean-square deviation (RMSD < 0.35 nm) and fluctuation (RMSF) values. Radius of gyration

Table 7

MM/GBSA data for the BaP-ACL, TPhP-ACL, and SDZ-ACL complexes are presented as mean $\pm$ standard error. All energy components are expressed in kJ/mol.

	BaP-ACL	TPhP-ACL	SDZ-ACL
ΔH	−100.25 $\pm$ 0.29	−130.00 $\pm$ 0.38	−112.59 $\pm$ 0.50
$-T\Delta S$	40.67 $\pm$ 0.00	38.07 $\pm$ 0.00	107 $\pm$ 0.00
ΔG	−59.58 $\pm$ 9.46	−91.88 $\pm$ 11.46	−4.81 $\pm$ 15.61

and solvent-accessible surface area analyses indicated that ligand binding did not induce major structural destabilization. Critically, MM/GBSA binding free energy calculations confirmed thermodynamically favorable and spontaneous interactions for all complexes, with triphenyl phosphate exhibiting the most negative ΔG (−91.88 kJ/mol), followed

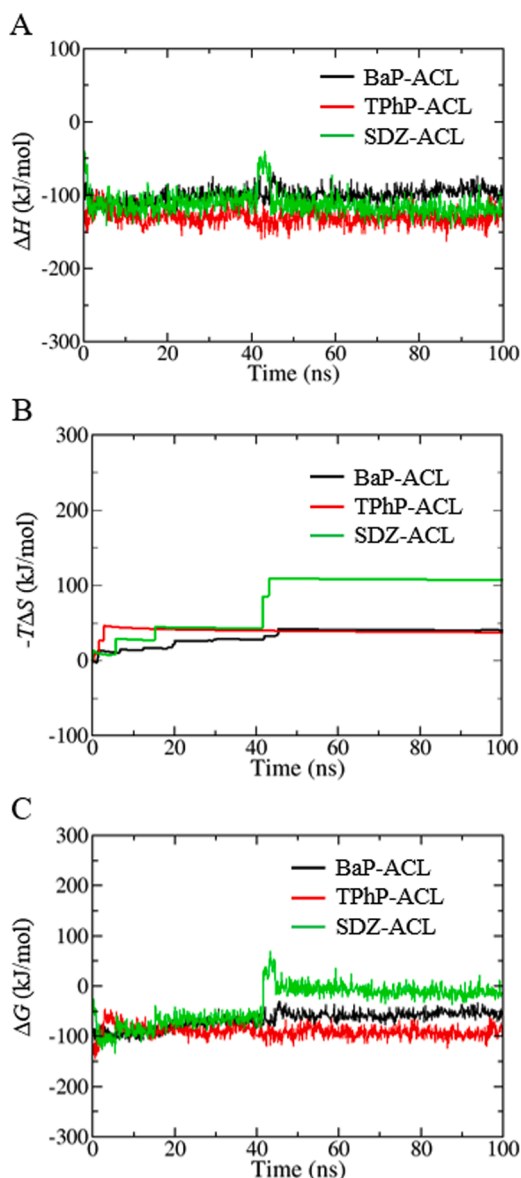


Fig. 9. Graph representations of (A) enthalpy, (B) entropy and (C) binding free energy derived from MM/GBSA data for BaP-ACL, TPhP-ACL, and SDZ-ACL complexes over 100 ns.

by benzo[a]pyrene (-59.58 kJ/mol). These results suggest that while all tested ligands can bind, triphenyl phosphate may form the most energetically stable complex under equilibrium conditions.

Collectively, these computational findings suggest that the acylamidase from *A. hoggarensis* possesses the structural and dynamic determinants to accommodate both PAHs and other organic micropollutants, providing a computational foundation for its application in enzymatic bioremediation. Nevertheless, further experimental investigations are required to validate these *in silico* predictions and to better elucidate the catalytic mechanisms involved, including enzyme activity assays, identification of degradation products, and directed mutagenesis studies to confirm the functional roles of key active-site residues identified in this work.

CRediT authorship contribution statement

Sara Zergoun: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Investigation, Formal analysis, Data curation. **Ikrame Abayahia:** Writing – review & editing, Writing –

original draft, Visualization, Validation, Software, Investigation, Formal analysis, Data curation. **Mahfoud Bakli:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Formal analysis, Data curation, Conceptualization. **Fatih Şaban Beriş:** Writing – review & editing, Writing – original draft. **Ahmad Hidayat Fadhlurrahman:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Investigation, Formal analysis. **Saharuddin Bin Mohamad:** Writing – review & editing, Writing – original draft. **Muhammad Munir:** Writing – review & editing, Writing – original draft, Funding acquisition.

Funding

The authors extend their gratitude to the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Kingdom of Saudi Arabia, for funding the publication of this work under the Ambitious Researcher program (Project No. KFU263256).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

- [1] Ehrl BN, Kundu K, Gharasoo M, Marozava S, Elsnar M. Rate-limiting mass transfer in micropollutant degradation revealed by isotope fractionation in chemostat. *Environ Sci & Technol* 2019;53(3):1197–205. <https://doi.org/10.1021/acs.est.8b05175>.
- [2] OECD. Pharmaceutical Residues in Freshwater: Hazards and Policy Responses. OECD Publishing; 2019. <https://doi.org/10.1787/c936f42d-en>.
- [3] Zhu X, Liu S, Gao X, Gu Y, Yu Y, Li M, et al. Typical emerging contaminants in sewage treatment plant effluent, and related watersheds in the Pearl River Basin: Ecological risks and source identification. *J Hazard Mater* 2024;476:135046. <https://doi.org/10.1016/j.jhazmat.2024.135046>.
- [4] Wolowicz A, Munir HMS. Emerging organic micropollutants as serious environmental problem: A comprehensive review. *Sci Total Environ* 2025;958:177948. <https://doi.org/10.1016/j.scitotenv.2024.177948>.
- [5] Wang X, Zhai Y, Deng X, van der Meer W, Pabst M, van Loosdrecht M. Prospects for Using Multi-omics to Advance the Bioremediation of Organic Micropollutants in Aquatic Environments: A Review. *Water & Ecol* 2026;2(1):100036. <https://doi.org/10.1016/j.wateco.2026.100036>.
- [6] Ali A, Zedan TH, Hassan MHU, Ali N, Naeem Z, Devi K, et al. A review of CRISPR-engineered microbial consortia for biodegradation of petroleum pollutants. *Desalin Water Treat* 2026;325:101612. <https://doi.org/10.1016/j.dwt.2025.101612>.
- [7] Sinha S. Construction and Functional Screening of Metagenomic Libraries for D-Stereospecific Amidases and Their Catalytic Promiscuity. *Appl Biochem Microbiol* 2024;60(4):614–26. <https://doi.org/10.1134/S0003683823601889>.
- [8] Lavrov KV, Yanenko AS. Cloning of new acylamidase gene from *Rhodococcus erythropolis* and its expression in *Escherichia coli*. *Russ J Genet* 2013;49(10):1078–81. <https://doi.org/10.1134/S1022795413070090>.
- [9] Boudjelal F, Zitouni A, Bouras N, Schumann P, Spröer C, Sabaou N, et al. *Actinoboliteichus hoggarensis* sp. nov., an actinomycete isolated from Saharan soil. *Int J Syst Evol Microbiol* 2015;65(Pt 6):2006–10. <https://doi.org/10.1099/ijso.0.000216>.
- [10] Xie F, Pathom-Aree W. Actinobacteria From Desert: Diversity and Biotechnological Applications. *Front Microbiol* 2021;12:765531. <https://doi.org/10.3389/fmicb.2021.765531>.
- [11] Bakli M, Bouras N, Paşcalău R, Şmuleac L. *In silico* Structural and Functional Characterization of an Endoglucanase from *Actinoboliteichus hoggarensis* 2023;7(1):135–41. <https://doi.org/10.2478/arls-2023-0015>.
- [12] Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press; 2005. p. 571–607. <https://doi.org/10.1385/1-59259-890-0:571>.
- [13] Nielsen H. Practical Applications of Language Models in Protein Sorting Prediction: SignalP 6.0, DeepLoc 2.1, and DeepLocPro 1.0. In: Kc DB, editor. *Large Language Models (LLMs) in Protein Bioinformatics*. New York, NY: Springer US; 2025. p. 153–75. https://doi.org/10.1007/978-1-0716-4623-6_10.
- [14] Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories

- and predictive capabilities for all prokaryotes. *Bioinforma (Oxf Engl)* 2010;26(13): 1608–15. <https://doi.org/10.1093/bioinformatics/btq249>.
- [15] Geourjon C, Deléage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci* 1995;11(6):681–4. <https://doi.org/10.1093/bioinformatics/11.6.681>.
- [16] Buchan DWA, Jones DT. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res* 2019;47(W1):W402–7. <https://doi.org/10.1093/nar/gkz297>.
- [17] Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinforma (Oxf Engl)* 2006;22(2):195–201. <https://doi.org/10.1093/bioinformatics/bti770>.
- [18] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 2015;10(6):845–58. <https://doi.org/10.1038/nprot.2015.053>.
- [19] Fleming J, Magana P, Nair S, Tsenkov M, Bertoni D, Piduchna I, et al. AlphaFold Protein Structure Database and 3D-Beacons: New Data and Capabilities. *J Mol Biol* 2025;437(15):168967. <https://doi.org/10.1016/j.jmb.2025.168967>.
- [20] Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 2015;12(1):7–8. <https://doi.org/10.1038/nmeth.3213>.
- [21] Du Z, Su H, Wang W, Ye L, Wei H, Peng Z, et al. The trRosetta server for fast and accurate protein structure prediction. *Nat Protoc* 2021;16(12):5634–51. <https://doi.org/10.1038/s41596-021-00628-9>.
- [22] Heo L, Park H, Seok C. GalaxyRefine: Protein structure refinement driven by side-chain repacking (Web Server issue). *Nucleic Acids Res* 2013;41:W384–8. <https://doi.org/10.1093/nar/gkt458>.
- [23] Laskowski R., MacArthur M., Thornton J. PROCHECK: validation of protein-structure coordinates. 2006.
- [24] Mahuri M, Paul M, Sahu SK, Thatoi H. Computational investigations involving the structural, functional, and molecular dynamics analysis of bacterial laccase to unravel its role in lignin biodegradation. *J Biomol Struct Dyn* 2026;44(3):1581–94. <https://doi.org/10.1080/07391102.2025.2499216>.
- [25] Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwazd M, Lu S, et al. The conserved domain database in 2023. *Nucleic Acids Res* 2022;51(D1):D384–8. <https://doi.org/10.1093/nar/gkac1096>.
- [26] Liu Y, Yang X, Gan J, Chen S, Xiao Z-X, Cao Y. CB-Dock2: improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Res* 2022;50(W1):W159–64. <https://doi.org/10.1093/nar/gkac394>.
- [27] Liu Y, Grimm M, Dai W-t, Hou M-c, Xiao Z-X, Cao Y. CB-Dock: a web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacol Sin* 2020;41(1):138–44. <https://doi.org/10.1038/s41401-019-0228-6>.
- [28] Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem Substance and Compound databases. *Nucleic Acids Res* 2015;44(D1):D1202–13. <https://doi.org/10.1093/nar/gkv951>.
- [29] Croitoru A, Park S-J, Kumar A, Lee J, Im W, MacKerell Jr AD, et al. Additive CHARMM36 Force Field for Nonstandard Amino Acids. *J Chem Theory Comput* 2021;17(6):3554–70. <https://doi.org/10.1021/acs.jctc.1c00254>.
- [30] Bugnon M, Goullieux M, Röhrig UF, Perez MAS, Daina A, Michielin O, et al. SwissParam 2023: A Modern Web-Based Tool for Efficient Small Molecule Parametrization. *J Chem Inf Model* 2023;63(21):6469–75. <https://doi.org/10.1021/acs.jcim.3c01053>.
- [31] Herrera-Rodríguez AM, Miletić V, Aponte-Santamaría C, Gräter F. Molecular Dynamics Simulations of Molecules in Uniform Flow. *Biophys J* 2019;116(9): 1579–85. <https://doi.org/10.1016/j.bpj.2018.12.025>.
- [32] Simmonett AC, Brooks BR. A compression strategy for particle mesh Ewald theory. *J Chem Phys* 2021;154(5):054112. <https://doi.org/10.1063/5.0040966>.
- [33] Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA, Moreno E. gmx-MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J Chem Theory Comput* 2021;17(10):6281–91. <https://doi.org/10.1021/acs.jctc.1c00645>.
- [34] Bhattacharya M, Hota A, Kar A, Sankar Chini D, Chandra Malick R, Chandra Patra B, et al. *In silico* structural and functional modelling of Antifreeze protein (AFP) sequences of Ocean pout (*Zoarces americanus*, Bloch & Schneider 1801). *J Genet Eng Biotechnol* 2018;16(2):721–30. <https://doi.org/10.1016/j.jgeb.2018.08.004>.
- [35] Farjami T, Sharma A, Jensen I-J, Falch E. A proteomics–bioinformatics approach to assess the molecular and physicochemical properties of brewer's spent grain protein concentrates: Effect of alkali-heat treatment. *J Agric Food Res* 2024;18: 101543. <https://doi.org/10.1016/j.jafr.2024.101543>.
- [36] Sahay A, Piprodhe A, Pise M. *In silico* analysis and homology modeling of strictosidine synthase involved in alkaloid biosynthesis in *Catharanthus roseus*. *J Genet Eng Biotechnol* 2020;18(1):44. <https://doi.org/10.1186/s43141-020-00049-3>.
- [37] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 1982;157(1):105–32. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0).
- [38] Bouqellah NA, Farag PF. *In Silico* Evaluation, Phylogenetic Analysis, and Structural Modeling of the Class II Hydrophobin Family from Different Fungal Phytopathogens 2023;11(11):2632. <https://doi.org/10.3390/microorganisms11112632>.
- [39] El-Sayed ASA, Elghamry HN, Yassin MA. Biochemical Characterization of Thermostable Acrylamide Amidohydrolase from *Aspergillus fumigatus* with Potential Activity for Acrylamide Degradation in Various Food Products. *Curr Microbiol* 2023;81(1):30. <https://doi.org/10.1007/s00284-023-03544-1>.
- [40] Terwilliger TC, Liebschner D, Croll TI, Williams CJ, McCoy AJ, Poon BK, et al. AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination. *Nat Methods* 2024;21(1):110–6. <https://doi.org/10.1038/s41592-023-02087-4>.
- [41] Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018;46(W1):W296–303. <https://doi.org/10.1093/nar/gky427>.
- [42] Chadha A, Yanai Y, Oide H, Wakana Y, Inoue H, Saha S, et al. *Legionella* uses host Rab GTPases and BAP31 to create a unique ER niche. *Cell Rep* 2024;43(12). <https://doi.org/10.1016/j.celrep.2024.115053>.
- [43] Fong JH, Geer LY, Panchenko AR, Bryant SH. Modeling the Evolution of Protein Domain Architectures Using Maximum Parsimony. *J Mol Biol* 2007;366(1): 307–15. <https://doi.org/10.1016/j.jmb.2006.11.017>.
- [44] Sun S, Chen W, Peng K, Chen X, Chen J. Characterization of a novel amidohydrolase with promiscuous esterase activity from a soil metagenomic library and its application in degradation of amide herbicides. *Environ Sci Pollut Res* 2024;31(14):20970–82. <https://doi.org/10.1007/s11356-024-32362-6>.
- [45] T'Syen J, Tassoni R, Hansen L, Sorensen SJ, Leroy B, Sekhar A, et al. Identification of the amidase bdda that initiates biodegradation of the groundwater micropollutant 2,6-dichlorobenzamide (BAM) in *Aminobacter* sp. MSH1. *Environ Sci & Technol* 2015;49(19):11703–13. <https://doi.org/10.1021/acs.est.5b02309>.
- [46] Marti TD, Schweizer D, Yu Y, Schärer MR, Probst SI, Robinson SL. Machine Learning Reveals Signatures of Promiscuous Microbial Amidases for Micropollutant Biotransformations. *ACS Environ Au* 2025;5(1):114–27. <https://doi.org/10.1021/acsenvironau.4c00066>.
- [47] Hidalgo KJ, Sierra-Garcia IN, Dellagennazze BM, de Oliveira VM. Metagenomic Insights into the Mechanisms for Biodegradation of Polycyclic Aromatic Hydrocarbons in the Oil Supply Chain 2020;11:2020. <https://doi.org/10.3389/fmicb.2020.561506>.
- [48] Abdel-Shafy HI, Mansour MSM. A review on polycyclic aromatic hydrocarbons: Source, environmental impact, effect on human health and remediation. *Egypt J Pet* 2016;25(1):107–23. <https://doi.org/10.1016/j.ejpe.2015.03.011>.
- [49] Lee S, Park E-H, Ko H-J, Bang WG, Kim H-Y, Kim KH, et al. Crystal structure analysis of a bacterial aryl acylamidase belonging to the amidase signature enzyme family. *Biochem Biophys Res Commun* 2015;467(2):268–74. <https://doi.org/10.1016/j.bbrc.2015.09.177>.
- [50] Sinha S, Tam B, Wang SM. Applications of molecular dynamics simulation in protein study. *Membranes* 2022;12(9). <https://doi.org/10.3390/membranes12090844>.
- [51] Halder SK, Sultana I, Shuvo MN, Shil A, Himel MK, Hasan MA, et al. *In Silico* Identification and Analysis of Potentially Bioactive Antiviral Phytochemicals against SARS-CoV-2: A Molecular Docking and Dynamics Simulation Approach. *BioMed Res Int* 2023;2023(1):5469258. <https://doi.org/10.1155/2023/5469258>.
- [52] Soremekun C, Jjingo D, Kateete D, Nash O, Grallert H, Peters A, et al. Structural insights into conformational stability of both wild-type and mutant Insulin Receptor Gene. *Research* 2024;1(2):100041. <https://doi.org/10.1016/j.nexres.2024.100041>.
- [53] Alawi MM, Gattan HS, Alharbi AS, Alruhaili MH, Al-Zahrani IA, Masaud SBAL, et al. Suppressing lipid biosynthesis in *Mycobacterium tuberculosis* through polyketide synthase 13 thioesterase inhibition: Insights from computational analysis. *J Infect Public Health* 2025;18(9):102835. <https://doi.org/10.1016/j.jiph.2025.102835>.
- [54] Du X, Li Y, Xia Y-L, Ai S-M, Liang J, Sang P, et al. Insights into protein–ligand interactions: mechanisms, models, and methods. *Int J Mol Sci* 2016;17(2):144.
- [55] Garbett NC, Chaires JB. Thermodynamic studies for drug design and screening. *Expert Opin Drug Discov* 2012;7(4):299–314. <https://doi.org/10.1517/17460441.2012.666235>.
- [56] Ross PD, Subramanian S. Thermodynamics of protein association reactions: forces contributing to stability. *Biochemistry* 1981;20(11):3096–102. <https://doi.org/10.1021/bi00514a017>.