Journal of Biomolecular Structure and Dynamics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tbsd20

In silico analysis of a putative dehalogenase from the genome of halophilic bacterium *Halomonas smyrnensis* AAD6T

Habeebat Adekilekun Oyewusi, Kolajo Adedamola Akinyede, Roswanira Abdul Wahab & Fahrul Huyop

To cite this article: Habeebat Adekilekun Oyewusi, Kolajo Adedamola Akinyede, Roswanira Abdul Wahab & Fahrul Huyop (2023) *In silico* analysis of a putative dehalogenase from the genome of halophilic bacterium *Halomonas smyrnensis* AAD6T, Journal of Biomolecular Structure and Dynamics, 41:1, 319-335, DOI: 10.1080/07391102.2021.2006085

To link to this article: https://doi.org/10.1080/07391102.2021.2006085

JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
2023, VOL. 41, NO. 1, 319–335
https://doi.org/10.1080/07391102.2021.2006085

In silico analysis of a putative dehalogenase from the genome of halophilic bacterium *Halomonas smyrnensis* AAD6T

Habeebat Adekilekun Oyewusi[a,b,c], Kolajo Adedamola Akinyede[c,d], Roswanira Abdul Wahab[b,e] and Fahrul Huyop[a,b]

[a]Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, Johor, Malaysia; [b]Enzyme Technology and Green Synthesis Research Group, Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, Johor, Malaysia; [c]Department of Science Technology, Biochemistry unit, The Federal Polytechnic P.M.B, Ado Ekiti, Ekiti State, Nigeria; [d]Department of Medical Bioscience, University of the Western Cape, Bellville, Cape Town, South Africa; [e]Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, Johor, Malaysia

Communicated by Ramaswamy H. Sarma

ABSTRACT

Microbial-assisted removal of natural or synthetic pollutants is the prevailing green, low-cost technology to treat polluted environments. However, the challenge with enzyme-assisted bioremediation is the laborious nature of dehalogenase-producing microorganisms' bioprospecting. This bottleneck could be circumvented by *in-silico analysis of certain microorganisms' whole-genome* sequences to predict their protein functions and enzyme versatility for improved biotechnological applications. Herein, this study performed structural analysis on a dehalogenase (*DehHsAAD6*) from the genome of *Halomonas smyrnensis* AAD6 by molecular docking and molecular dynamic (MD) simulations. Other bioinformatics tools were also employed to identify substrate preference (haloacids and haloacetates) of the *DehHsAAD6*. The *DehHsAAD6* preferentially degraded haloacids and haloacetates (−3.2–4.8 kcal/mol) and which formed three hydrogen bonds with Tyr12, Lys46, and Asp182. MD simulations data revealed the higher stability of *DehHsAAD6*-haloacid- (RMSD 0.22–0.3 nm) and *DehHsAAD6*-haloacetates (RMSF 0.05–0.14 nm) complexes, with the *DehHsAAD6*-L-2CP complex being the most stable. The detail of molecular docking calculations ranked complexes with the lowest binding free energies as: *DehHsAAD6*-L-2CP complex (−4.8 kcal/mol) = *DehHsAAD6*-MCA (−4.8 kcal/mol) < *DehHsAAD6*-TCA (−4.5 kcal/mol) < *DehHsAAD6*-2,3-DCP (−4.1 kcal/mol) < *DehHsAAD6*-D-2CP (−3.9 kcal/mol) < *DehHsAAD6*-2,2-DCP (−3.5 kcal/mol) < *DehHsAAD6*-3CP (−3.2 kcal/mol). In a nutshell, the study findings offer valuable perceptions into the elucidation of possible reaction mechanisms of dehalogenases for extended substrate specificity and higher catalytic activity.

ARTICLE HISTORY
Received 10 June 2021
Accepted 9 November 2021

KEYWORDS
Bioremediation; dehalogenase; *Halomonas smyrnensis* AAD6; marine; MD simulation

1. Introduction

The marine environment is affected by extreme pressure, temperature, salinity, nutrient availability, which invariably drives the microbial inhabitants' molecular adaptability and gives rise to their unique biochemistry characteristics. The same characteristics are now exploited in different fields to solve an array of issues. For example, scientists have successfully expedited molecular changes in certain microbial enzymes to mimic those produced by microbes living in extreme environmental conditions for different applications, viz. to degrade proteins, polysaccharides, and many environmental pollutants (Gupta & Diwan, 2017; Nwodo et al., 2012; Salama et al., 2016). The versatility of extremophilic microorganisms to utilize different substrates is associated with their unique intracellular and extracellular enzymes.

Natural or synthesized chemicals contamination has a profound effect on the marine environment. To date, marine bacteria, algae, sponges, fungi, and insects are documented to produce over 4000 natural organohalides, in addition to several other xenobiotic substances (Gribble, 2003, 2012). Commonplace marine contaminants encompassing halogenated organic compounds include haloacids [(2, 2-dichloropropionic acid (2, 2-DCP), 2, 3-dichloropropionic acid (2,3-DCP), (L-2-chloropropionic acid (L-2-CP), D-2-chloropropionic acid (D-2-CP), 3-chloropropionic acid (3CP)] and haloacetates comprising monochloroacetate (MCA), Dichloroacetate (DCA) trichloroacetate (TCA)]. Fortunately, nature compensates and curbs their accumulation in the environment by having microbes producing specialized enzymes called dehalogenases. This group of biotechnologically important enzymes is capable of detoxifying an array of halogenated compounds by cleaving the carbon–halogen bond.

However, the downside of using such enzymes for the bioremediation of halogen-contaminated marine

CONTACT Habeebat Adekilekun Oyewusi habbyfat@gmail.com Department of Science Technology, Biochemistry Unit, The Federal Polytechnic P.M.B, Ado Ekiti, Ekiti State 5351, Nigeria; Fahrul Huyop fahrul@utm.my Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor, Malaysia

Supplemental data for this article can be accessed online at https://doi.org/10.1080/07391102.2021.2006085

environments is the laborious nature of bioprospecting for suitable microorganisms. The culturing, isolation and the identification of exceptional dehalogenase-producing microorganisms may take several years and warrants a high cost. To date, several dehalogenases have been amplified from cultivable fractions of marine microbes (Edbeib et al., 2017). To overcome the two above issues, novel enzymes derived from a specific microbe are possible using robust bioinformatics analyses on the whole–genome sequences of known microorganisms. The computational investigations include the structure-to-function relationship of dehalogenase, homology modeling, substrate docking, and molecular dynamics simulations to uncover the dynamics of enzyme-substrate complexes and functions, particularly those related to the diversity in binding affinity, specificity, and flexibility. The computationally gathered data could serve as a basis for empirical study and identify specific catalytic and substrate binding residues responsible for the success of the various dehalogenation reactions. The above approach could expedite studies on halophilic bacteria's genetic and metabolic diversity for bioremediation purposes (Oyewusi et al., 2021a).

Genome analysis of *Halomonas smyrnensis* AAD6T isolated from soil samples of the Çamaltı Saltern area (Turkey) was previously reported (Diken et al., 2015; Sogutcu et al., 2012). The bacterium is Gram-negative, non-motile, and is a moderately halophilic aerobe (Poli et al., 2013). Interestingly, the bacterium excretes high levels of levan exopolysaccharide which has applications in nanostructured thin films and peptide-based drug nanocarrier systems (Ates et al., 2013; Aydin et al., 2018, Poli et al., 2009; Sarilmiser et al., 2015; Sezer et al., 2017; Tohme et al., 2018). The bacterial levan exopolysaccharide also found value as a flocculating- and anti-cancer agent (Sam et al., 2011; Sarilmiser & Oner, 2014). Because of the high-value biopolymer the *Halomonas smyrnensis* AAD6T produces, this study conducted an *in-silico* analysis on the genome of *Halomonas smyrnensis* AAD6 which encodes for a dehalogenase, to predict the enzyme's pollutants (halogenated organic compounds) degradation potential. It is worth mentioning here that this study provides the first evidence to understand the function and structure of the predicted dehalogenase, encoded in the genome of *Halomonas smyrnensis* AAD6T. The data obtained from this investigation may prove useful for future empirical analysis.

2. Materials and methods

2.1. Whole-genome sequence retrieval

Halomonas smyrnensis AAD6 genome was retrieved, and the annotated genomes were downloaded from http://www.ncbi.nlm.nih.gov/genome (GenBank accession number: AJKS00000000 (AJKS02000001 to AJKS02000034 for the second version). Keywords dehalogenase searches were first done to identify the gene from the whole genome before predicting the dehalogenase gene organization using the annotated genome location and other dehalogenase accessories.

2.2. Primary and secondary structure prediction

The amino acid sequence of the dehalogenase designated as *DehHsAAD6* (protein_id: WP_016855717.1, locus tag: 72875_73549) from the annotated genome of *Halomonas smyrnensis* AAD6 (GenBank accession number: AJKS02000001 to AJKS02000034) was retrieved from the National Center for Biotechnology Information (NCBI) database. The amino acid sequence was then subjected to an in-depth characterization of the physicochemical properties on the ExPASy server via the Protparam tools (https://web.expasy.org/protparam/) (Gasteiger et al., 2005). The entry name was employed in BRENDA (The Comprehensive Enzyme Information System) (http://www.brenda-enzymes.org/) to identify the Enzyme Commission number for the *DehHsAAD6*. The secondary structure of the enzyme protein was predicted by the Self-Optimized Prediction Method with Alignment (SOPMA) (Geourjon & Deleage, 1995), PSIPRED (Buchan & Jones, 2019), PHD (Rost & Sander, 1994). and GOR4 (Combet et al., 2000) servers. The servers also estimated the percentage of helix and sheet regions on the *DehHsAAD6*.

2.3. Construction of 3D protein model

The *DehHsAAD6* amino acid sequence was submitted to the SWISS-MODEL Automatic Comparative Protein Modelling Server website (https://swissmodel.expasy.org) (Waterhouse et al., 2018) to build the *DehHsAAD6* model template, based on the architecture of a homologous protein of the haloacid dehalogenase of a *Polaromonas* sp. strain JS666 (PDB ID: 3um9). PyMOL (DeLano, 2002) was used to view the generated model. The software packages in VERIFY-3D was used to structurally evaluate the compatibility of sequences to its structure (Lüthy et al., 1992), stereochemical quality assessment on the protein used PROCHECK (Laskowski et al., 1993), and the non-bonded interactions were evaluated by ERRAT (Colovos & Yeates, 1993).

2.4. Model refinement by molecular dynamic simulation and validation

The generated *DehHsAAD6* homology model was further refined by molecular dynamic (MD) simulation, to guarantee that the attained native state was truly at the global minimum (Feig, 2017), and the protein model was found to be free from significant errors compared to its native structure (Park et al., 2018). MD simulation of *DehHsAAD6* was carried out on a parallel version of GROMACS 5.1.2 that employed the Gromos96 53a7 forcefield (https://www.gromcs.org/About_Gromacs). In this study, energy minimization used the steepest descent and conjugate gradient methods at 515 steps and MD simulation was carried out under a constant temperature, and pressure of 300 K and 1 atm, respectively. The production simulation was triplicated, where the *DehHsAAD6* protein was placed in a defined cubic box with a 1.0 nm minimum distance between the solutes and the box's edge. The solutes were solvated using the simple point charge (SPC) water model, and seven sodium ions were

added to the protein simulation box to neutralize the system. The equilibrated structure was simulated for 100 ns with an integration time step of 2 fs. All outputs were plotted as finished Xmgrace graphs to analyze the simulation trajectories (Lindahl et al., 2001), where the snapshots were extracted from the trajectories based on published protocols (Lemkul, 2018). The geometry of the refined 3D model was generated using GROMACS 'gmx trjconv' functions and again validated using PROCHECK, ERRAT, and Verify-3D at the SAVEs server (http://servicesn.mbi.ucla.edu/SAVES/). Finally, the dynamic behavior and structural changes of the *DehHsAAD6* protein were then gauged using the root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) functions.

2.5. Identification of the DehHsAAD6 catalytic residues

The prediction of the *DehHsAAD6* active-site residue used the multiple sequence alignment method constructed by the software Multalin (http://multalin.toulouse.inra.fr/multalin/). Visualization of the primary structure alignment used the ESPript, and PyMOL (DeLano, 2002) detected the catalytic residues. This was done by superimposing the putative active-site residues over the template structure (Oyewusi et al., 2020). The COACH meta-server was also employed to predict the protein-ligand binding site, in which the COFACTOR, FINDSITE, and ConCavity function were combined to generate the final ligand binding site predictions (Yang et al., 2013).

2.6. Preparation of 3D ligand's structure

The 3D structures of the ligands, haloacids: [2,2-dichloropropionic acid (2,2-DCP), 2,3-dichloropropionic acid (2,3-DCP), L-2-chloropropionic acid (L-2-CP), D-2-chloropropionic acid (D-2-CP), 3-chloropropionic acid (3CP)], haloacetate [monochloroacetate (MCA) and, trichloroacetate (TCA)] (Figure 1) were retrieved from PubChem database (3D modelling conformer). The SDF files of the ligands were converted to PDB files (version 2.3.1) using BABEL before hydrogens were added to the 3 D model of *DehHsAAD6* using the Avogadro software. Finally, structures were submitted to the Automated Topology Builder (ATB) and Repository version 3.0 server to optimize the geometries of the molecules.

2.7. Substrate docking

Molecular docking is valuable for elucidating interactions between a protein and ligand. In this study, the Autodock version 4.2.6 was used for molecules preparation (i.e. protein and ligands preparation) and molecular docking between each single ligand and *DehHsAAD6* was conducted using AutoDock Vina, which combines certain advantages of knowledge-based potentials and empirical scoring functions. It extracts empirical information from both the conformational preferences of the protein-ligand complexes and the experimental affinity measurements (Trott & Olson, 2010). Briefly, the investigation started with the blind docking method followed by specific docking into the predicted active site of *DehHsAAD6*. According to a previous study, specific molecular docking is substantially more precise than blind molecular docking. Water molecules were excluded from the DehHsAAD6 protein following the addition of polar hydrogens and non-polar hydrogens before assigning the total Kollman and Gasteigher charges. The ligands underwent the same treatment to ensure the correct adoption of torsions for rotation during docking. The binding region comprising the amino acid residues in the simulated protein was demarcated by the Autogrid tool in Autodock for *DehHsAAD6* was fixed at ±1.000 Å from the 6.283 Å, 1.500 Å and 9.117 Å coordinates with sizes 50, 62, and 46 (x, y, and z positions respectively). The simulation proceeded for 100 runs and 10 million energy evaluations per trail of each flexible ligand. The Vina program's exhaustiveness parameter was set to the default value of 8 Å. A maximum of 10 binding modes were generated, with a maximum energy difference of 3 Kcal/mol between the best and worst binding modes. A redocking experiment was carried out in this study to recreate the native binding poses with the co-crystal reference compound (sulphate ion) into the binding pocket of the template to validate the molecular docking protocol and algorithm. The result of the re-docking experiment of that of docked and co-crystalized reference molecule was superimposed using Pymol and the RMSD value for docked ligand with respect to reference ligands at crystal structure was determined. Then, AutoDock Vina estimated in the subsequent docking analysis where the best result for each substrate was chosen as the largest conformation cluster registering the lowest binding free energy (kcal/mol). Lastly, the 'pdbqt' file for each *DehHsAAD6* substrate complex was converted to the 'pdb' format and visualized by PyMOL to identify the amino acid positions and their corresponding hydrogen bond distances (Houston & Walkinshaw, 2013). LigPlot was then used to visualize both the hydrogen bond and hydrophobic interactions of each *DehHsAAD6* substrate complex.

2.8. MD simulations for enzyme-ligand complex

The parallel version of the GROMACS 5.1.2 software was used for the *DehHsAAD6* protein simulation using the 53a7 Gromos as the forcefield (http://www.gromacs.org/About_Gromacs). Coordinate and topology files of the *DehHsAAD6* protein and ligands were prepared separately to provide the input files for the subsequent MD simulations. For the contents of the system, the ligands were optimized by the Automated force field Topology Builder (ATB; http://compbio.biosci.uq.edu.au/atb/) and Repository (Malde et al., 2011). ATB server has been developed to provide interactive parameters for a wide range of molecules compatible with the GROMOS force Field (Oostenbrink et al., 2004). While the *DehHsAAD6* protein was prepared using the pdb2gmx program in the GROMACS package (Koziara et al., 2014; Malde et al., 2011). The protein and the ligands were merged using the minimized structure of the system before proceeding with the production run to obtain a complex structure. Simulation of the *DehHsAAD6* protein occurred in a

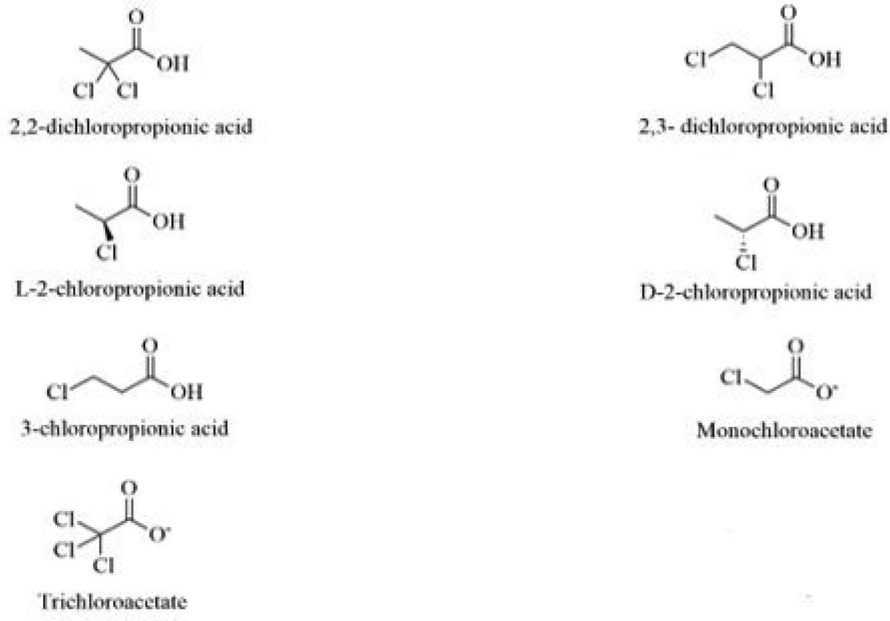

Figure 1. Structures of halogenated organic compounds assessed by this study.

12.571 nm x 11.646 nm x 5.173 nm cubic simulation box. The solutes were solvated with 180000 SPC water molecules for 100 ns under a constant temperature 300 K and pressure of 1.0 atm, and the enzyme was electrically neutralized by adding seven Na^{+} ions.

Before performing the MD simulations, energy minimization was performed with a maximum force per complex not greater than 1000 kJ/mol A. The system was energy minimized using the steepest descent algorithm up to a maximum of 10,000 steps or until the maximum force (Fmax) was no longer greater than 1000 kJ mol^{-1} nm^{-1} (the default threshold). This step was essential to remove steric clashes or inappropriate geometry in the solvated protein-ligand complex system. The *DehHsAAD6* protein was fixed at the center in the system using Periodic boundary conditions (PBC) to make a shift of the protein or protein-ligand complex at the center of the box to prevent edge effect before GROMACS analysis, and Visual Molecular Dynamics was used to visualize the position of the simulation trajectories. The final outputs were then plotted as Xmgrace graphs (Lindahl et al., 2001). RMSD, RMSF, Radius of gyration (Rg), and hydrogen bonds trajectory were used to assess interactions between the *DehHsAAD6* protein and ligands and the corresponding structural changes dynamic behavior of the protein.

3. Results and discussions

3.1. Genetic organization of dehalogenase gene in Halomonas smyrnensis AAD6 genome

The advent of genome sequencing coupled with advances in bioinformatics promises invaluable insights into the genomic analysis. Such technological development has led to new knowledge in gene evolution, ecology, and the design of related therapeutic interventions. In the current study, *H. smyrnensis* AAD6 genome was screened for dehalogenases. The result showed that a single 674 bp-sized haloacid dehalogenase type II (typeII) (72875_73549) was present in the *H. smyrnensis* AAD6 genome and designated as *DehHsAAD6*. The gene was located in the plasmid (contig 10) with a 69.48% G + C content. The detection of type II dehalogenase is somewhat expected because a type II haloacid dehalogenase was common than D-specific dehalogenases (type I). Literature has shown that most naturally occurring halogenated compounds exist in the L-form (Adamu et al., 2016).

Moreover, the genetic organization of *DehHsAAD6* was predicted to further analyse other dehalogenase accessory genes such as regulatory and an uptake gene that could be important for the enzyme expression and regulation. As can be seen in Figure 2, the study identified a putative uptake gene (77166_77845) that encodes for permease protein belonging to ABC Transporter Permease subunit (designated as *DehHsAAD6Pt*). The gene is likely responsible for haloacid uptake. A gene adjacent upstream of *DehHsAAD6* is a gene that encodes DUF3047 domain-containing protein (59855_59983). An adjacent regulator gene was not observed except the GntR family transcriptional regulator (59773_58370) located downstream of *DehHsAAD6*. However, analysis by InterPro revealed that the DUF3047 domain-containing protein is a hypothetical protein with a possible role as the regulatory gene. Herein, we hypothesized that this hypothetical protein might control the dehalogenase of *H. smyrnensis* AAD6 (DehHsAAD6) functions, given their proximity in the genome. Hence, further assessment using site-directed mutagenesis (SDM) is useful to identify the functions of the dehalogenase and its accessory genes.

3.2. Sequence analysis of DehHsAAD6

The *DehHsAAD6* enzyme (EC 3.8.1.2) from *H. smyrnensis* AAD6 genome has a total of 224 amino acid residues, and the BRENDA data revealed it to be a true dehalogenase. The

Figure 2. Gene organization of dehalogenase from the genome of *H. smyrnensis* AAD6. Hypothetical protein *(hyp),* DUF3047of the domain protein *(DUF3047),* GntR family transcriptional regulator, dehalogenase enzyme *(DehHsAAD6),* permease protein *(DehHsAAD6Pt)*, and the phosphotransferase system *(PTS)*.

Table 1. Summary of physicochemical properties of *DehHsAAD6* protein determined by ExPASy's ProtParam.

Details	*DehHsAAD6*
Amino acid residues	224
Molecular weight (Da)	25 066
Theoretical pI	4.93
Negatively charged residues	35
Positively charged residues	24
Molecular formula	$C_{1115}H_{1722}N_{312}O338S5$
Total number of atoms	3492
Aliphatic index (%)	82.90
Instability index (%)	49.49
GRAVY (Grand Average of Hydropathicity)	−0.267

Protparam analysis on the DehHsAAD6 primary structure (Table 1) showed that the dehalogenase's amino acid sequence comprised 3492 atoms with a corresponding molecular formula and molecular weight $C_{1115}H_{1722}N_{312}O_{338}S_5$ and 25066.13 Daltons. Notably, a computed theoretical pI value of below 7 indicates acidic characters (Gasteiger et al., 2005). The *DehHsAAD6's* low theoretical pI (4.93) is the enzyme's unique feature, as it is similar to the hypersaline-adapted dehalogenase from *Bacillus thuringiensis* H2 (*DehHsAAD6*) (pI = 4.58) and *Pseudomonas halophila* HX (DehHX) (pI = 3.89) (Edbeib et al., 2017; Oyewusi et al., 2020). The findings indicated that the acidic nature of the *DehHsAAD6* protein correlated with the high number of acidic residues (Asp and Glu = 35), a well-known adaptive strategy towards increased salinity in brackish water (Oren et al., 2005; Oyewusi et al., 2020).

Likewise, the instability index of *DehHsAAD6* of 49.49 conveyed that the protein is unstable *in-vitro* in aqueous environment. The high aliphatic index of 82.90 thus further affirmed the protein's thermal stability. For clarity, stability- and aliphatic indices of <40 and ≥ 40, respectively, are the collective attributes of a thermally stable protein (Dutta et al., 2018). In this study, the aliphatic index expresses the total volume occupied by aliphatic side chains. A high value (≥ 40) seen for the *DehHsAAD6* protein reflects a more thermally stable protein (Dutta et al., 2018).

Literature has shown that halophilic enzyme proteins tended to have lower hydrophilic amino acids (Paul et al., 2008). This fact thus correlated well with the low percentage of hydrophobic residues and the appreciably high percentage of negative charge in the amino acid sequence of the *DehHsAAD6* (24%). Equally, the data seen here consistently agreed with the high percentage of acidic residues in the *DehHsAAD6* and its negatively valued Grand Average of Hydropathicity (GRAVY = −0.267). GRAVY index indicates solubility of proteins, and a negative value of GRAVY defines it as hydrophilic in nature. Notably, a negative sign preceding the value of GRAVY points to a hydrophilic feature of the protein (Oyewusi et al., 2020). It was clear that the *DehHsAAD6* proteins' higher population of hydrophilic amino acids is an adaptation to the marine environment, wherein the surplus acidic residues promote better solvation of the dehalogenase protein. The negatively/positively charged residues could ionically attract the surrounding water molecules to bind to the *DehHsAAD6* protein, averting dehydration and loss of active conformation. The same distinctive features were also reported for other halophilic proteins (Hutcheon et al., 2005; Lanyi, 1974; Oyewusi et al., 2020).

The ability of halophilic enzymes to function in a highly saline condition is their ability to maintain a balance between protein flexibility for catalytic activity and adequate rigidity to avert excessive unfolding (Zorgani et al., 2014). Varied approaches for predicting protein structure yield different results, such as the presence or absence of α-helices, extended strands, β-turns, and random coils, as well as the percentage of residues in these regions. It has been widely accepted that amino acid sequences carry all the information required to construct three-dimensional structures (Mustafa et al., 2017). Consequently, protein structures (2D and 3D) can theoretically be predicted based on their animo acids sequences. In this study, different prediction tools such as SOPMA, PSIPRED, PHD and GOR4 secondary structure prediction servers were used. Table 2 lists the results obtained using these tools. Surprisingly, for secondary structure elements, all the tools yielded comparative results. SOPMA revealed that *DehHsAAD6* has all four elements (α-helix, extended strands, β-turns and random coils), although PSIPRED, PHD and GOR4 secondary structure prediction server did not predict β-turns. Similarly, SOPMA, PSIPRED, PHD and GOR4 all yielded essentially identical results, with 58.93%, 54.91%, 58.48% and 57.59% of residues belonging to the α-helix, respectively. All tools showed that *DehHAAD6* has 11.62%, 13.39%, 12.05% and 10.27% residues belonged to extended strand, while 26.78%, 31.69%, 29.46%, and 32.14% residues belonged to SOPMA, PSIPRED, PHD and GOR4 respectively. whereas 2.86% residues belonged to β-turn according to SOPMA (Table 2).

According to Geourjon and Deleage (1995), a high population of α-helices and coiled region indicates better conservation and stability of the protein. Thus, this outcome indicated a sufficiently flexible *DehHsAAD6* protein, capable of averting functionality loss and maintaining the enzyme's conformational structure under extreme salt stress. Most importantly, the *DehHsAAD6* protein could remain hydrated in a highly saline condition because of its higher ability to attract ambient water molecules. This observation agreed with another hydrophilic dehalogenase protein, DehH2, that strongly binds water molecules under a hypersaline environment (Oyewusi et al., 2020a, 2021b). Therefore, adequate flexibility and hydration are relevant factors in ensuring

Table 2. Secondary structure of *DehHsAAD6* protein predicted by different savers.

	SOPMA		PSIPRED		PHD		GOR4	
Tools properties	Value	%	Value	%	Value	%	Value	%
α-helix	132	58.93	124	54.91	131	58.48	129	57.59
Extended strand	26	11.62	30	13.39	27	12.05	23	10.27
β-turn	6	2.86	NP	NP	NP	NP	NP	NP
Random coil	60	26.78	65	29.02	66	29.46	72	32.14

NP: Not Predicted.

effective catalysis and activity of enzymes in highly saline environments (Zorgani et al., 2014). As a result, while predicting protein secondary structure, one should not rely on a single tool but rather attempt as many as possible before manually selecting the best prediction. To gain a clearer picture of the presence and location of helices, strands, turns and coils, the tertiary structure of proteins should also be anticipated.

3.3. Homology modeling, model refinement, and structure validation

The amino acid sequence of dehalogenase from *H. smyrnensis* AAD6 genome was subjected to homology modeling using the SWISS-MODEL web server (Waterhouse et al., 2018). The three-dimensional (3D) model of *DehHsAAD6* was constructed from the crystal structure of haloacid dehalogenase of a *Polaromonas* sp. strain JS666 (PDB ID: 3um9) (Figure 3) that shares a 37.79% sequence similarity with the *DehHsAAD6*, based on our BLAST search. The 37.79% sequence similarity seen here affirmed (>30%) the suitability of Bpro0530 haloacid dehalogenase of a *Polaromonas* sp. strain JS666 as a template for the homology modelling study. The root mean square deviation (RMSD) between the *DehHsAAD6* template and predicted model was 2.129 Å, which was within the permitted range for comparative structural modelling. In practice, comparative structural modelling is employed for proteins with homologous templates, and most predicted structures have an RMSD of less than 3 Å compared to experimental structure, achieving the accuracy of medium-resolution NMR or low-resolution X-ray structures in some cases (Zhang, 2009). RMSD is no longer a valid metric of modelling quality for poor precision models (say, RMSD > 3 Å) since a small misorientation of tails or loops might result in a large overall RMSD even if the core region of the model is right. The Ramachandran plot for the *DehHsAAD6* model before minimization is shown in Table 2. The model was energy minimized using Gromos97 to remove local strains within the 3D *DehHsAAD6* model, as minor errors from atomic overlap or bad Van der Waals interaction could exist in the original protein structure (Oyewusi et al., 2020). Energy minimization of enzymes helped remove any local strain after the addition of hydrogens. PROCHECK, ERRAT, and Verify-3D were used as indicators to verify and validate this data after energy minimization (Table 2).

Then, energy minimization was performed to observe the atomic-level and the dynamical properties of the *DehHsAAD6*. The conformational sampling during

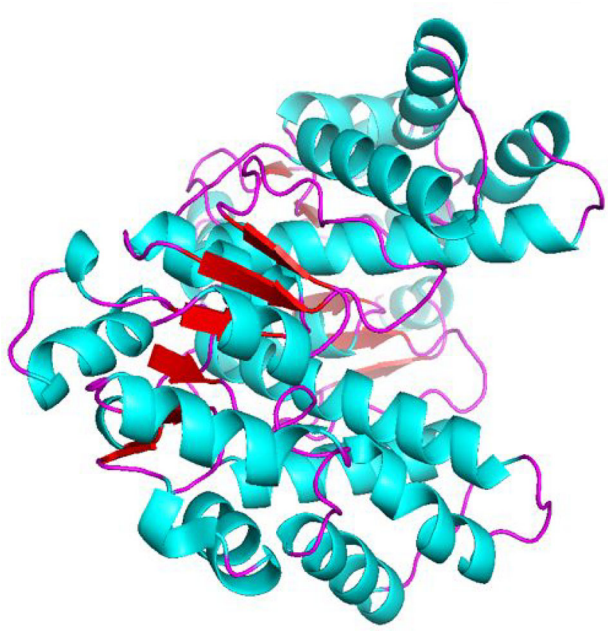

Figure 3. The SWISS-MODEL generated 3 D structure of *DehHsAAD6* protein *(Polaromonas* sp. strain JS666 (PDB ID: 3um9). Blue, red and purple represent the α-helices, β-sheet and loop of the enzyme, respectively.

simulation could establish the protein's most stable state in relevance to its function (Oyewusi et al., 2021). The RMSD and RMSF values of the *DehHsAAD6* model are depicted in Figures 4a and 4b. Notably, a low RMSD value (average RMSD ∼ 0.25 nm) implies the high stability of the *DehHsAAD6* structure, while less stable protein structures have higher values (Anuar et al., 2020; Oyewusi et al., 2020). The lower RMSD and RMSF values seen here were proof of the enzyme's ability to remain adequately flexible and stable in a highly saline environment, indicating a good model for *DehHsAAD6*.

Next, the refined output structure from the MD simulation depicted a good protein folding comparable to a native protein. The refinement step had effectively reduced side chain clashes and steric hindrances. The same treatment also minimized the protein's potential energy by the steepest gradient method, yielding the most stable 3D structure of the *DehHsAAD6* protein. Data of Verify −3D, and ERRAT revealed that 96.0% of residues occupied the most favorable region, with 4.0% in the additional allowed region, and none are in the disallowed region (Table 3). The good quality of the *DehHsAAD6* homology model was verified by >90% of the residues being cited in the most favorable region (Figure 5), thus corroborating earlier studies (Anuar et al., 2020; Oyewusi et al., 2020). The ERRAT score for the non-energy minimized *DehHsAAD6* model exceeded the rejection limit of 95% (96.279%) but was appreciably improved to 98.144% after energy minimization (Table 3). It should be noted that an ERRAT score of >50% is the accepted range for a good protein model (Rosdi et al., 2018). Also, the *DehHsAAD6* homology model attained satisfactory side-chain environments, as shown by the high Verify-3D scores of before and after energy minimization (>0) (Table 3). A deemed satisfactory protein model has a Verify-3D score of >80% (Rosdi et al., 2018).

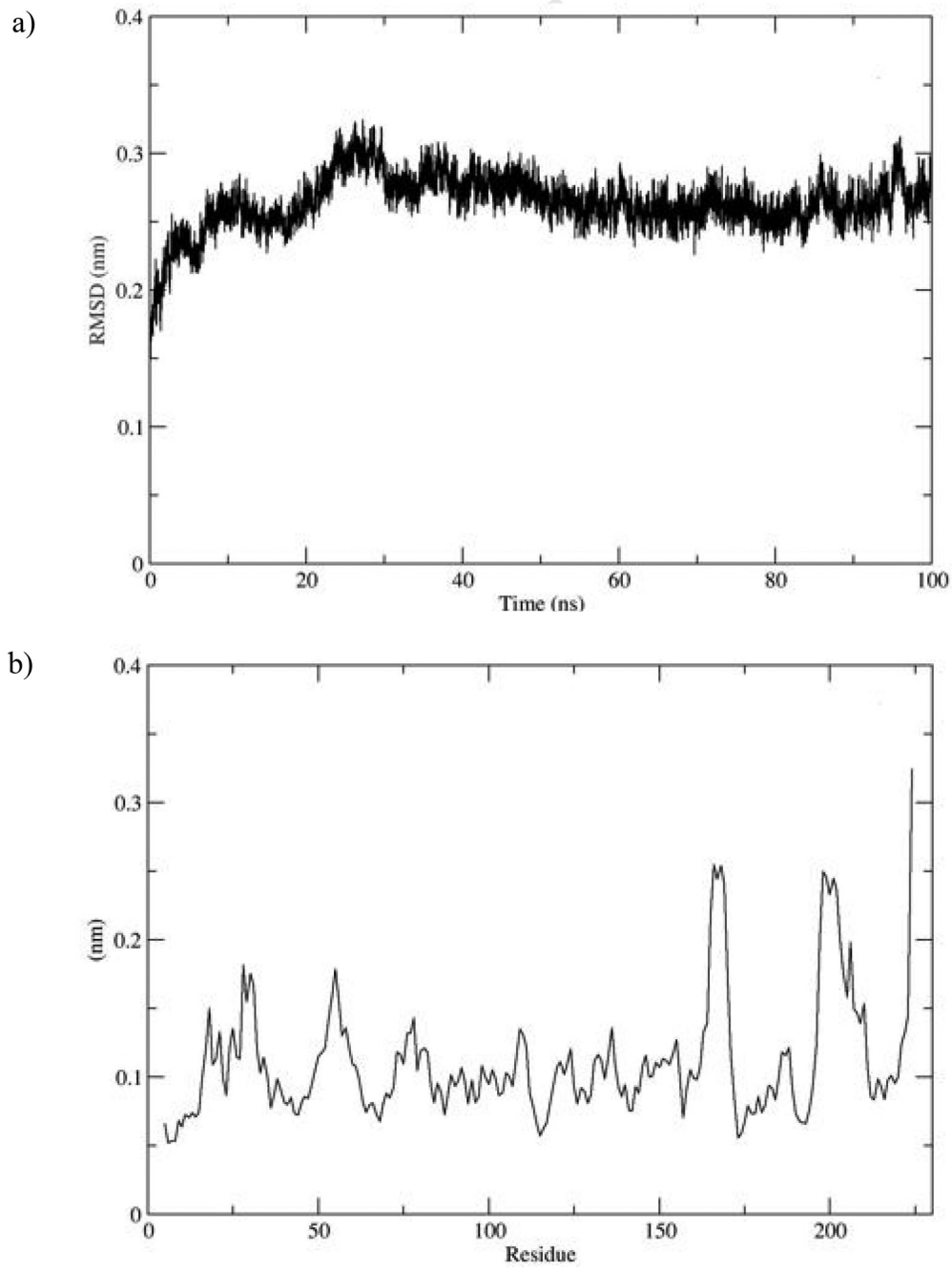

Figure 4. MD simulation of *DehHsAAD6*. (a) average RMSD and (b) average RMSF of the backbone conformation are shown as a function of simulation time (100 ns) at 300 K.

Table 3. Results of structural validation of *DehHsAAD6* protein model with the SAVEs Server (before and after).

			Validation of analysis (%)		
Validation	Parameter	Scheme	Before	After	Ranges scores (%)
ERRAT		Over quality factor	96.28	98.14	>50
Verify-3D		Amino acid compatibility	99.33	100	>80
Procheck	Most favored region (A,B,L)	Stereochemical quality	92.8	96.0	>90
	Additional favored (a,b,l,p)		6.8	4.0	
	Generously allowed regions (□a, □b, □l, □p)		0.5	0	

3.4. Active-site residues prediction

This investigation used COACH, a meta-server approach to predict protein-ligand binding site. This server generates complementary ligand binding site predictions using two comparative methods, TM-SITE and S-SITE and also combined with results from other methods such as COFACTOR, FINDSITE and ConCavity to generate final ligand binding site predictions (Yang et al., 2013). Additionally, blind docking, protein-ligand binding site prediction and multiple sequence alignment to predict the catalytic triad of *DehHsAAD6*.

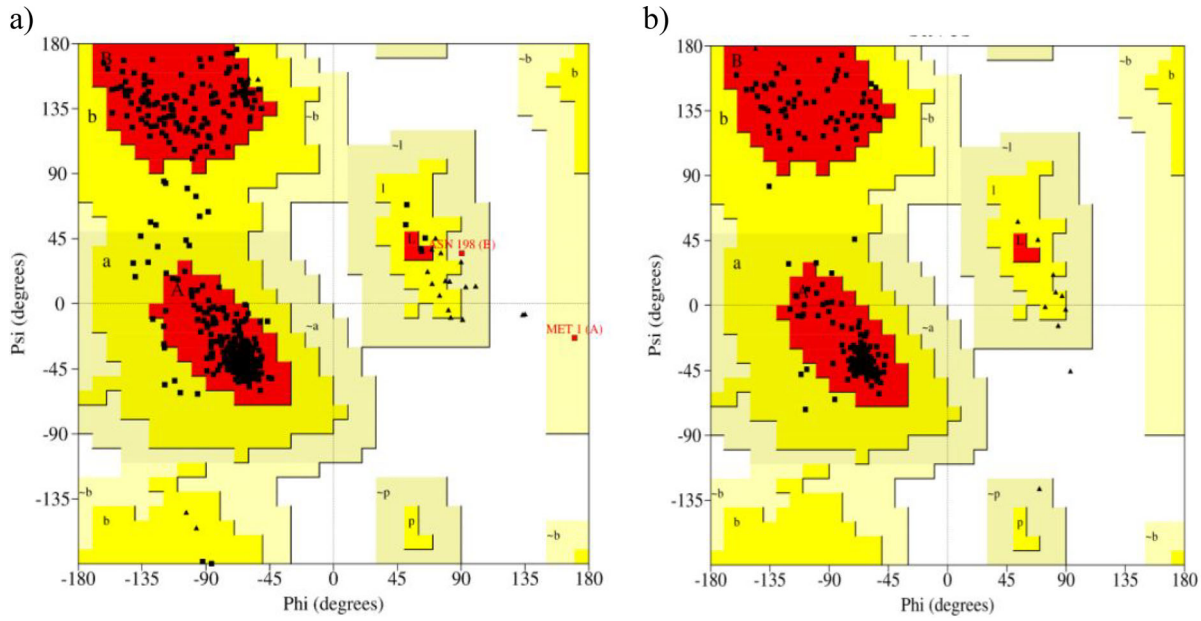

Figure 5. Ramachandran plot of a polypeptide backbone torsion angles psi (w) against phi (u) of amino acid present in the homology modelled structures of DehHsAAD6 generated by SWISS-MODEL (a) before and (b) after minimization. The most favored region are colored in red. The additional allowed regions [a, b, l, p] are colored in yellow. The generously allowed regions [~a, ~b, ~l, ~p] are colored in pale yellow. All non-glycine and proline residues are illustrated as filled black squares, and glycine (non-end) are indicated as filled black triangles. Disallowed residues are colored in white.

Specific docking at the predicted catalytic triad by blind docking was done to double-check the location of the active site residues. This step is to ensure that the substrates had bonded correctly to the *DehHsAAD6* active site, to identify the substrate-enzyme binding mechanism or for structural determination (Awasthi et al., 2018).

Blind docking of substrates haloalkanoic acids (2,2-DCP, 2,3-DCP, L-2CP, D-2CP, and 3-CP) and haloacetate (MCA and TCA) was performed over the whole surface of the *DehHsAAD6* to locate the possible active sites, as the active site pocket of the enzyme is unknown (Yan et al., 2016). Results revealed that all substrates except chlorpyrifos formed hydrogen bonds with active site residues, Tyr12, Lys46, and Asp182, of the *DehHsAAD6*. This data corresponded with the COACH data that showed a similar set of repeating numbers for Tyr12, Lys46, and Asp182. Pertinently, the findings corroborated the earlier blind docking result, suggesting that Tyr12, Lys46, and Asp182 are the possible active site residues of *DehHsAAD6* (Figure 6).

3.5. Molecular docking analysis

The catalytic triad plays a relevant role in the dehalogenation reaction in the bioremediation or biodegradation of halogenated organic compounds (Yu et al., 2020). Thus, the structural basis of an enzyme and its corresponding active site residues must be well-understood (Lemmon & Meiler, 2013). In this study, AutoDock Vina and AutoDock 4.2.6 software using the AutoGrid tools was employed for the docking analysis on the *DehHsAAD6*. The docking procedure and algorithm were validated utilising a re-docking experiment between the original co-crystal reference molecule (sulphate ion) and the target protein (DehHsAAD6 model). Based on the re-docking analysis the reference molecule exhibited an RMSD value of 2.129 Å, between the docked and co-crystal conformation. According to a recent study, docking solutions with values of 2.0, 2.0–3.0, and $\geq$3 Å are regarded as good, acceptable, and unacceptable respectively (Ramírez & Caballero, 2018). As a result, a little divergence in RMSD indicated that the molecular docking protocol, parameters and algorithm utilised in this experiment were reliable enough to mimic biological conformation of the molecules (Gurung et al., 2020). The results of the re-docking analysis showed that the docked and co-crystalized reference molecules were partially superimposed (S1); thus, the docking method was deemed adequate for further docking analysis. Eight different ligands used for the molecular docking study against the *DehHsAAD6* of *H. smyrnensis* AAD6 were 2,2-DCP, 2,3-DCP, L-2CP, D-2CP and 3-CP) and haloacetate (MCA and TCA). Specific docking of protein-ligand complex helps to anticipate the preferred orientation when bound to each other to form a stable enzyme-substrate complex (Oyewusi et al., 2020); hence the lowest energy (kcal/mol) illustrates a strong enzyme-substrate affinity (Mishra et al., 2019).

Therefore, AutoDock Vina provides an accurate and high-performance docking score with various orientations and conformations for a given ligand at a binding site, based on the Broyden–Fletcher–Goldfarb–Shanno algorithm long side with empirical and knowledge-based scoring functions (Elmezayen et al., 2020). Results showed that the three residues (Tyr12, Lys46, and Asp182) previously predicted to be the catalytic triad of *DehHsAAD6* in the blind docking study were found to form hydrogen bonds to the oxygen atom of C = O of the substrates. Table 4 enlists the docking scores of the *DehHsAAD6*-ligand complexes with the corresponding hydrogen bond distances. Figure 7 depicts the LigPlot showing the formed hydrogen bonds between the *DehHsAAD6*-ligands complexes.

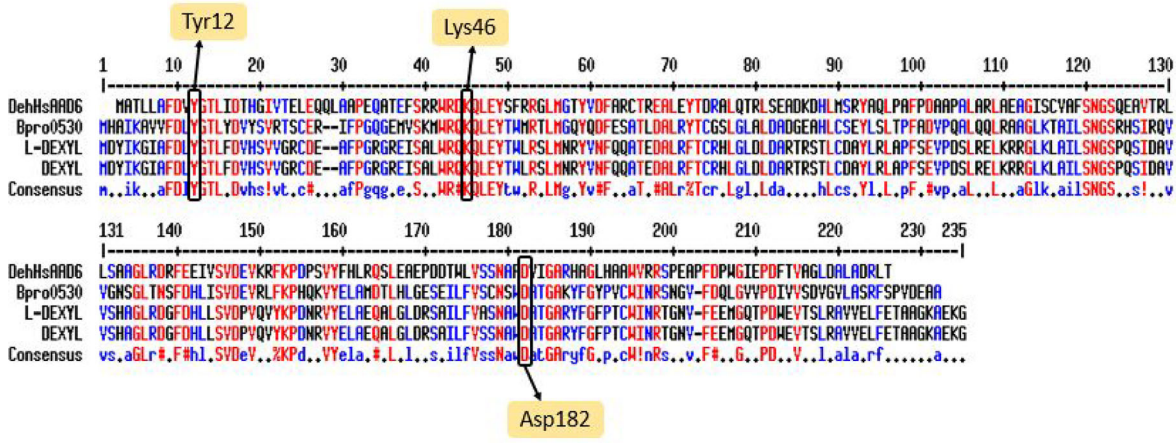

Figure 6. Multiple sequence alignment demonstrating the catalytic triad of *DehHsAAD6* that consisted of Tyr12, Lys46, and Asp182.

Table 4. Summary of AutoDock Vina scores from docking analysis of the *DehHsAAD6*-ligand complexes.

Protein-ligand complexes	Score (kcal/mol)	Interacting residues	Distance (Å)
DehHsAAD6-MCA	−4.8	Lys46	3.00 Å
		Tyr12	3.06 Å
DehHsAAD6-TCA	−4.5	Asp182	3.02 Å
		Lys46	3.16 Å
DehHsAAD6-3CP	−3.2	Asp182	2.92 Å
		Tyr12	2.86 Å
DehHsAAD6-2,3-DCP	−4.1	Lys46	3.34 Å
		Tyr12	2.70 Å
DehHsAAD6-D-2CP	−3.9	Lys46	3.18 Å
		Tyr12	3.31 Å
DehHsAAD6-2,2-DCP	−3.5	Asp182	3.07 Å
		Lys46	3.11 Å
DehHsAAD6-L-2CP	−4.8	Asp182	3.01 Å
		Lys46	2.96 Å
		Tyr12	2.83 Å

The lowest binding free energy reflects a stronger *DehHsAAD6*-ligand complex interaction, wherein the *DehHsAAD6*-L-2CP and *DehHsAAD6*-MCA complexes yielded the lowest binding free energy of −4.8 kcal/mol. The findings were seen to correlate well with the three hydrogen bonds formed by Tyr12 (2.83 Å), Lys46 (2.96 Å), and Asp182 (3.01 Å) in the *DehHsAAD6*-L-2CP complex. However, the *DehHsAAD6*-MCA complex formed only two hydrogen bonds with Tyr12 (3.06 Å) and Lys46 (3.00 Å) (Table 4). The *DehHsAAD6*-TCA and *DehHsAAD6*-2,3-DCP complexes also yielded comparable binding free energies of −4.5 kcal/mol and −4.1 kcal/mol, respectively. The former's hydrogen bond distances were slightly longer, formed with residues Lys46 (3.16 Å) and Asp182 (3.02 Å). Likewise, the *DehHsAAD6*-2,3-DCP complex formed hydrogen bonds with Lys46 (3.34 Å) and Tyr12 (2.70 Å). Higher binding free energies were estimated for the *DehHsAAD6*-D-2CP- (−3.9 kcal/mol), *DehHsAAD6*-2,2-DCP- (−3.5 kcal/mol) and *DehHsAAD6*-3CP (−3.2 kcal/mol) complexes. All three complexes formed two hydrogen bonds with the predicted three catalytic residues (Tyr12, Lys46, Asp182) (Table 4). Similarly, the complexes have very close hydrogen bond distances and form distinctively weaker bonds (Table 4).

Pertinently, the formed hydrogen bonds in the *DehHsAAD6*-ligand complexes were within the acceptable cut-off distances of intermolecular hydrogen bonds (<3.5 Å) (Fu et al., 2018). A longer hydrogen bond distance (>3.5 Å), on the one hand, conveys a lower affinity of an enzyme towards a substrate. Binding affinity is the strength of the binding interaction between the enzyme and its ligand. It is used to evaluate and rank order strengths of biomolecular interactions (Oyewusi et al., 2020; 2021b). In this sense, the lower the binding affinity, the more weakly the enzyme and ligand are attracted to and bind each other. Therefore, a more negative binding energy value shows a greater enzyme-ligand binding affinity (e.g. L-2CP or MCA) for the enzyme. In this study, the binding affinity was evaluated by molecular docking. The organohalide-binding affinity with the *DehHsAAD6* was −4.8 kcal/mol, −4.8 kcal/mol, −4.5 kcal/mol, −4.1 kcal/mol, −3.9 kcal/mol, −3.5 kcal/mol, −3.2 kcal/mol for L-2CP, MCA, TCA, 2,3- DCP, D-2CP, 2,2-DCP, and 3CP, respectively (Table 4). L-2CP, MCA, and 3CP showed the highest and lowest binding affinity values when it was bound to *DehHsAAD6*, respectively. This shows that L-2CP and MCA have a great binding affinity for *DehHsAAD6*, whereas 3CP is weakly attracted to *DehHsAAD6*. These results rank the strength of the binding interaction and degradation potential of these organohalide pollutants to the enzyme as follows: for L-2CP = MCA > TCA > 2,3- DCP > D-2CP > 2,2-DCP > 3CP. Our findings thus confirmed the efficacy and specificity of the *DehHsAAD6* to degrade haloacids and haloacetate. In contrast, the molecular docking results imply the substrates L-2CP and MCA is the preferred substrate of *DehHsAAD6,* but the binding affinity observed for all the ligands were observably lower compared to other studies in the field of medicinal chemistry (about −7 kcal/mol). A certain degree of discrepancy was somewhat expected for the L-2CP and MCA ligands, as binding affinity tends to differ from one ligand to another. Moreover, in this area of study i.e. bioremediation there is no benchmark for assigning a certain range of binding affinity in terms in relation to the preference/strength of binding between a protein and ligand. Likewise, in the future the study might also focus on ligand optimization to enhance the binding of the ligands with *DehHsAAD6*. Also, the findings in this work corroborated earlier studies showing a dehalogenase that interacted similarly with the two substrates and was capable of degrading them (Oyewusi et al., 2020; 2021b).

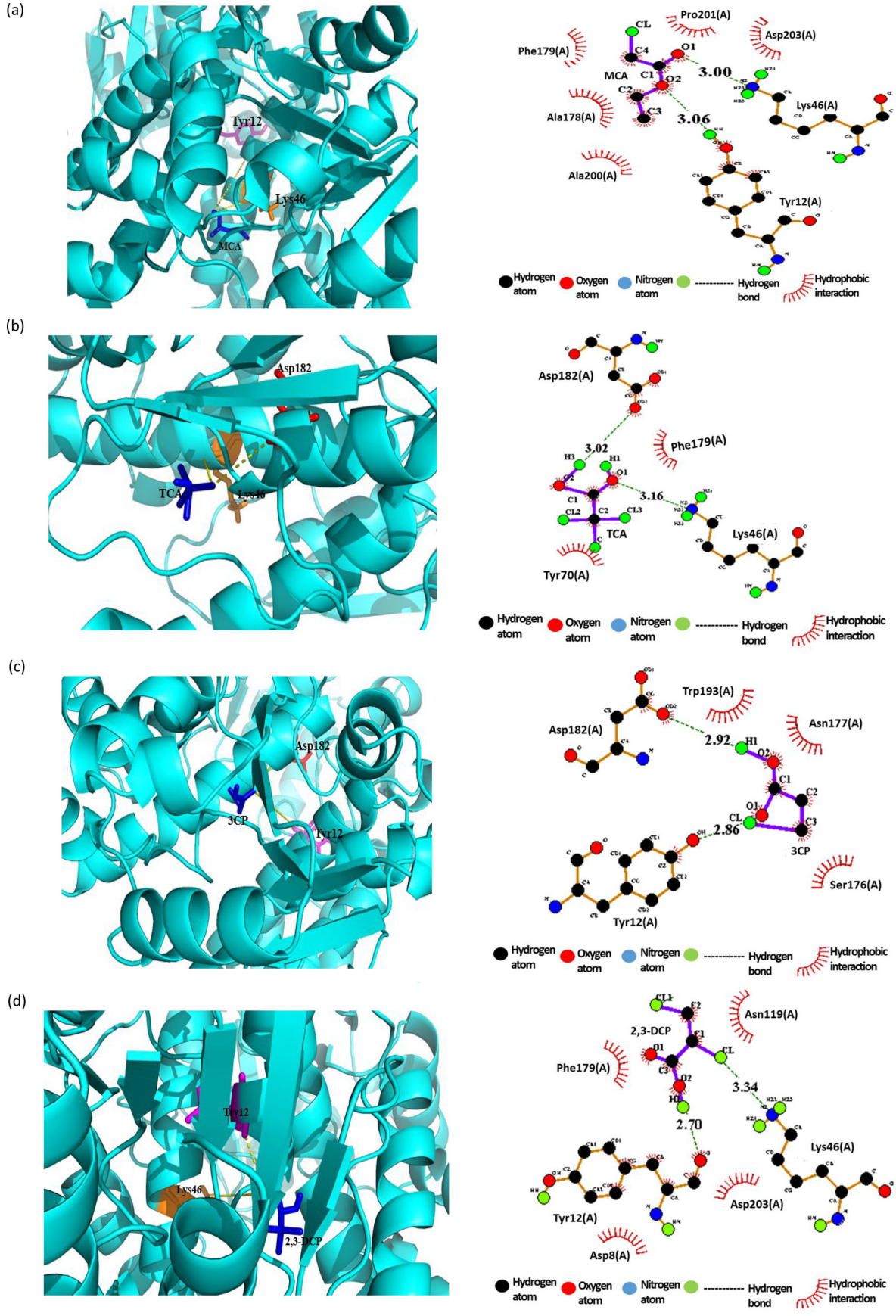

Figure 7. LigPlots for the *DehHsAAD6*-ligand interactions, which show hydrogen bonds with their equivalent distances and non-ligand residues involved in the hydrophobic interactions. (a) MCA (b) TCA (c) 3-CP (d) 2,3-DCP (e) D-2CP (f) 2,2-DCP and (g) L-2CP.

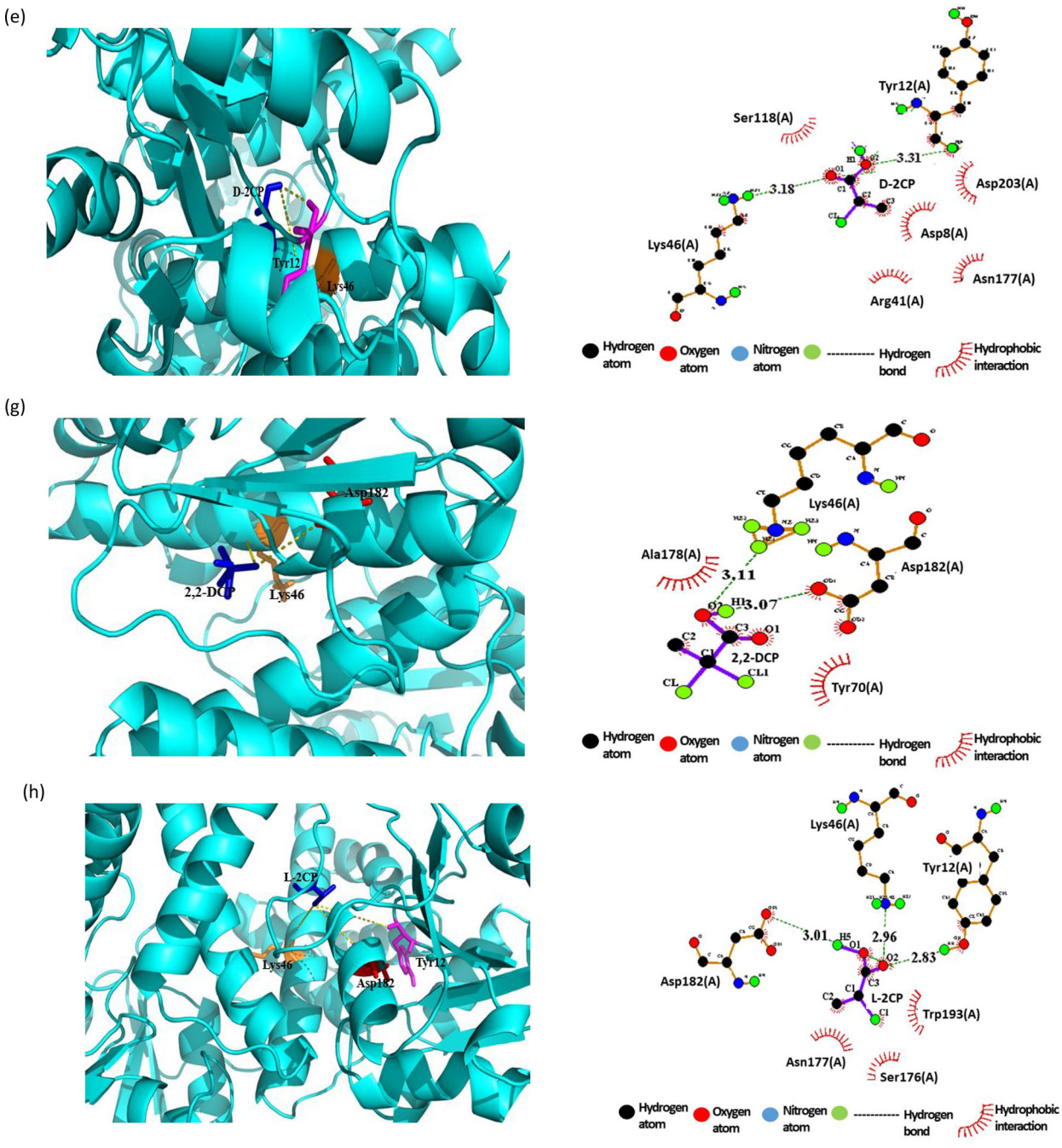

Figure 7. (Continued).

3.6. MD simulation of enzyme-ligand complexes

MD simulation allows an insight into the behavior and structural flexibility of protein in a system when bonded to different ligands or substrates (Lee et al., 2015). In this study, structural changes, stability, and flexibility of the *DehHsAAD6* protein that interacted with the ligands/substrates (haloalkanoic acid 2,2-DCP, 2,3-DCP, L-2CP, D-2CP, and 3-CP) and haloacetate (MCA and TCA) were observed. Their interactions were gauged by calculating the average values of Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), radius of gyration, and hydrogen bond distance in the triplicated 100 ns MD simulations.

3.6.1. Root mean square deviation (RMSD)

Notably, a low RMSD value (RMSD $\sim$ 0.2 – 0.3 nm) indicates high stability of the complex structures and vice versa (Anuar et al., 2020; Oyewusi et al., 2020). The RMSD values for all complexes were seen to fluctuate weakly throughout the 100 ns of MD simulation, indicating the adequacy of simulation time to equilibrate the *DehHsAAD6*-ligand complexes (Hamid et al., 2015a; 2015b). As shown in Figure 5, MD trajectory of the *DehHsAAD6*-MCA complex stabilized rapidly at 6 ns (average RMSD = 0.20 nm) onwards, compared to the *DehHsAAD6*-L-2CP complex. The latter required 20 ns (average RMSD = 0.22 nm) but exhibited a major

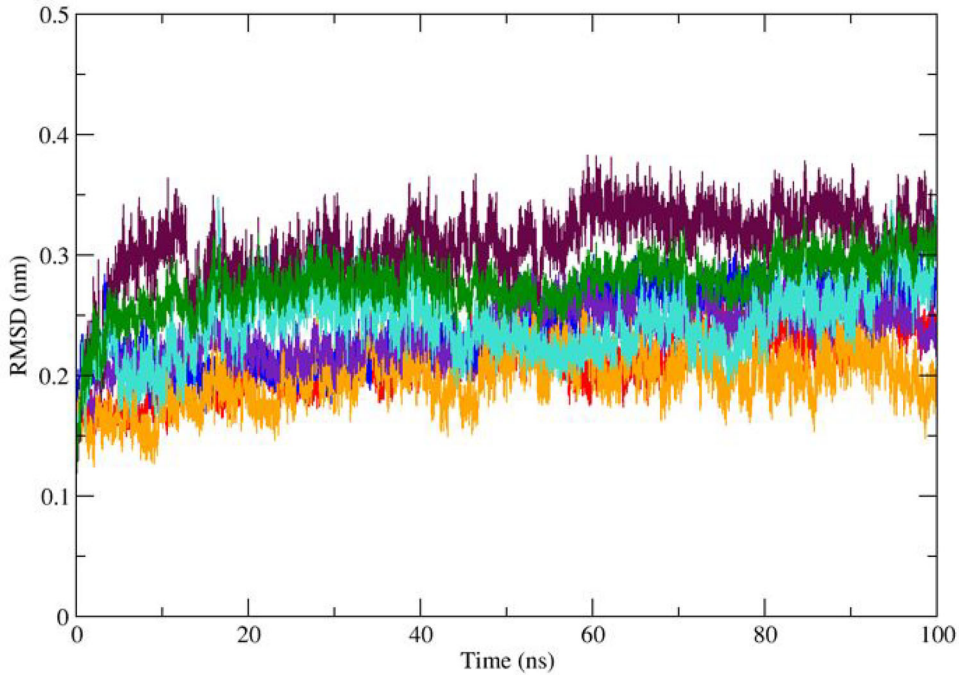

Figure 8. MD simulation of *DehHsAAD6* against the various halogenated organic compounds and chlorpyrifos. Average RMSD of the backbone conformation is shown as simulation time (100 ns) at 300 K and the overlayed RMSD plots of L-2CP in maroon, MCA in purple, 2,2-DCP in orange, D-2CP in green, 3CP in Red, 2,3-DCP in turquoise, and TCA in blue.

fluctuation at 63 ns (average RMSD = 0.30 nm). The *DehHsAAD6*-2,3-DCP complex achieved reasonable equilibrium from 9 ns (average RMSD = 0.20 nm) onwards with a major fluctuation at 43 ns (average RMSD = 0.25 nm) (Figure 8). Equilibration of complexes *DehHsAAD6*-D-2CP and *DehHsAAD6*-TCA (Figure 8) occurred a little later, in which each scored an average RMSD of 0.21 nm (15 ns) and 0.20 nm (28 ns), respectively. The MD trajectory of the *DehHsAAD6*-2,2-DCP complex was somewhat unusual with reasonable regular fluctuations at 27 ns, 56 ns (average RMSD = 0.25 nm), and 93 ns (0.29 nm) before stabilizing from 95 ns onwards (average RMSD = 0.2 nm). Conversely, the study found that the *DehHsAAD6*-3CP complex fluctuated until 60 ns production. The complex recorded the highest average RMSD value of 0.29 nm; thus it was the least stable enzyme-ligand complex suggesting dehalogenase enzymes that can degrade alpha-chloroalkanoate (2,2DCP; 2,3DCP, D- or L-2CP) were common compared to beta-substituted haloalkanoate (3CP).

The L-2CP and MCA were found to be the preferred substrate of *DehHsAAD6* followed by, 2,3-DCP, D-2CP, TCA, 2,2-DCP, and lastly, 3CP. In short, the MD data revealed that nearly all ligands were within the favorable range for interaction with *DehHsAAD6* (0.3 nm), which stabilized toward to end of production simulation. Hence, the substrates could form adequately strong bonds based on their hydrogen bond distances (<0.3 nm) to *DehHsAAD6*, as similarly described by similar studies (Hamid et al., 2015a, 2015b). The findings also validated the blind docking and molecular docking results which showed L-2CP and MCA (−4.8 kcal/mol) having the highest affinity to *DehHsAAD6* compared to TCA (−4.5 kcal/mol), 2,3-DCP (−4.1 kcal/mol), and D-2CP (−3.9 kcal/mol), 2,2-DCP (−3.5 kcal/mol) and 3CP (−3.2 kcal/mol). The general convergence from the initial structure for the *DehHsAAD6*-ligand configurations for substrates L-2CP, MCA, 2,3-DCP, D-2CP, and TCA in the first 56 ns, less so for 2,2-DCP and 3CP, under the constant temperature (300 K) and pressure (1 atm) signified their satisfactory stabilities.

3.6.2. Root mean square fluctuation (RMSF)

The RMSF value reveals the residue-specific flexibility by calculating the individual residue flexibility or the extent of any residue movement (fluctuates) along a principal axis during simulation. A high RMSF value describes a high degree of movement, while a low RMSF value signifies a stable and more rigid structure (Anuar et al., 2020; Junaid et al., 2014; Kumar et al., 2014). RMSF of >0.05 nm (0.5 Å) is the threshold value where a significant change in residue-specific flexibility occurs (Dong et al., 2018; Kovacic et al., 2016; Zhu et al., 2019).

As roughly seen in the RMSF plots (Figure 9), the estimated average RMSF values were low for complexes *DehHsAAD6*-L-2CP, *DehHsAAD6*-MCA (0.05–0.10 nm), followed by *DehHsAAD6*-TCA, *DehHsAAD6*-2,2-DCP, *DehHsAAD6*-D-2CP, *DehHsAAD6*-3CP complexes (0.05–0.12 nm) and the *DehHsAAD6*-2,3DCP complex (0.05–0.14 nm) (Figure 9). The good structural stability of the *DehHsAAD6*-L-2CP, *DehHsAAD6*-MCA complexes were well exemplified in their low RMSF values (0.05–0.10 nm), as similarly described by other studies (Hamid et al., 2015a, 2015b). The findings verified that the *DehHsAAD6* bonded tightly to the ligands in relation to their average positions. However, comparisons of RMSF values in all tested complexes did show some suppression of conformational dynamics in certain residues, notably residues 1–5. The outcome implied the absence of

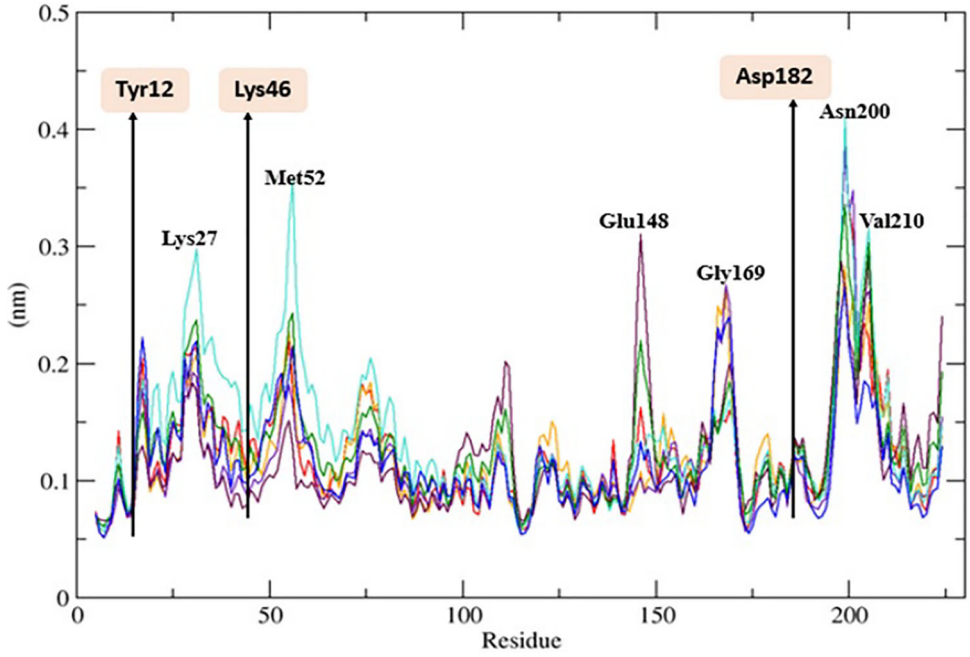

Figure 9. Average RMSF plots of Cα atoms for *DehHsAAD6* showing all RMSD plotted together throughout the 100 ns production simulations. The L-2CP is depicted in maroon, MCA in purple, 2,2-DCP in orange, D-2CP in green, 3CP in red, 2,3-DCP in turquoise, and TCA in blue. The labeled residues in pink box (Tyr12, Lys46 & Asp182) denote the catalytic triad while others (Lys27, Met52, Glu148, Gly169, Asn200 & Val210) are amino acids with the high fluctuation.

interactions between the enzyme residues with the ligands, as previously described by Cohen et al. (2019).

In general, the *DehHsAAD6*-ligands complexes showed that lysine, methionine, glutamate, glycine, asparagine, and valine were at positions 27, 52, 148, 169, 200, and 210, respectively, were the fluctuating residues (Figure 9). Complex *DehHsAAD6*-2,3-DCP (Figure 6) exhibited markedly fluctuating residues, namely the lysine-27 (0.3 nm) and methionine-52 (0.36 nm) and asparagine-200 (0.41 nm), corresponding to 0.200 nm (88 ns), 0.180 nm (60 ns) and 0.195 nm (78 ns), respectively. Similarly, glutamate, glycine, asparagine and valine in complexes *DehHsAAD6*-L-2CP, *DehHsAAD6*-TCA, *DehHsAAD6*-3CP, *DehHsAAD6*-2,2-DCP and *DehHsAAD6*-D-2CP fluctuated at the corresponding positions 148 (0.32 nm), 169 (0.28 nm), 200 (0.41 nm) and 210 (0.32 nm) (Figure 9). The higher *DehHsAAD6*-complexes RMSF values seen here ascribed the ability of the enzyme to bind to all the ligands firmly (Figure 9). Conversely, large RMSF fluctuations of certain residues in the *DehHsAAD6*-complexes (Figure 9) were related to their location in the highly flexible α-helix region compared to their average positions (Nemaysh & Luthra, 2017).

Most importantly, the low RMSF values of the catalytic triad (Tyr12-Lys46-Asp182) in the DehHsAAD6 point to their overall tight bonding with the ligands; behavior consistent with their location in the protein core (Figure 9). Quite the reverse, residues located on exposed loops tend to record higher RMSF values (Fuentes et al., 2018). The lower RMSF values in all *DehHsAAD6*-ligand complexes also validated their higher selectivity for halogenated organic substrates. A comparable outcome was also predicted by another related study (Oyewusi et al., 2020, 2021b). However, the slight difference in binding free energy was insufficient to reliably establish the substrate preference of *DehHsAAD6*, as previously indicated by a similar study by Anbarasu and Jayanthi (2018). Another corresponding parameter must also be measured.

3.6.3. Radius of gyration (Rg)

The Rg value of the *DehHsAAD6*-substrate complexes was measured to observe changes in compaction and the overall protein dimensions during the MD simulation. A reasonably constant Rg value describes a stable folded structure, while an unfolded structure will cause the Rg value to fluctuate throughout the simulation (Liao et al., 2014).

In this study, low average Rg values for *DehHsAAD6*-TCA, *DehHsAAD6*-L-2CP, and *DehHsAAD6*-MCA, (Figure 10) complexes (1.68–1.78 nm) were observed, which indicated minor fluctuations. All the three complexes mildly fluctuated (average Rg value = 1.72 nm), which corroborated their stability and good protein folding. The *DehHsAAD6*-D-2CP average Rg value appeared to stabilize moderately at the start of MD simulation but showed noticeable spikes at 42 ns and 63 ns (1.75 nm). The *DehHsAAD6*-3CP complex also behaved similarly and equilibrated to an average Rg near the end of the simulation (average Rg value = 1.68 nm). For *DehHsAAD6*-2,3-DCP, the Rg value gradually decreased from 1.79 to 1.67 nm and briefly peaked at 80 ns (Figure 10) before reaching equilibrium (average Rg −1.75 nm). The Rg value of the *DehHsAAD6*-2,2-DCP gradually decreased from 1.74–1.64 nm (Figure 10).

Based on the computer modelling research, the shift in Rg is best explained in terms of protein compactness with ligands. Comparatively, the Rg of DeHsAAD6 protein was presented in S1. Consequently, enhancement in protein

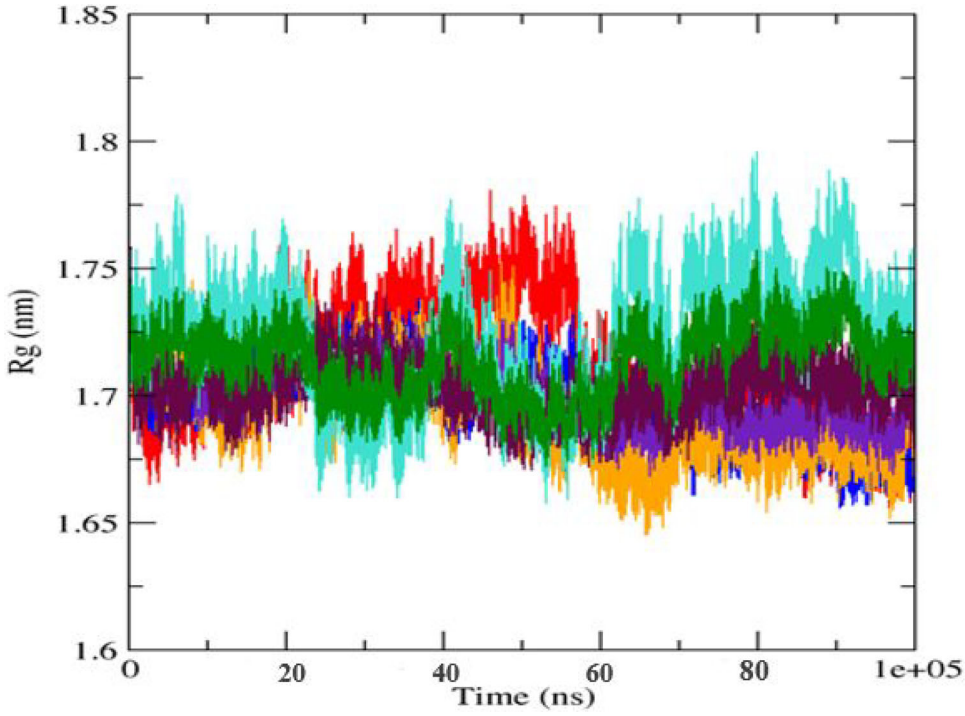

Figure 10. Average of radius of gyration is shown as a 100 ns simulation time function at 300 K, showing *DehHsAAD6* against L-2CP in maroon, MCA in purple, 2,2-DCP in orange, D-2CP in green, 3CP in Red, 2,3-DCP in turquoise, and TCA in blue showing radius of gyration (Rg) plotted together.

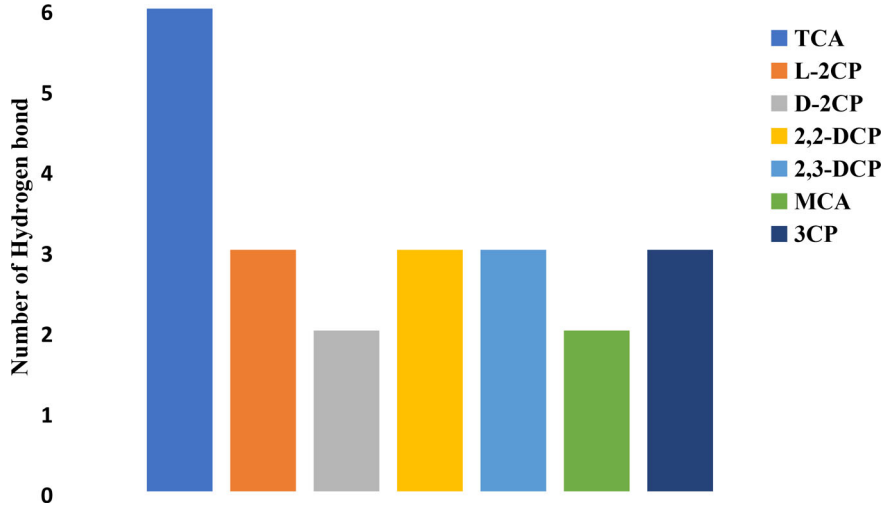

Figure 11. Number of intermolecular hydrogen bonds at 100 ns simulation for *DehHsAAD6* when bonded to different haloacetate (TCA & MCA) and haloacid (L-2CP, 3CP, 2,2-DCP, 2,3-DCP & D-2CP).

compactness on ligand binding can be elucidated by low values of Rg. The Rg plots for the trajectories exhibited better structural compactness of all the ligands with DehHsAAD6 (Rg range 1.64–1.78 nm) than Rg of DehHsAAD6 protein (1.77–1.83 nm) (S2). It should be noted that the compactness of DehHsAAD6 did not remain constant throughout the simulations, fluctuating somewhat at 5 ns, 30 ns, 45 ns, 60 ns and 85 ns, with an average Rg of 1.83 nm (S2). According to Lobanov et al. (2008), the highest Rg plot proposes a looser packing of amino acids and vice versa when the Rg value is at the lowest. Hence, the relatively low Rg values of all complexes represented their tight bonding to *DehHsAAD6*. The findings also point to highly compact *DehHsAAD6*-ligand complexes.

3.6.4. Hydrogen bonds analysis

According to Chen et al. (2016), hydrogen bonds and their relative strength in a water environment are vital to enable protein-ligand binding, particularly when the mechanism of action involves hydrolysis. In any case, the important role of water in the breakdown of a compound cannot be dismissed. Figure 8 illustrates the intermolecular hydrogen bonds formed between *DehHsAAD6* and the substrates. The study found that six hydrogen bonds (H-bonds) were formed in the *DehHsAAD6*-TCA (Figure 8) complex compared to only three bonds in complexes *DehHsAAD6*-L-2CP, *DehHsAAD6*-3CP, *DehHsAAD6*-2,2-DCP, and *DehHsAAD6*-2,3-DCP (Figure 11). The *DehHsAAD6*-MCA and *DehHsAAD6*-D-2CP (Figure 11) complexes showed two hydrogen bonds.

The *DehHsAAD6*-ligand outputs seen in the MD simulation have more hydrogen bonds than those observed in the corresponding molecular docking study. This is because MD simulation incorporates water into the system, which closely resembles the actual enzymatic hydrolysis system (Anuar et al., 2020; Oyewusi et al., 2020). Thus, the number of formed hydrogen bonds constantly changed throughout the simulation time. However, they were more profound in the *DehHsAAD6*-2,3-DCP and *DehHsAAD6*-2,2-DCP complexes.

In this investigation, RMSD, RMSF, Rg and hydrogen bonding were the parameters computed and analysed after a 100 ns dynamics trajectory. In fact, the context in which RMSD and RMSF are performed is because such stabilities and flexibilities are required to acquire good binding interaction. Also, low Rg values can reveal an increase in protein compactness on ligand binding. While a high number of hydrogen bonds between the ligand and key residues associated with protein activity indicates a high binding activity. It must be stressed that the binding free energies derived from molecular docking analysis for all the complexes were found to agree with the RMSD and RMSF. Consequently, it is necessary to strike a balance between computational reliability and efficacy with experimental investigation especially in biological processes or systems. For this reason, an experimental investigation is needed to provide more trustworthy information on the enzyme activity with ligands, which can be based on enzymatic kinetics and thermodynamics. In particular, future study on enzyme kinetics will be required to confirm the binding affinity.

4. Conclusion

Molecular docking and MD simulations successfully predicted the order of substrate preference by the putative dehalogenase DehHsAAD6 derived from genomic *H. smyrnensis* AAD6. Undoubtedly, the DehHsAAD6 showed preferential interactions with different haloalkanoates and haloacetates to form stable complexes in the molecular docking and MD simulation. The study discovered that the DehHsAAD6 has a broad substrate specificity although L-2CP and MCA were shown to be the enzymes' top two most favored substrates. Hence, it was confirmed that the putative DehHsAAD6 dehalogenase is a possible exceptional enzyme for treating halogen-contaminated environments provided that all these substrates were substrates for the bacterium as well.

According to study halophilic bacterium *Halomonas smyrnensis* AAD6T could be used to detoxify or degrade hazardous halogenated organic pollutants in the environment. The existence of a putative dehalogenase enzyme (DehHsAAD6) in the genome of this bacterium suggests its functionality and usefulness in degrading or absorbing pollutants as well as presents a clear picture of its use in bioremediation of polluted environments. Having said this, the *in-silico* would speed up the process of selecting a specific pollutant using computational scrutiny, which could provide valuable insights into the explanation of dehalogenase (DehHsAAD6) mechanism for extended substrate selectivity and catalytic activity. It must be noted that the study needed to be validated in a wet laboratory level degradation assay to thoroughly screen for degradation and fate of concern pollutants, and then scaled up for real-time use on polluted sites for environmental safety.

Acknowledgments

The authors would like to express their gratitude to the Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia for their facilities.

Disclosure statement

There are no conflicts to declare.

ORCID

Fahrul Huyop http://orcid.org/0000-0003-3978-4087

References

Adamu, A., Wahab, R. A., & Huyop, F. (2016). L-2-Haloacid dehalogenase (DehL) from Rhizobium sp. RC1. *SpringerPlus*, *5*(1), 1–17. https://doi.org/10.1186/s40064-016-2328-9

Anbarasu, K., & Jayanthi, S. (2018). Identification of curcumin derivatives as human LMTK3 inhibitors for breast cancer: A docking, dynamics, and MM/PBSA approach. *3 Biotech*, *8*(5), 228. https://doi.org/10.1007/s13205-018-1239-6

Anuar, N. F. S. K., Wahab, R. A., Huyop, F., Amran, S. I., Hamid, A. A. A., Halim, K. B. A., & Hood, M. H. M. (2020). Molecular docking and molecular dynamics simulations of a mutant *Acinetobacter haemolyticus* alkaline-stable lipase against tributyrin. *Journal of Biomolecular Structure and Dynamics*, *39*(6), 2079–2091.

Ates, O., Arga, K. Y., & Oner, E. T. (2013). The stimulatory effect of mannitol on levan biosynthesis: Lessons from metabolic systems analysis of *Halomonas smyrnensis* AAD6T. *Biotechnology Progress*, *29*(6), 1386–1397. https://doi.org/10.1002/btpr.1823

Awasthi, M. K., Wong, J. W., Kumar, S., Awasthi, S. K., Wang, Q., Wang, M., Ren, X., Zhao, J., Chen, H., & Zhang, Z. (2018). Biodegradation of food waste using microbial cultures producing thermostable α-amylase and cellulase under different pH and temperature. *Bioresource Technology*, *248*(Pt B), 160–170.

Aydin, B., Ozer, T., Oner, E. T., & Arga, K. Y. (2018). The genome-based metabolic systems engineering to boost levan production in a halophilic bacterial model. *Omics: A Journal of Integrative Biology*, *22*(3), 198–209. https://doi.org/10.1089/omi.2017.0216

Buchan, D. W., & Jones, D. T. (2019). The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Research*, *47*(W1), W402–W407.

Chen, D., Oezguen, N., Urvil, P., Ferguson, C., Dann, S. M., & Savidge, T. C. (2016). Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Science Advances*, *2*(3), e1501240.

Cohen, I., Coban, M., Shahar, A., Sankaran, B., Hockla, A., Lacham, S., Caulfield, T. R., Radisky, E. S., & Papo, N. (2019). Disulfide engineering of human Kunitz-type serine protease inhibitors enhances proteolytic stability and target affinity toward mesotrypsin. *The Journal of Biological Chemistry*, *294*(13), 5105–5120.

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science: A Publication of the Protein Society*, *2*(9), 1511–1519.

Combet, C., Blanchet, C., Geourjon, C., & Deleage, G. (2000). NPS@: Network protein sequence analysis. *Trends in Biochemical Sciences*, *25*(3), 147–150. https://doi.org/10.1016/s0968-0004(99)01540-6

DeLano, W. L. (2002). Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on Protein Crystallography*, *40*(1), 82–92.

Diken, E., Ozer, T., Arikan, M., Emrence, Z., Oner, E. T., Ustek, D., & Arga, K. Y. (2015). Genomic analysis reveals the biotechnological and

industrial potential of levan producing halophilic extremophile. *Halomonas smyrnensis AAD6T*. *SpringerPlus*, *4*(1), 1–11.

Dong, Y. W., Liao, M. L., Meng, X. L., & Somero, G. N. (2018). Structural flexibility and protein adaptation to temperature: Molecular dynamics analysis of malate dehydrogenases of marine molluscs. *Proceedings of the National Academy of Sciences*, *115*(6), 1274–1279. https://doi.org/10.1073/pnas.1718910115

Dutta, B., Banerjee, A., Chakraborty, P., & Bandopadhyay, R. (2018). In silico studies on bacterial xylanase enzyme: Structural and functional insight. *Journal, Genetic Engineering & Biotechnology*, *16*(2), 749–756.

Edbeib, M. F., Wahab, R. A., Kaya, Y., & Huyop, F. (2017). In silico characterization of a novel dehalogenase (DehHX) from the halophile Pseudomonas halophila HX isolated from Tuz Gölü Lake, Turkey: Insights into a hypersaline-adapted dehalogenase. *Annals of Microbiology*, *67*(5), 371–382. https://doi.org/10.1007/s13213-017-1266-2

Elmezayen, A., Al-Obaidi, A., Şahin, A., Yelekçi, K. (2020). Drug repurposing for coronavirus (COVID-19): in silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes. *Journal of Biomolecular Structure and Dynamics*, 1–13.

Feig, M. (2017). Computational protein structure refinement: Almost there, yet still so far to go. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *7*(3), e1307.

Fu, Y., Zhao, J., & Chen, Z. (2018). Insights into the molecular mechanisms of protein-ligand interactions by molecular docking and molecular dynamics simulation: A case of oligopeptide binding protein. *Computational and Mathematical Methods in Medicine*, *2018*, 3502512–3502514. https://doi.org/10.1155/2018/3502514

Fuentes, D., Muñoz, N. M., Guo, C., Polak, U., Minhaj, A. A., Allen, W. J., Gustin, M. C., & Cressman, E. N. (2018). A molecular dynamics approach towards evaluating osmotic and thermal stress in the extracellular environment. *International Journal of Hyperthermia: The Official Journal of European Society for Hyperthermic Oncology, North American Hyperthermia Group*, *35*(1), 559–567.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In Walker J.M. (Ed.), *The proteomics protocols handbook* (pp. 571–607). Springer.

Geourjon, C., & Deleage, G. (1995). SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer Applications in the Biosciences: CABIOS*, *11*(6), 681–684. https://doi.org/10.1093/bioinformatics/11.6.681

Gribble, G. W. (2003). The diversity of naturally produced organohalogens. *Chemosphere*, *52*(2), 289–297. https://doi.org/10.1016/S0045-6535(03)00207-8

Gribble, G. W. (2012). Recently discovered naturally occurring heterocyclic organohalogen compounds. *Heterocycles*, *84*(1), 157–207.

Gupta, P., & Diwan, B. (2017). Bacterial exopolysaccharide mediated heavy metal removal: A review on biosynthesis, mechanism and remediation strategies. *Biotechnology Reports (Amsterdam, Netherlands)*, *13*, 58–71.

Gurung, A., Ali, M., Lee, J., Farah, M., Al-Anazi, K. (2020). Unravelling lead antiviral phytochemicals for the inhibition of SARS-CoV-2 Mpro enzyme through in silico approach. *Life sciences*, *255*, 117831.

Hamid, A. A. A., Hamid, T. H. T. A., Wahab, R. A., Omar, M. S. S., & Huyop, F. (2015a). An S188V mutation alters substrate specificity of non-stereospecific α-Haloalkanoic acid Dehalogenase E (DehE). *PLoS One*, *10*(3), e0121687. https://doi.org/10.1371/journal.pone.0121687

Hamid, A. A. A., Tengku Abdul Hamid, T. H., Wahab, R. A., & Huyop, F. (2015b). Identification of functional residues essential for dehalogenation by the non-stereospecific α-haloalkanoic acid dehalogenase from Rhizobium sp. RC1. *Journal of Basic Microbiology*, *55*(3), 324–330. https://doi.org/10.1002/jobm.201300526

Houston, D. R., & Walkinshaw, M. D. (2013). Consensus docking: Improving the reliability of docking in a virtual screening context. *Journal of Chemical Information and Modeling*, *53*(2), 384–390. https://doi.org/10.1021/ci300399w

Hutcheon, G. W., Vasisht, N., & Bolhuis, A. (2005). Characterisation of a highly stable α-amylase from the halophilic archaeon *Haloarcula hispanica*. *Extremophiles: Life Under Extreme Conditions*, *9*(6), 487–495.

Junaid, M., Muhseen, Z. T., Ullah, A., Wadood, A., Liu, J., & Zhang, H. (2014). Molecular modeling and molecular dynamics simulation study of the human Rab9 and Rhobtb3 C-terminus complex. *Bioinformation*, *10*(12), 757–763. https://doi.org/10.6026/97320630010757

Kovacic, F., Mandrysch, A., Poojari, C., Strodel, B., & Jaeger, K. E. (2016). Structural features determining thermal adaptation of esterases. *Protein Engineering, Design & Selection: PEDS*, *29*(2), 65–76.

Koziara, K. B., Stroet, M., Malde, A. K., & Mark, A. E. (2014). Testing and validation of the Automated Topology Builder (ATB) version 2.0: Prediction of hydration free enthalpies. *Journal of Computer-Aided Molecular Design*, *28*(3), 221–233.

Kumar, C. V., Swetha, R. G., Anbarasu, A., & Ramaiah, S. (2014). Computational analysis reveals the association of threonine 118 methionine mutation in PMP22 resulting in CMT-1A. *Advances in Bioinformatics*, *2014*, 502618. https://doi.org/10.1155/2014/502618

Lanyi, J. K. (1974). Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriological Reviews*, *38*(3), 272–290. https://doi.org/10.1128/br.38.3.272-290.1974

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(2), 283–291. https://doi.org/10.1107/S0021889892009944

Lee, H. S., Qi, Y., & Im, W. (2015). Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. *Scientific Reports*, *5*, 8926. https://doi.org/10.1038/srep08926

Lemkul, J. (2018). From proteins to perturbed Hamiltonians: A suite of tutorials for the GROMACS-2018 molecular simulation package. *Living Journal of Computational Molecular Science*, *1*(1), 5068. [article v1. 0].

Lemmon, G., & Meiler, J. (2013). Towards ligand docking including explicit interface water molecules. *PLoS One*, *8*(6), e67536.

Liao, K. H., Chen, K. B., Lee, W. Y., Sun, M. F., Lee, C. C., & Chen, C. Y. C. (2014). Ligand-based and structure-based investigation for Alzheimer's disease from traditional Chinese medicine. *Evidence-Based Complementary and Alternative Medicine*, *2014*, 1–16. https://doi.org/10.1155/2014/628712

Lindahl, E., Hess, B., & Van Der Spoel, D. (2001). Gromacs 3.0: A package for molecular simulation and trajectory analysis. *Journal of Molecular Modeling*, *7*(8), 306–317. https://doi.org/10.1007/s008940100045

Lobanov, M. Y., Bogatyreva, N., & Galzitskaya, O. (2008). Radius of gyration as an indicator of protein structure compactness. *Molecular Biology*, *42*(4), 623–628. https://doi.org/10.1134/S0026893308040195

Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, *356*(6364), 83–85.

Malde, A. K., Zuo, L., Breeze, M., Stroet, M., Poger, D., Nair, P. C., Oostenbrink, C., & Mark, A. E. (2011). An automated force field Topology Builder (ATB) and repository: Version 1.0. *Journal of Chemical Theory and Computation*, *7*(12), 4026–4037.

Mishra, R., Mazumder, A., Mazumder, R., Mishra, P. S., & Chaudhary, P. (2019). Docking study and result conclusion of heterocyclic derivatives having urea and acyl moiety. *Asian Journal of Biomedical and Pharmaceutical Sciences*, *9*(67), 13. https://doi.org/10.35841/2249-622X.67.19-082

Mustafa, G., Iqbal, M. J., Hassan, M., & Jamil, A. (2017). Bioinformatics characterization of growth differentiation factor 11 of *Oryctolagus cuniculus*. *Journal of the Chemical Society of Pakistan*, *39*(6), 1089–1089.

Nemaysh, V., & Luthra, P. M. (2017). Computational analysis revealing that K634 and T681 mutations modulate the 3D-structure of PDGFR-β and lead to sunitinib resistance. *RSC Advances*, *7*(60), 37612–37626. https://doi.org/10.1039/C7RA01305A

Nwodo, U. U., Green, E., & Okoh, A. I. (2012). Bacterial exopolysaccharides: Functionality and prospects. *International Journal of Molecular Sciences*, *13*(11), 14002–14015.

Oostenbrink, C., Villa, A., Mark, A. E., & Van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, *25*(13), 1656–1676.

Oren, A., Larimer, F., Richardson, P., Lapidus, A., & Csonka, L. N. (2005). How to be moderately halophilic with broad salt tolerance: Clues

from the genome of *Chromohalobacter salexigens*. *Extremophiles: Life Under Extreme Conditions*, *9*(4), 275–279. https://doi.org/10.1007/s00792-005-0442-7

Oyewusi, H. A., Huyop, F., & Wahab, R. A. (2020). Molecular docking and molecular dynamics simulation of *Bacillus thuringiensis* dehalogenase against haloacids, haloacetates and chlorpyrifos. *Journal of Biomolecular Structure and Dynamics*, 1–16. https://doi.org/10.1080/07391102.2020.1835727

Oyewusi, H. A., Huyop, F., Wahab, R. A., & Hamid, A. A. A. (2021b). In silico assessment of dehalogenase from *Bacillus thuringiensis* H2 in relation to its salinity-stability and pollutants degradation. *Journal of Biomolecular Structure and Dynamics*, 1–15. https://doi.org/10.1080/07391102.2021.1927846

Oyewusi, H. A., Wahab, R. A., & Huyop, F. (2021a). Whole genome strategies and bioremediation insight into dehalogenase-producing bacteria. *Molecular Biology Reports*, *48*(3), 2687–2615. https://doi.org/10.1007/s11033-021-06239-7

Park, H., Ovchinnikov, S., Kim, D. E., DiMaio, F., & Baker, D. (2018). Protein homology model refinement by large-scale energy optimization. *Proceedings of the National Academy of Sciences*, *115*(12), 3054–3059. https://doi.org/10.1073/pnas.1719115115

Paul, S., Bag, S. K., Das, S., Harvill, E. T., & Dutta, C. (2008). Molecular signature of hypersaline adaptation: Insights from genome and proteome composition of halophilic prokaryotes. *Genome Biology*, *9*(4), R70.

Poli, A., Kazak, H., Gürleyendağ, B., Tommonaro, G., Pieretti, G., Öner, E. T., & Nicolaus, B. (2009). High level synthesis of levan by a novel Halomonas species growing on defined media. *Carbohydrate Polymers*, *78*(4), 651–657. https://doi.org/10.1016/j.carbpol.2009.05.031

Poli, A., Nicolaus, B., Denizci, A. A., Yavuzturk, B., & Kazan, D. (2013). *Halomonas smyrnensis* sp. nov., a moderately halophilic, exopolysaccharide-producing bacterium. *International Journal of Systematic and Evolutionary Microbiology*, *63*(Pt 1), 10–18.

Ramírez, D., & Caballero, J. (2018). Is it reliable to take the molecular docking top scoring position as the best solution without considering available structural data? *Molecules*, *23*(5), 1038. https://doi.org/10.3390/molecules23051038

Rosdi, M. N., Mohd Arif, S., Abu Bakar, M. H., Razali, S. A., Mohamed Zulkifli, R., & Ya'akob, H. (2018). Molecular docking studies of bioactive compounds from Annona muricata Linn as potential inhibitors for Bcl-2, Bcl-w and Mcl-1 antiapoptotic proteins. *Apoptosis: An International Journal on Programmed Cell Death*, *23*(1), 27–40.

Rost, B., & Sander, C. (1994). Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins: Structure, Function, and Genetics*, *19*(1), 55–72. https://doi.org/10.1002/prot.340190108

Salama, Y., Chennaoui, M., Sylla, A., Mountadar, M., Rihani, M., & Assobhei, O. (2016). Characterization, structure, and function of extracellular polymeric substances (EPS) of microbial biofilm in biological wastewater treatment systems: A review. *Desalination and Water Treatment*, *57*(35), 16220–16237. https://doi.org/10.1080/19443994.2015.1077739

Sam, S., Kucukasik, F., Yenigun, O., Nicolaus, B., Oner, E. T., & Yukselen, M. A. (2011). Flocculating performances of exopolysaccharides produced by a halophilic bacterial strain cultivated on agro-industrial waste. *Bioresource Technology*, *102*(2), 1788–1794.

Sarilmiser, H. K., Ates, O., Ozdemir, G., Arga, K. Y., & Oner, E. T. (2015). Effective stimulating factors for microbial levan production by *Halomonas smyrnensis* AAD6T. *Journal of Bioscience and Bioengineering*, *119*(4), 455–463.

Sarilmiser, H. K., & Oner, E. T. (2014). Investigation of anti-cancer activity of linear and aldehyde-activated levan from *Halomonas smyrnensis* AAD6T. *Biochemical Engineering Journal*, *92*, 28–34.

Sezer, A. D., Kazak Sarılmışer, H., Rayaman, E., Çevikbaş, A., Öner, E. T., & Akbuğa, J. (2017). Development and characterization of vancomycin-loaded levan-based microparticular system for drug delivery. *Pharmaceutical Development and Technology*, *22*(5), 627–634.

Sogutcu, E., Emrence, Z., Arikan, M., Cakiris, A., Abaci, N., Öner, E. T., Üstek, D., & Arga, K. Y. (2012). Draft genome sequence of *Halomonas smyrnensis* AAD6T. *Genome Announcement*, *194*(20), 5690–5991.

Tohme, S., Hacıosmanoğlu, G. G., Eroğlu, M. S., Kasavi, C., Genç, S., Can, Z. S., & Oner, E. T. (2018). *Halomonas smyrnensis* as a cell factory for co-production of PHB and levan. *International Journal of Biological Macromolecules*, *118*(Pt A), 1238–1246.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Lepore, R. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303.

Yan, C., Xu, X., & Zou, X. (2016). Fully blind docking at the atomic level for protein-peptide complex structure prediction. *Structure (London, England: 1993)*, *24*(10), 1842–1853. https://doi.org/10.1016/j.str.2016.07.021

Yang, J., Roy, A., & Zhang, Y. (2013). Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics*, *29*(20), 2588–2595. https://doi.org/10.1093/bioinformatics/btt447

Yu, J., Shi, J., Zhang, Y., & Yu, Z. (2020). Molecular docking and site-directed mutagenesis of dichloromethane dehalogenase to improve enzyme activity for dichloromethane degradation. *Applied Biochemistry and Biotechnology*, *190*(2), 487–505.

Zhang, Y. (2009). Protein structure prediction: When is it useful? *Current Opinion in Structural Biology*, *19*(2), 145–155. https://doi.org/10.1016/j.sbi.2009.02.005

Zhu, J., Qi, R., Liu, Y., Zhao, L., & Han, W. (2019). Mechanistic insights into the effect of ligands on structural stability and selectivity of sulfotransferase 2A1 (SULT2A1). *ACS Omega*, *4*(26), 22021–22034.

Zorgani, M. A., Patron, K., & Desvaux, M. (2014). New insight in the structural features of haloadaptation in α-amylases from halophilic Archaea following homology modeling strategy: Folded and stable conformation maintained through low hydrophobicity and highly negative charged surface. *Journal of Computer-Aided Molecular Design*, *28*(7), 721–734.