Environmental Sciences Europe

Bioinformatics analysis and molecular dynamics simulations of azoreductases (AzrBmH2) from *Bacillus megaterium* H2 for the decolorization of commercial dyes

Habeebat Adekilekun Oyewusi[1,3,4*], Roswanira Abdul Wahab[2,3], Kolajo Adedamola Akinyede[4], Ghadeer M. Albadrani[5], Muath Q. Al-Ghadi[6], Mohamed M. Abdel-Daim[7,8], Basiru Olaitan Ajiboye[9] and Fahrul Huyop[1,3*]

Abstract

The present study aimed to investigate the decolorization of various commercial dyes by azoreductases (AzrBmH21, AzrBmH22/3, and AzrBmH24/5) through bioinformatics means, comprising molecular docking, molecular dynamics simulation, and molecular mechanics Poisson–Boltzmann surface area (MM-PBSA). Therefore, four commercial dyes, namely acid orange 7, cresol red, methylene blue, and malachite green, were selected as potential targets for degradation by the above said azoreductases derived from *Bacillus megaterium* H2. The prediction of ligand binding or catalytic sites for AzrBmH21, AzrBmH22/3, and AzrBmH24/5 were performed using a machine learning algorithm based on the Prank Web and DeepSite chemoinformatic tool. This analysis revealed that several amino acids of AzrBmH2 interacted with the tested dyes, indicating the presence of distinct ligand-binding sites for AzrBmH2-dye complexes. Likewise, the binding affinity for AzrBmH21, AzrBmH22/3, and AzrBmH24/5 ranged from − 9.4 to − 5.5 kcal/mol, − 9.2 to − 5.4 kcal/mol, and − 9.0 to − 5.4 kcal/mol, respectively, with each complex stabilized at a minimum of 0–5 hydrogen bonds. MD simulations revealed stable AzrBmH2-dye complexes with RMSD and RMSF values ranging from 0.15 to 0.42 nm and 0.05 to 0.48 nm, respectively, with Rg values between 1.75 and 1.88 nm. MM-PBSA calculations indicated that the AzrBmH2–dye complexes, except for AzrBmH2–malachite green, exhibited the lowest binding energy (− 191.05 ± 7.08 to 314.19 ± 6.88 kcal/mol), with prevalent hydrophobic interactions (− 268.25 ± 12.25 to − 418.92 ± 29.45 kcal/mol) through van der Waals forces. Therefore, this study was able to highlight the potential role of enzymes, specifically azoreductases from *Bacillus megaterium* H2, in predicting the decolorization of commercial dyes. These findings could contribute to our understanding of the azoreductases' mechanisms in bioremediation and for biotechnological applications.

Keywords Azoreductases, *Bacillus megaterium*, Degradation, Synthetic dyes, Molecular docking, Molecular dynamics simulation

*Correspondence:
Habeebat Adekilekun Oyewusi
haoyewusi@utm.my
Fahrul Huyop
fahrul@utm.my
Full list of author information is available at the end of the article

Introduction

Textile industry is one of the global giant and oldest industries that has a significant challenge with wastewater treatment. The discharged wastewater contains salts, dyes, heavy metals, and levelling agents that pose a threat to the environment. It has been reported that approximately 20% of textile dyes are lost during manufacturing [35]. The discharge of textile wastewater into water bodies can disrupt aquatic ecosystems and decrease oxygen solubility, posing severe environmental and health risks [14]. Consequently, releasing azo-dye wastewater into the environment can result in ecological hazards. Azo dyes are extensively used in textile industries due to their cost-effectiveness, versatility in color range, ease of manufacturing, and resistance to damage. Efficient color removal and treatment of harmful by-products are crucial for the reuse of treated water. Biological decolorization and degradation of azo dyes offer an eco-friendly, efficient, and cost-effective option with lower sludge production and higher efficiency for a wide range of azo dyes [7, 22]. Bacterial systems are particularly interesting for degradation due to their ease of synthesis, maintenance, ability to survive in stressful conditions, and the presence of various oxidative and reductive enzymes for biotransformation.

Enzymes catalyzing the reduction of azo dyes are called azoreductases (EC 1.7.1.6), utilizing NADH/FMN and/or NADPH as electron donors to decolorize these dyes by reductively cleaving azo bonds to form aromatic amines as the degradation products. The rate-limiting step is decolorization, followed by aerobic mineralization of the colorless aromatic amines. Scientists worldwide focus on using in silico methods before in situ or ex situ biodegradation of toxicants to identify ecologically sustainable strains for field-level applications. This in silico approach saves time and resources, making way for future bioremediation technologies [17, 32]. Bioinformatics tools provide insights into bioremediation mechanisms by predicting the propensity of dye degradation via oxidoreductive enzymes [1, 20]. The systems biology approach for microbial bioremediation reduces the time and cost involved in preliminary screening for identifying feasible strains for effective treatment of textile effluents. Molecular docking and molecular dynamics simulation can optimize and enhance our understanding of protein–ligand interactions [21]. In particular, docking can be used to predict amino acid residues involved in enzyme–dye binding.

Halophilic microorganisms growing in extreme environments, like hypersaline lakes, also show promise in degrading azo dyes. Combining bioinformatics and bacterial remediation, this study aims to validate innovative approaches for treating textile wastewater. *Bacillus thuringiensis* H2, isolated from hypersaline Lake Tuz in Turkey, has shown an innate ability to render pollutants such as haloacids, haloacetates, and chlorpyrifos less detrimental in the presence or absence of salt [19, 24]. Previously, an azoreductase enzyme (AzrBmH2) was identified from the bacterium genome and analyzed using silico methods [18]. Integrating bioinformatics and bacterial remediation in this study is an interesting approach to validating innovative strategies to treat textile wastewater. Furthermore, this in silico strategy, prior to empirical ones, conserves time and resources, making way for future bioremediation technologies.

The main objective of the present study is to validate a novel approach combining bioinformatics and bacterial remediation using non-native strains for decolorizing and degrading azo dyes. Specifically, we focused on using the identified AzrBmH2 enzyme to decolorize textile azo dyes (methylene blue, malachite green, cresol red, and acid orange 7). Here, we performed in silico studies (molecular docking and active site residue analysis) to investigate the interaction between dyes and oxidoreductive enzymes, such as azoreductases (AzrBmH2) present in the bacterium genome, which are responsible for dye degradation. Through this approach, we successfully identified the amino acid residues involved in enzyme–dye binding, which might prove valuable in eliminating time-consuming and cumbersome preliminary screening procedures in future bioremediation endeavors.

Materials and methods
Proteins and ligands preparation

The homology-modeled azoreductase (AzrBmH2) used in this study was previously reported [18]. Briefly, five amino acid sequences of azoreductases, namely AzrBmH21, AzrBmH22, AzrBmH23, AzrBmH24, and AzrBmH25, were annotated from the genome of *Bacillus megaterium* H2 (GenBank accession number: CP071466-CP071484) [19] and modeled using the SWISS-MODEL web server [37]. The three-dimensional (3-D) structure of the AzrBmH21 protein was reported in our previous study [18], resembling the crystal structure of FMN-dependent NADH-azoreductase from *Bacillus* sp. A01 (PDB ID: 6qu0). Similarly, the 3-D structures of AzrBmH22/3 and AzrBmH24/5 were predicted to resemble the crystal structures of FMN-dependent NADH-indigo reductase from *Bacillus smithii* type strain DSM 4216 (PDB ID: 6JXS) and *Bacillus* sp. B29 (PDB ID: 3W77), respectively [18]. The 3D structures of the dye molecules, including methylene blue, malachite green, cresol red, and acid orange 7, were obtained from the PubChem database as 3D modeling conformers. Figure 1 depicts the structures of the tested dyes prepared using the University of California San Francisco (UCSF)

Oyewusi *et al. Environmental Sciences Europe*   (2024) 36:31

Page 3 of 20
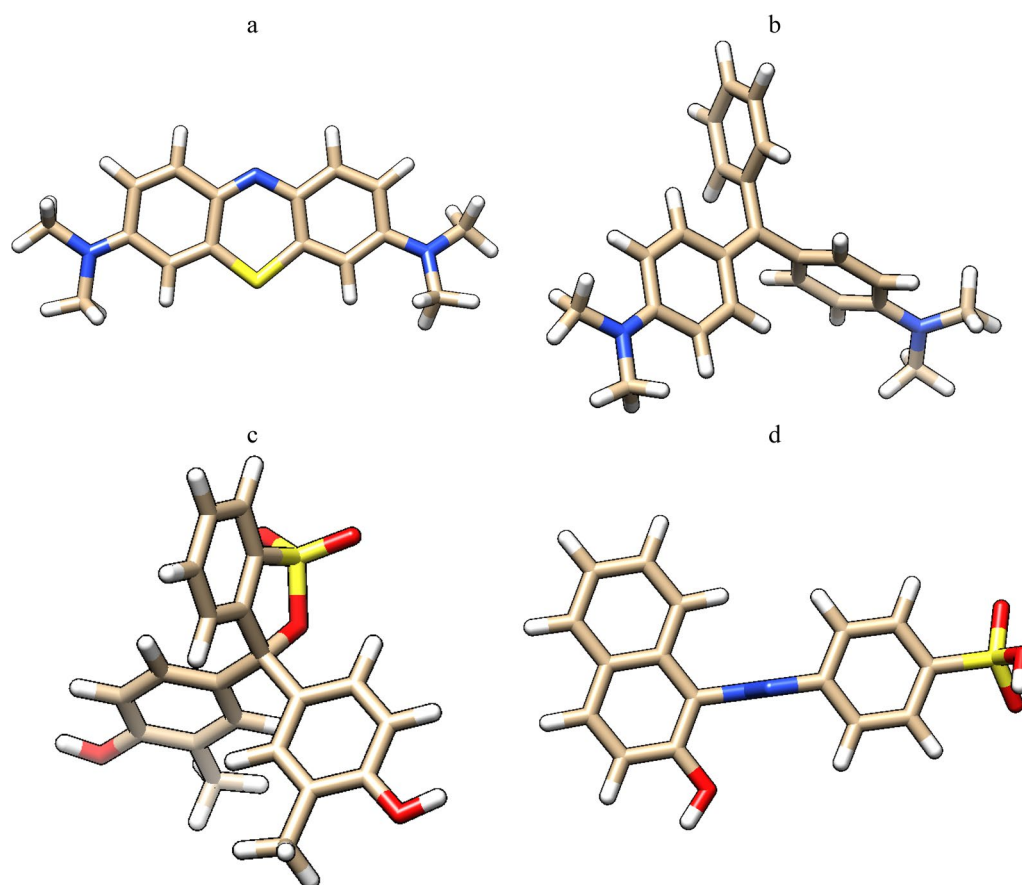
Fig. 1 Structures of synthetic dyes assessed by this study **a** methylene blue, **b** malachite green, **c** cresol red, and **d** acid orange 7

Chimera (Version 1.17.3) [3, 9]. The SDF files of the ligands were converted to PDB files using O-BABEL (version 2.3.1), and hydrogens were added to the 3D models of the ligands using Avogadro software. Finally, the structures were submitted to the Automated Topology Builder (ATB) and Repository version 3.0 server to optimize the molecular geometries.

Ligand binding sites prediction

The 3-D structures of novel potential targets were analyzed to identify binding sites for small organic molecules and assess their suitability for bioremediation. To achieve this, we employed the PrankWeb and DeepSite web servers. PrankWeb and DeepSite utilize machine learning algorithms to predict ligand-binding sites based on various factors, including protein 3D structure, sequence information, binding pocket lists, and visual analysis. PrankWeb and DeepSite can predict new binding sites and are designed to handle multidomain protein structures (Krivak and Hoksza, [11]). This analysis was to examine the binding site coordinates and identify the specific amino acids within the pockets of the azoreductase proteins. To accomplish this, we uploaded the 3D models of the azoreductases (6QU0, 6JXS, and 3W77) to the PrankWeb and DeepSite web servers and submitted them for analysis.

To reconfirm the predicted catalytic amino acids of the AzrBmH2 proteins, COACH-D algorithm (https://yanglab.qd.sdu.edu.cn/COACH-D/) [6, 27] and CAVER Web 1.0 tools (https://loschmidt.chemi.muni.cz/caverweb/) were employed in this study. COACH-D determined and validated the best active site before the output was subjected to quantum tunneling using CAVER Web 1.0 tools. The latter is useful in comprehending the intended ligand modalities accommodated within the AzrBmH2 proteins.

Molecular docking

To determine the optimal conformation and orientation of ligands when interacting with the target protein, molecular docking was employed in this study. The purpose was to assess the binding affinity of four azo dyes in complex with different AzrBmH2 proteins (AzrBmH21, AzrBmH22/3, and AzrBmH24/5). All molecular docking

simulations were conducted using Autodock version 4.2.6 and AutoTools 1.5.6 software. AutoDock Tools was utilized in this study to identify the rotatable bonds in the ligands, generating various conformers for the docking process. Additionally, AutoDock Tools was used to create PDBQT files for both the ligands and AzrBmH2 proteins. In these files, Gasteiger and Kollman charges were calculated for each atom of the AzrBmH2 molecules and ligands, ensuring accurate torsion adoption for rotation during the docking simulation. This step was carried out using AutoDock version 4.2.6. Three-dimensional grids were created to identify the ligand binding sites and binding residues, as well as to generate affinity maps for all the atom types. For AzrBmH21, the grid dimensions were set to $56 \times 80 \times 42$ Å with a spacing of ± 1.000 Å. Similarly, for AzrBmH22/3, the grid dimensions were $58 \times 68 \times 40$ Å, and for AzrBmH24/5, the grid dimensions were $66 \times 54 \times 44$ Å. These grids were generated using AutoGrid, and they allowed for the calculation of affinity values for all the atom types present in the system.

In the docking simulations, a total of 100 runs were performed for each flexible ligand. Within each run, 10 million energy evaluations were conducted to explore the conformational space and evaluate the binding affinities. The docking analysis was carried out using AutoDock Vina, which allowed for the identification of energetically favorable binding poses. The resulting conformations were then subjected to a conformation cluster analysis based on the root mean square deviation (RMSD) values, using the starting structure as the reference. The conformations were ranked according to their binding energy in units of kcal/mol, with lower energy values indicating stronger binding interactions. Furthermore, to visualize the hydrogen bond and hydrophobic interactions between the ligands and the binding site residues, LigPlot + software was used [26]. The LigPlot + generated schematic diagrams illustrating the best docking result for each substrate, where the interactions were depicted along with the corresponding binding energy in units of kcal/mol. This analysis provided a visual representation of the key molecular interactions involved in the ligand–protein binding [36].

Molecular dynamics (MD) simulation

Molecular dynamics (MD) simulations were conducted using the parallel version of the GROMACS 5.1.2 software, employing the gromos96 53a7 force field (http://www.gromacs.org/About_Gromacs). These simulations aimed to assess the stability and flexibility of the docked AzrBmH2–dyes complexes. The simulation system required all the complexes to be solvated in a cubic water box with dimensions of $90.0 \times 90.0 \times 90.0$ nm. The water molecules were modeled using the Simple Point Charge (SPC) water model, as proposed by Wu et al. [38]. Sodium ions were introduced to neutralize the entire system, by substituting water molecules to ensure charge neutrality of the AzrBmH2 proteins. Before the molecular dynamic simulations, an energy minimization step was performed using the steepest descent algorithm. The minimization process involved 735 steps for AzrBmH21, 978 steps for AzrBmH22/3, and 938 steps for AzrBmH24/5. This energy minimization step aimed to alleviate steric clashes or inappropriate geometries in the solvated protein–ligand complex system, ensuring a more stable starting configuration for the subsequent simulations.

Following the energy minimization, the minimized system was subjected to equilibration in two phases: NVT (constant Number of particles, Volume, and Temperature) and NPT (constant Number of particles, Pressure, and Temperature). The equilibration process was performed for a duration of 1000 picoseconds (ps). During the NVT phase, a position restraint molecular dynamics simulation was carried out, applying a constant force of 1000 kJ mol^{-1} nm^{-2} on the heavy atoms of both the protein and substrate. This restraint helped maintain the stability of the system and allowed for the adjustment of internal energies and atomic fluctuations. The NVT conditions ensured that the system's temperature remained constant throughout the equilibration process. After the equilibration phase, the well-equilibrated systems of AzrBmH2–dyes complexes were ready for the production run. The production run was carried out at a temperature of 300 K (Kelvin) and a pressure of 1 atmosphere for 100 ns (ns). During this production run, the dynamic behavior and stability of each AzrBmH2–dyes complex were analyzed using various parameters. The root mean square deviation (RMSD) was calculated to measure the deviation of each complex's structure from the starting conformation throughout the simulation. The radius of gyration (Rg) was determined to assess the compactness and overall size of the complexes during the simulation. The root means square fluctuation (RMSF) provided information about the flexibility and fluctuation of individual residues within the complexes. The solvent-accessible surface area (SASA) was calculated to inspect the exposure of the complexes to the surrounding solvent molecules. Furthermore, the hydrogen bond profile was analyzed to study the formation and stability of hydrogen bonds between the protein and ligand molecules. These analyses were performed using built-in tools in the Gromacs software, widely used for molecular dynamics simulations. By examining these parameters, we gained insights into the dynamic behavior, stability, and intermolecular interactions of the AzrBmH2–dyes complexes throughout the 100-ns production run.

Calculation of binding free energy

The binding free energies for the AzrBmH2–dye complexes were calculated using the molecular mechanics/Poisson Boltzmann surface area (MM-PBSA) approach. This analysis was performed using the g_mmpbsa tool in Gromacs. To determine the contribution of each residue to the binding free energy, we utilized the MmPbSaDecomp.py Python script. The g_mmpbsa package is a standalone tool that incorporates subroutines from the GROMACS and APBS packages, providing a detailed estimation of the molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) interaction. This allowed us to validate the results obtained from the molecular dynamics (MD) investigation. The binding free energy for each AzrBmH2–dye complex was calculated following the methods described in our previous studies [29], 2022c). The presented results are the average of triplicated MM-PBSA calculations.

Results and discussion

Ligand binding site prediction

The protein structures of the azoreductases (AzrBmH2) from *Bacillus megaterium* H2 were predicted using the DeepSite and PrankWeb tools. PrankWeb predicted a higher number of ligand-binding pockets compared to DeepSite. Specifically, AzrBmH21, AzrBmH22/3, and AzrBmH24/5 were predicted to have 7, 6, and 3 binding pockets by PrankWeb, respectively. DeepSite, on the other hand, predicted 6, 4, and 2 binding pockets for the three AzrBmH2 proteins. Considering the higher number of binding pockets and binding residues predicted by PrankWeb, the results obtained from this server were further analyzed. The binding pocket with the highest probability and pocket score, referred to as Pocket 1, was selected for validation. Pocket 1 exhibited a probability score of 0.929, 0.951, and 0.868, with corresponding pocket scores of 32.46, 37.82, and 22.95 for AzrBmH21, AzrBmH22/3, and AzrBmH24/5, respectively. The investigation revealed that Pocket 1 consisted of 15, 20, and 21 amino acids for AzrBmH21, AzrBmH22/3, and AzrBmH24/5, respectively. Complete details of the AzrBmH2 binding-site predictions are presented in Table 1 and Fig. 2.

The predictions of the ligand binding site of protein structure have many applications related to pollutants bioremediation. Several methods for detecting enzyme-binding cavities have been developed over the years based on the protein's structural, geometric, and chemical features [28]. In the present study, the prediction of binding active sites of azoreductase (AzrBmH2) proteins from *Bacillus megaterium* H2 was performed using machine learning algorithm-based deep convolutional neural networks (DeepSite and PrankWeb servers). The

two servers permit fast visualization of results and give valid predictions that other tools could not attain [10]. After analyzing Prank Web and Deep Site findings, the study noticed both servers agreed that the AzrBmH2 proteins have more than one perspective binding regions. Interestingly, the results for the machinery servers differ in the number of predicted regions. PrankWeb predicts more possible regions, whereas DeepSite predicts fewer regions. Upon closer examination of the outcomes, it is observed that the amino acids predicted by DeepSite largely overlap with those predicted by PrankWeb, albeit in different pockets. This agreement conveys confidence that the pockets in AzrBmH2 proteins are highly likely to be the predicted regions or pockets indicated by PrankWeb, as illustrated in Table 1 and Fig. 2.
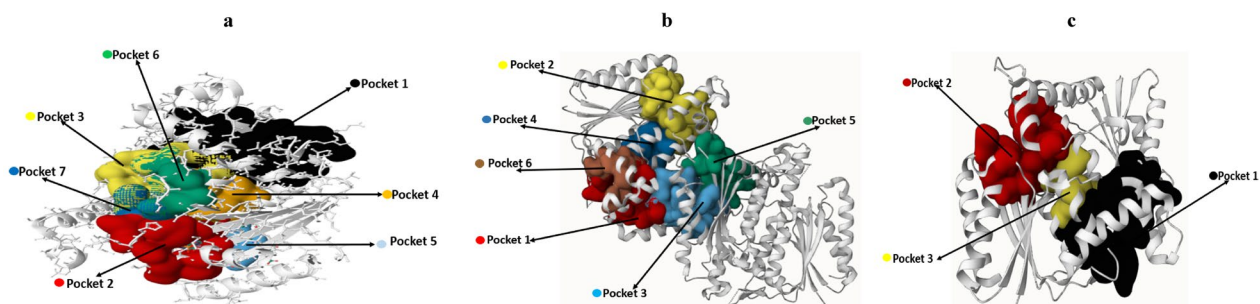
We then used the COACH-D algorithm and quantum tunnel prediction to reconfirm the active sites of the AzrBmH2 proteins (AzrBmH21, AzrBmH22/3, and AzrBmH24/5), guided by protocols reported by two similar studies [6, 27]. The result revealed five poses of the active site, with the predicted binding affinities of the best poses for the AzrBmH21, AzrBmH22/3, and AzrBmH24/5 proteins being − 7.1 kcal/mol, − 6.4 kcal/mol, and − 6.3 kcal/mol, respectively (Fig. 3). For authentication of the abovesaid result and perfect analysis by supermolecular docking, the fifth, first, and first binding poses for the corresponding AzrBmH21, AzrBmH22/3, and AzrBmH24/5 proteins were subjected to the Caver 3 webserver for quantum tunnelling of the best active site binding pose [15, 16]. As can be seen, the fifth active site position for the AzrBmH21 revealed six tunnels corresponding to a length and radius of 3.8 Å and 1.5–3.6 Å, respectively (Fig. 4a); whereas the 1st and 5th active site positions for the AzrBmH22/3 and AzrBmH24/5 proteins gave three and two tunnels with lengths and radii of 1.4 Å and 3.5 Å and ranging between 0.4–2.5 Å and 1.6–3.5 Å, respectively (Fig. 4b, c). The quantitative analysis revealed that AzrBmH21 has the highest radius (give value here) compared to two other proteins (Fig. 4), in which the radius and the tunnel lengths hinge on dependent and independent variables, respectively. The radius of a protein is valuable in deducing the ligand's length, height, and width [15].
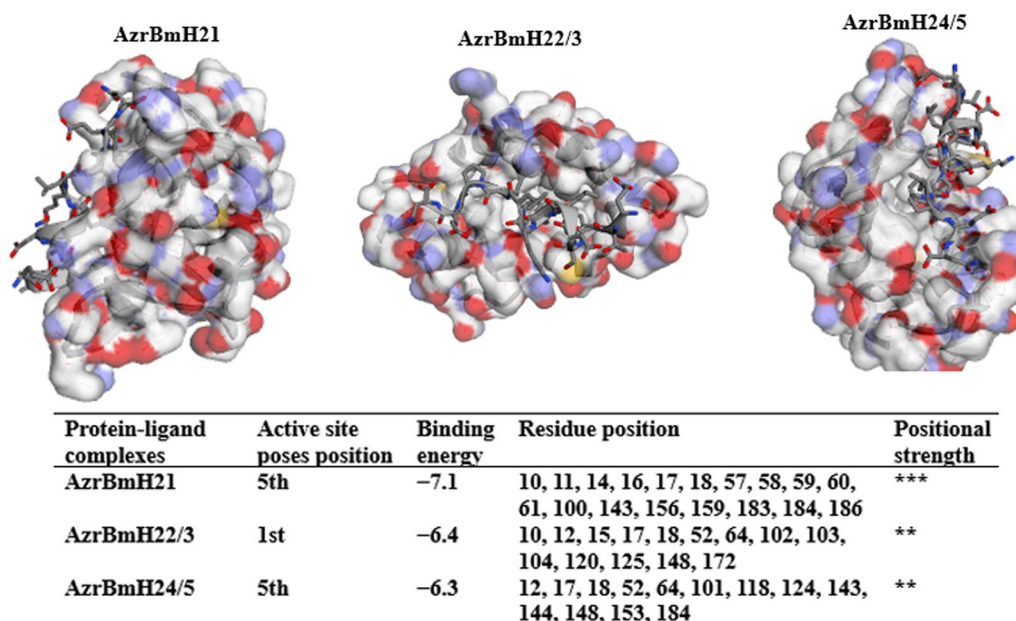
Molecular docking analysis

A bioinformatic analysis provides a detailed understanding of the substrate interaction and the amino acid residues involved in the interaction of AzrBmH2. In this study, molecular docking experiments were conducted to determine the binding affinity of AzrBmH2 and four different ligands (dyes) complexes. Molecular docking enables a comprehensive examination of the interaction patterns between the ligands and the protein, allowing

Table 1 Details of binding pocket as predicted by PrankWeb server

Pocket rank	AzrBmH21 (PDB: 6QU0)							AzrBmH22/3 (PDB: 6JXS)					AzrBmH24/5 (PDB: 3W77)			
	1	2	3	4	5	6	7	1	2	3	4	5	6	1	2	3
Pocket score	32.64	19.35	3.36	3.14	3.08	2.99	1.67	37.82	34.55	15.09	6.51	4.93	1.90	22.95	22.19	3.42
Probability score	0.929	0.829	0.123	0.109	1.105	0.099	0.027	0.951	0.939	0.75	0.356	0.233	0.037	0.868	0.860	0.127
Amino acid count	15	13	5	5	7	7	4	20	11	12	7	5	3	21	12	3
Amino acid residues	I52, D53, F57, S58, G59, W60, G61, D116, A119, Y125, T127, Q130, G131, P132, F172	H10,P11, E14,Q16, F18,P101, L102,W103, N104, A146, G145, N184, G186	H51, I52, P109, A113, D116	S106, F107, K112, M161, Y195	F105, G148, G149, I150, Y151, A156, M159	L44, Y45, I49, A51, P109, V110, A113	Y45, W103, P109, V110	W10, Q16, S17,L12, L18, K101, M102 W103, N104, F105, A146 Y148, D149, I150, Y151, A156, M159, Q186, V187, A188	I52, F57, W60, D116, T119, A121, F125, G131, A132, N171, F172	Y6, W10, P11, L12, D13, E14, M20, K24, H41, D43, Y45, K46	K3, I40, E47, R82, E85, Q89, A92	S106, F107, K112, M161, Y165	L44, Y45, K47	N10, R12, A17, Q16, V18, S19, P98, L99, W100, N101, L102, A143, R144, G145, G146, Y148, A153, V156, H183, N184, Q185	V52, M56, I57, Q60, D113, N116, F122, Y124, P129, S165, M166, F169	T103, H109, Y162

Fig. 2 PrankWeb prediction of binding active sites of AzrBmH2 proteins from *Bacillus megaterium* H2 **a** AzrBmH21 (PDB 6QU0), **b** AzrBmH22/3(PDB 6JXS), and **c** AzrBmH24/5 (PDB 3W77)

| Protein–ligand complexes	Active site poses position	Binding energy	Residue position	Positional strength
AzrBmH21	5th	−7.1	10, 11, 14, 16, 17, 18, 57, 58, 59, 60, 61, 100, 143, 156, 159, 183, 184, 186	***
AzrBmH22/3	1st	−6.4	10, 12, 15, 17, 18, 52, 64, 102, 103, 104, 120, 125, 148, 172	**
AzrBmH24/5	5th	−6.3	12, 17, 18, 52, 64, 101, 118, 124, 143, 144, 148, 153, 184	**

Fig. 3 The best active site poses of AzrBmH2 proteins predicted by COACH-D by considering the binding energy (kcal/mol) and the number of the amino acids involved at the docking region

for calculating their binding energy. A suitable algorithm can identify the specific amino acid residues in contact with the evaluated substrates, further enhancing our understanding of the ligand–protein interaction. The present study carried out molecular docking experiments to obtain the binding energy or binding affinity for the complexes formed between AzrBmH2 and the four different ligands (dyes). Docking results are either considered energetically unfavorable when binding energy for AzrBmH2–ligands complex ≥ 0 kcal/mol pointing extremely high or complete absence of affinity. Herein, the overall molecular docking results are shown in Table 2 and Fig. 5.

The results of molecular docking revealed that inter-atomic interactions, such as hydrogen bonding and hydrophobic interactions, occurred between the AzrBmH2 proteins and the ligands (dyes). The observed binding energies were between − 9.4 and − 5.5 kcal/mol, − 9.2 to − 5.4, and − 9.0 to − 5.4 for AzrBmH21, AzrBmH22/3, AzrBmH24/5, respectively, whereby each complex was stabilized by at least 0–5 hydrogen bonds (Table 2). The lowest binding energy reflects the formation of a stronger AzrBmH2–ligand complex interaction, wherein AzrBmH21, AzrBmH22/3 and AzrBmH24/5 complexes with cresol red yielded the lowest binding energy of − 9.4, − 9.2 and − 9.0 kcal/mol, respectively. This corresponds to four hydrogen bonds for AzrBmH21–cresol red complex with distances 3.08 Å, 2.80 Å, 2.91 Å and 2.65 Å, that link to Ala143, Gly145, Asn148 and Asp53, respectively (Table 2). At
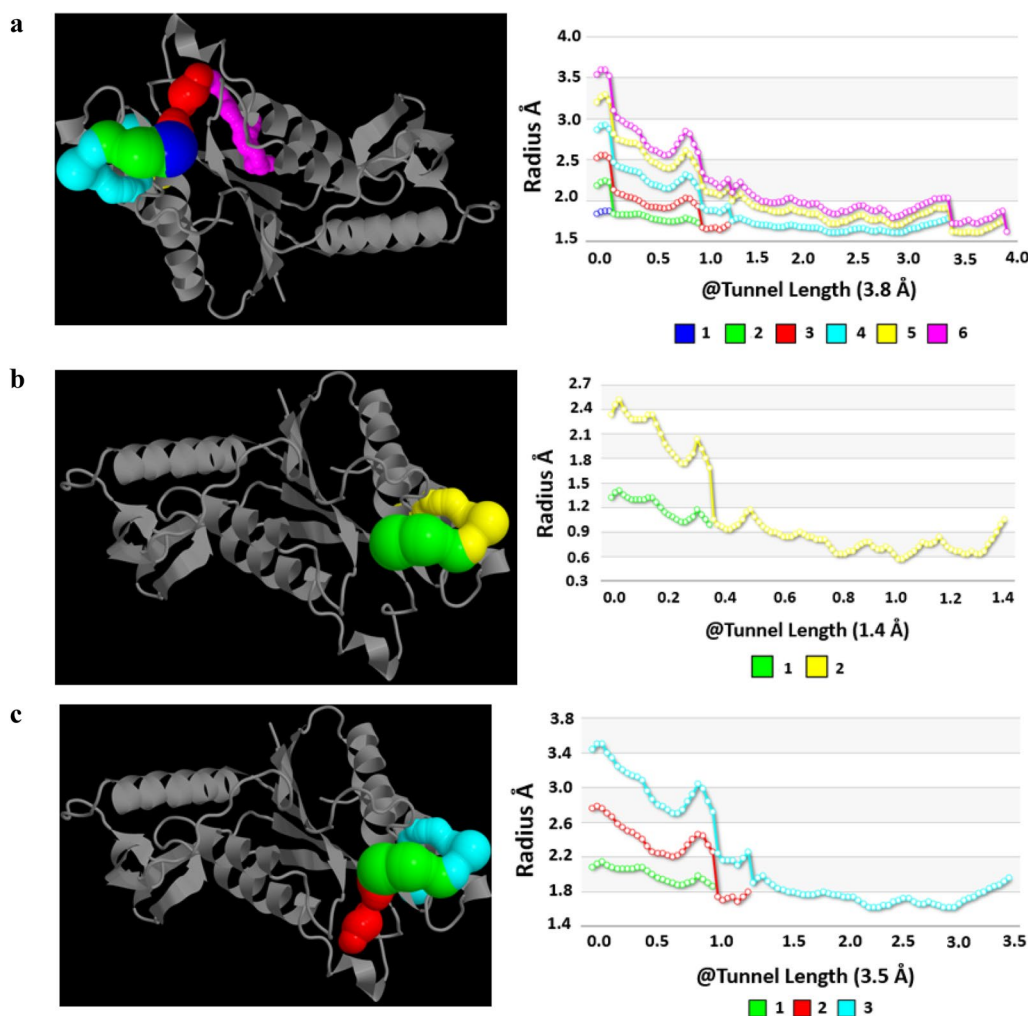
Fig. 4 Protein tunnels demonstration along with the tunnel length (Å) and radius (Å) of the three models of the best tunnels for allocating any ligand in the docking position

the same time, the AzrBmH22/3–cresol red complex formed two hydrogen bonds at 2.80 Å and 2.82 Å with residues Asn104 and Asp149, respectively. AzrBmH24/5-cresol red complexes formed two hydrogen bonds, at 2.80 Å and 2.90 Å with Arg12 and Gln117, respectively (Table 2). Similar results were observed when docking the AzrBmH2 proteins with acid orange 7, and methylene blue, yielding comparable binding energy between − 9.1 and 7.0 kcal/mol (Table 2). However, all the AzrBmH2–malachite green complexes showed binding energies between − 5.5 and − 5.4 kcal/mol with no intermolecular hydrogen bonds formation (Table 2, Fig. 5).

As can be seen in Fig. 3, AzrBmH21–malachite green complexes showed hydrophobic bonding with 10 amino acids (Asn184, Ala143, His10, Trp100, Ser58, His183, Phe18, Trp60, Phe57, Gly59); whereas, the AzrBmH22/3–malachite green complexes produced

nine hydrophobic bonds through Ser121, Arg64, Trp10, Val120, Thr119, Leu52, Trp103, Asn104, Phe125. Similarly, the same complex formed nine hydrophobic bonds with Tyr124, Pro129, Tyr148, Asn116, Asn101, Phe169, Phe102, Ser153, and Pro152 (Fig. 5). Conversely, the negative binding energy observed for the AzrBmH2–malachite green complexes was likely due to hydrophobic interactions.

Generally, the binding energy observed for AzrBmH2, and ligands (methylene blue, cresol red and acid orange 7) was the most negative interaction energy (strong interaction) compared with the remaining AzrBmH2–malachite green conformation complexes maybe due to the high number of hydrogen and hydrophobic interactions associated to predicted ligand binding residues (Asp53, Gly61, Gly145, Asn184, His10, Glu14, Gln16, Phe18, Phe57, Ser58, Trp60, Met102, Tyr151, Asn104, Tyr148,

Table 2 Summary of docking analysis of AzrBmH2 proteins with various dyes substrates from AutoDock Vina scores

Pollutants	AzrBmH21				AzrBmH22/3				AzrBmH24/5			
	Binding energy (kcal/mol)	Hydrogen bonds	Hydrogen bond distance	Hydrophobic interactions	Binding energy (kcal/mol)	Hydrogen bonds	Hydrogen bond distance	Hydrophobic interactions	Binding energy (kcal/mol)	Hydrogen bonds	Hydrogen bond distance	Hydrophobic interactions
Methylene blue (MB)	− 7.7	Asn184	3.23 Å	Phe18, Phe57, His138, Trp60, Gly61, Trp100, Met99, Ser58	− 7.3	Met102	2.95 Å	Leu18, Ser17, Leu12, Trp103, His10, Asn104, Gln16	− 7.0	Tyr148	2.99 Å	Asn101, Asn184, Gly145, Ala118
Malachite green (MG)	− 5.5	–	–	Asn184, Ser58, Ala143, Phe18, His10, Trp100, His183, Trp60, Phe57, Gly59	− 5.4	–	–	Ser121, Arg64, Trp10, Val120, Thr119, Leu52, Trp103, Asn104, Phe125	− 5.4	–	–	Tyr124, Pro129, Tyr148, Asn116 Asn101, Phe169, Phe102, Ser153, Pro152
Acid orange 7 (AO7)	− 9.4	Met99 Gly145 Asn184 Gly61	3.03 Å 3.20 Å 2.88 Å 2.92 Å	His10, Phe18 Arg144, Glu14 Trp100, Gln16 His183, Ser17,	− 9.2	Tyr151 Arg147 Asn104	2.76 Å 3.04 Å 2.89 Å	Phe105, Thr119, Ser121, Trp103, Phe172	− 9.0	Val18 Arg12 Ser19 Asn101 Leu99	3.26 Å 3.05 Å 3.17 Å 3.10 Å 2.97 Å	Asn184, Gly145, Try148, Arg144, Ala143, Ala17
Cresol red (CR)	− 9.1	Ala143 Gly145 Asn184 Asp53	3.04 Å 2.80 Å 2.91 Å 2.65 Å	Arg144, His10, Trp100, Pro11, Glu14, Met99, Gln16	− 8.9	Asn104 Asp149	2.80 Å 2.82 Å	Tyr151, Ala132, Tyr127, Val187, Phe125, Arg147, Gly148, Phe172	− 8.9	Arg12 Gln117	2.80 Å 2.90 Å	Gln60, Asn184, Leu52, Ala64, Ile57, Phe61

Amino acids highlighted in yellow are ligand or catalytic binding sites predicted by the PrankWeb server
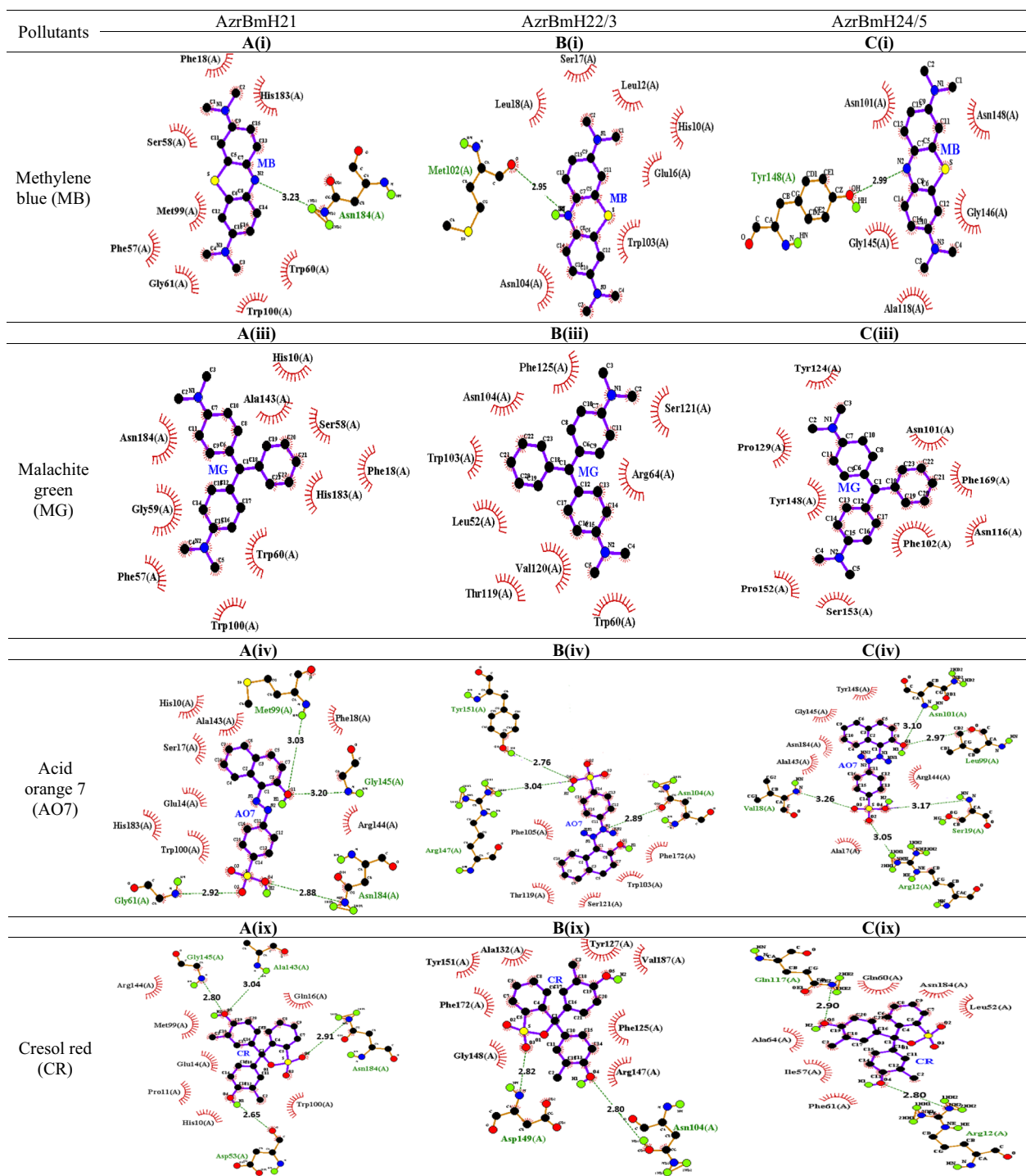
Fig. 5 LigPlot analysis for AzrBmH2–ligand interaction showing hydrogen bond with its equivalent distance and non-ligand residues involved in the hydrophobic interaction. **A** AzrBmH21–ligands, **B** AzrBmH22/3–ligands, **C** AzrBmH24/5–ligands interaction (i) methylene blue (MB), (ii) malachite green (MG), (iii) acid orange 7 (AO7), (iv) cresol red (CR)

Leu12, Leu18, Trp103, Ser121, Leu52, Phe125, Phe105, Phe172, Ala132, Val187, Gly148, Tyr148, Val18, Arg12, Ser19, Leu99, Asn101, Asn116, Tyr124, Phe102, Ser153, Arg144, Ala143, Ala17, Ile57) prophesized by PrankWeb (Table 1). However, the weaker interactions of malachite green with all three AzrBmH2 proteins might be due to

only hydrophobic interactions associated to predicted ligand AzrBmH2-residues like Asn184, Ala143, His10, Trp100, Ser58, His183, Phe18, Trp60, Phe57, Gly59, Ser121, Arg64, Trp10, Val120, Thr119, Leu52, Trp103, Asn104, Phe125, Tyr124, Pro129, Tyr148, Asn116, Asn101, Phe169, Phe102, Ser153 and Pro152 (Table 2). It is worth mentioning here that most of these amino acids are predicted to be catalytic by the PrankWeb server (Table 1).

The study was performed to validate the predicted active site and confirm the specific residues that play a crucial role in the catalytic activity of the AzrBmH2 enzyme when interacting with different synthetic dyes. By examining the enzyme's active site and analyzing its interaction with the dyes, the study sought to provide experimental evidence and verification of the residues responsible for the enzyme's catalytic function. The docking result of AzrBmH2 enzymes against cresol red, acid orange 7 and methylene blue was analyzed using LigPlot as shown in Table 2 and Fig. 3. The docking results for AzrBmH2–ligand complexes were considered energetically favorable since the binding energy ≤ 0 kcal/mol pointed to the affinity and interaction between the protein and ligand. On comparing the docking result with the result obtained by the PrankWeb server, several amino acid residues were found to be similar to hydrogen bonds and hydrophobic interactions. This indicated that the substrates were binding at the predicted active site. Commonly found amino acid residues in bacteria azoreductases were Pro, Phe, Gly, Asn, Ser, Lys Ile, Thr and Tyr [18, 31]. In our study, these amino acids were also present in each binding interaction. These residues were commonly in the binding interactions with all the model compounds and some residues are represented in our study.

Molecular dynamics simulation

Molecular dynamics (MD) simulation is a powerful method/tool that aid bioremediation strategies by anticipating the protein–ligand complex's structural changes, stability, and flexibility for an efficient technology transfer to real-time set-up [13, 23]. Remarkably, several properties, such as the chemical nature of pollutants, biodegradation pathways, and microorganisms' bioremediation abilities at the environmental levels, can be predicted by the in silico approach [13]. In this study, structural changes, stability, and flexibility of the AzrBmH2 proteins (AzrBmH21, AzrBmH22/3 and AzrBmH24/5) with respect to the ligands/substrates (acid orange 7, methylene blue, cresol red and malachite green) were revealed to predict their bioremediation potential. MD simulation of AzrBmH2–ligands complexes was performed for 100 ns using Gromacs software. root mean square

deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA) and hydrogen bond plots were generated for an in-depth analysis.

Root mean square deviation (RMSD)

Specifically, the minimum value of the (RMSD $\sim 0.2-\sim 0.3$ nm) means the protein complex is stable [34, 40]. As can be seen in Fig. 4 the structures of the AzrBmH21–ligand complexes are highly stable and lies between 0.15 and 0.3 nm, followed by AzrBmH24/5–ligand complexes (between 0.15 and 0.39 nm) and AzrBmH22/3–ligands complexes (between 0.2 and 0.42 nm). As revealed in Fig. 6a, the RMSD of the AzrBmH21–acid orange 7 and cresol red complexes achieved stability after 20 ns with an RMSD of 0.21 nm. A similar result was observed for the AzrBmH21–methylene blue complex with an average RMSD of 0.24 nm. From 0 to 90 ns, excessive variation in RMSD of the AzrBmH21–malachite green complex was observed due to extreme deviation of the ligand from its initial position as shown in Fig. 6a. The AzrBmH22/3 showed a similar stability and conformation change pattern with ligands during the simulation (Fig. 6b). After 49 ns, the ligands bind to the enzyme (AzrBmH22/3), and a stable conformation is achieved with an average RMSD of 0.29 and 0.3 nm for acid orange 7 and cresol red, respectively. Meanwhile, methylene blue and malachite green–enzyme complexes fluctuated to 0.32 nm (Fig. 6b). On analyzing the MD simulation trajectory of the AzrBmH24/5–ligand complexes, it was seen that AzrBmH24/5–methylene blue and acid orange 7 complexes were stabilized at 48 ns, corresponding to an average RMSD of 0.2 nm. The AzrBmH24/5–cresol red and malachite green complexes stabilized after 52 and 71 ns, averaging RMSDs of 0.25 and 0.3, respectively (Fig. 6c). It is worth stating here that the RMSD for all the AzrBmH2–ligand complexes was less than or equal to 0.3 nm. This showed that the ligands were present within the binding cleft of the proteins during the simulation time, which was considered based on the docking result.

Root mean square fluctuation (RMSF)

RMSF measures the displacement of amino acid residues during a given simulation time. By analyzing the RMSF, the highly flexible residue can be detected. The RMSD investigations of the AzrBmH2–ligand complex structures were found to be stable with minor changes. Generally, the RMSF plot revealed that few amino acid residues moved extensively (Fig. 7). As can be seen, the RMSF plots of AzrBmH21–ligand complexes showed that Gln16 and 68 amino acid residues were highly fluctuating, while residues, Leu18, Asp71, Ser73, Thr74 and
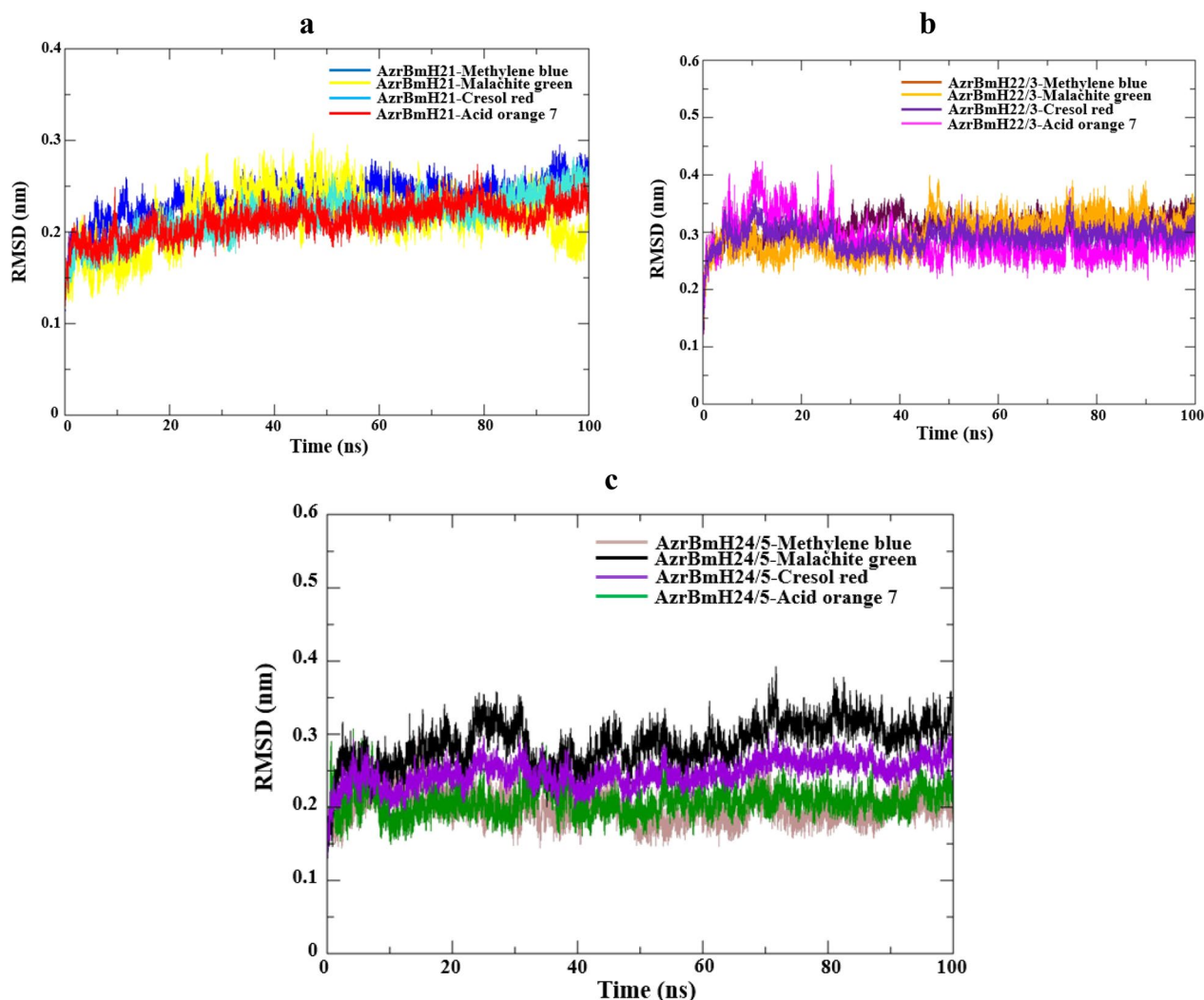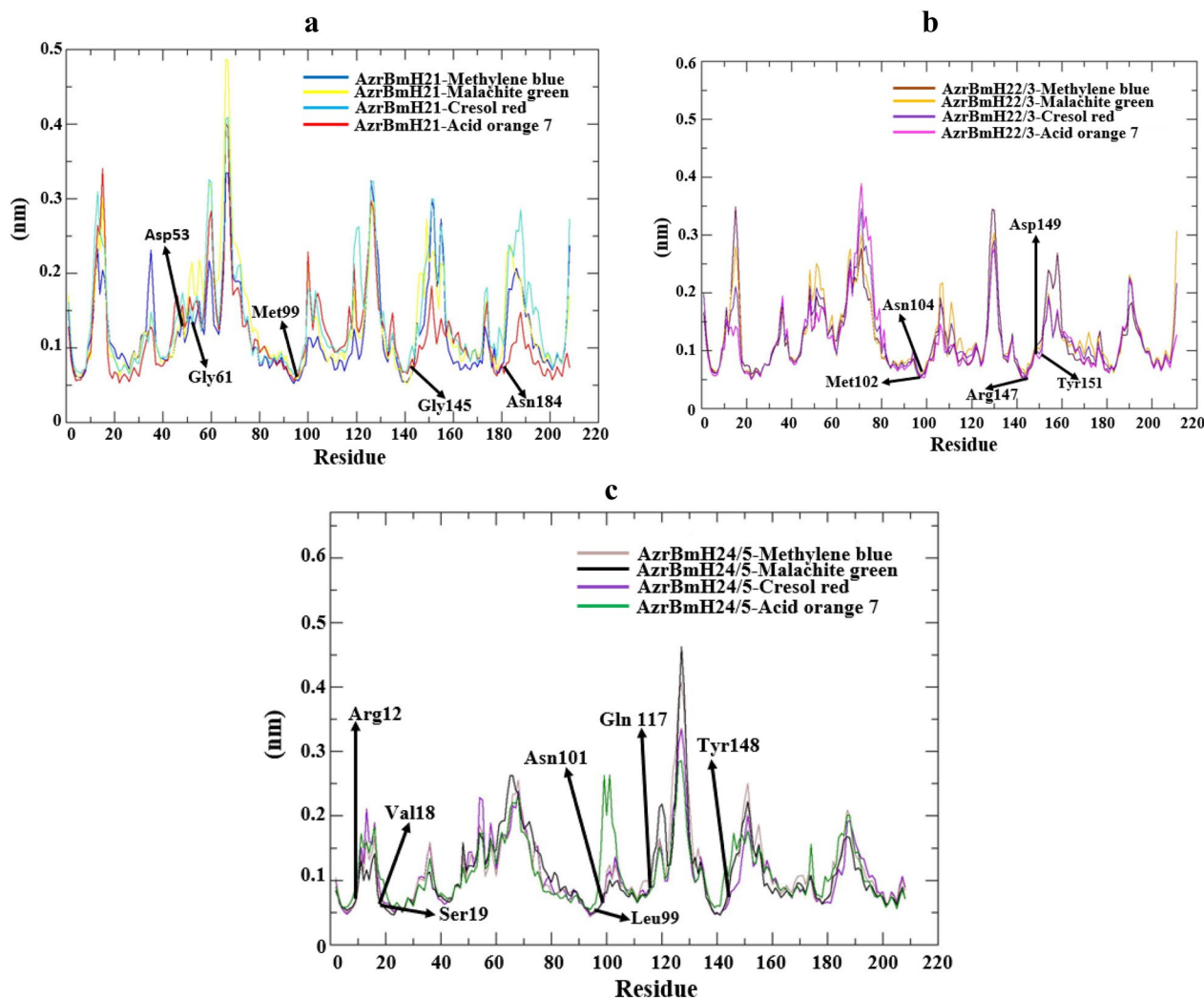
Fig. 6 MD simulation of AzrBmH2 proteins against the various synthetic dye. The average RMSD of the backbone conformation is shown as a function of simulation time (100 ns) at 300 K, and the overlayed RMSD plots of **a** AzrBmH21 of methylene blue in blue, malachite green in yellow, acid orange in red, cresol red in turquoise; **b** AzrBmH22/3 of methylene blue in orange, malachite green in indigo, acid orange in maroon, cresol red in magenta; **c** AzrBmH24/5 of methylene blue in brown, malachite green in black, acid orange in green, cresol red in violet

Thr119 behaving similarly in the AzrBmH22/3–ligand complexes. Likewise, Pro129 in AzrBmH24/5–ligand complexes exhibit an RMSF value that exceeded 0.3 nm (Fig. 7). The reason behind the little fluctuations seen in the RMSF of the AzrBmH2–ligand complexes might be the attachment of the dyes to the AzrBmH2 proteins, which led to the increase in the movement of amino acid residues for the preparation of the catalytic process. From the RMSF plot, it was observed that predicted catalytic residues Met99, Gly145, Asn184, Gly61, and Asp53 for AzrBmH21–ligand complexes (Fig. 7a), Tyr151, Arg147, Asn104, Met102 and Asp149 for AzrBmH22/3–ligand complexes (Fig. 7b) and Val18, Arg12, Ser19, Asn101, Leu99, Gln 117 and Tyr148 for

AzrBmH2–ligand complexes (Fig. 7c) showed fluctuations below 0.3 nm. All these amino acid residues were found in the loop of AzrBmH2 protein structures. This supports the fact that catalytic residues or binding sites are buried in an enzyme's loop or inner core, known for their dynamic binding ability [4]. Similarly, the amino acid residues involved in hydrophobic interactions exhibited low fluctuation, ranging from 0.05 to 0.3 nm, across all three AzrBmH2 protein structures (Fig. 7).

The RMSF detailed the protein flexibility and structural fluctuations for each residue contributing to the protein dynamics [12]. The threshold range for RMSF was between 0.05 and 0.3 nm, determining the significant changes in residue-specific flexibility that

Fig. 7 MD simulation of AzrBmH2 proteins against the various synthetic dyes. The average RMSF of the backbone conformation is shown as a function of simulation time (100 ns) at 300 K, and the overlayed RMSF plots of **a** AzrBmH21 of methylene blue in blue, malachite green in yellow, acid orange in red, cresol red in turquoise; **b** AzrBmH22/3 of methylene blue in orange, malachite green in indigo, acid orange in maroon, cresol red in magenta; **c** AzrBmH24/5 of methylene blue in brown, malachite green in black, acid orange in green, cresol red in violet

gauges an enzyme's substrate specificity [41]. Therefore, a high RMSF value represents a higher degree of fluctuation, while a low value implies limited structural movement and a more stable structure [39]. The AzrBmH2 proteins exhibited low RMSF values for the predicted catalytic or binding site residues, indicating strong molecular interactions with the dyes (ligands). The study observed that the four dyes bonded with the AzrBmH2 proteins in distinct patterns and remained stable during the MD simulations. Additionally, the differences in binding energies provide evidence of the substrate specificity of AzrBmH2.

Radius of gyration (Rg)

Rg is the root mean square (RMS) distance of the protein's atom and ligand's atoms from the common center of gravity [36]. To analyze the overall shape of AzrBmH2–ligand complexes, the Rg plot was examined. Figure 8 illustrates that the Rg values for AzrBmH21–ligands, AzrBmH22/3–ligands, and AzrBmH24/5–ligand complexes range from 1.75 to 1.83 nm, 1.78 to 1.88 nm, and 1.75 to 1.85 nm, respectively. Specifically, in the case of AzrBmH21–ligand complexes, the average Rg value for AzrBmH21–methylene blue starts at 1.8 nm and decreases to 1.76 nm in
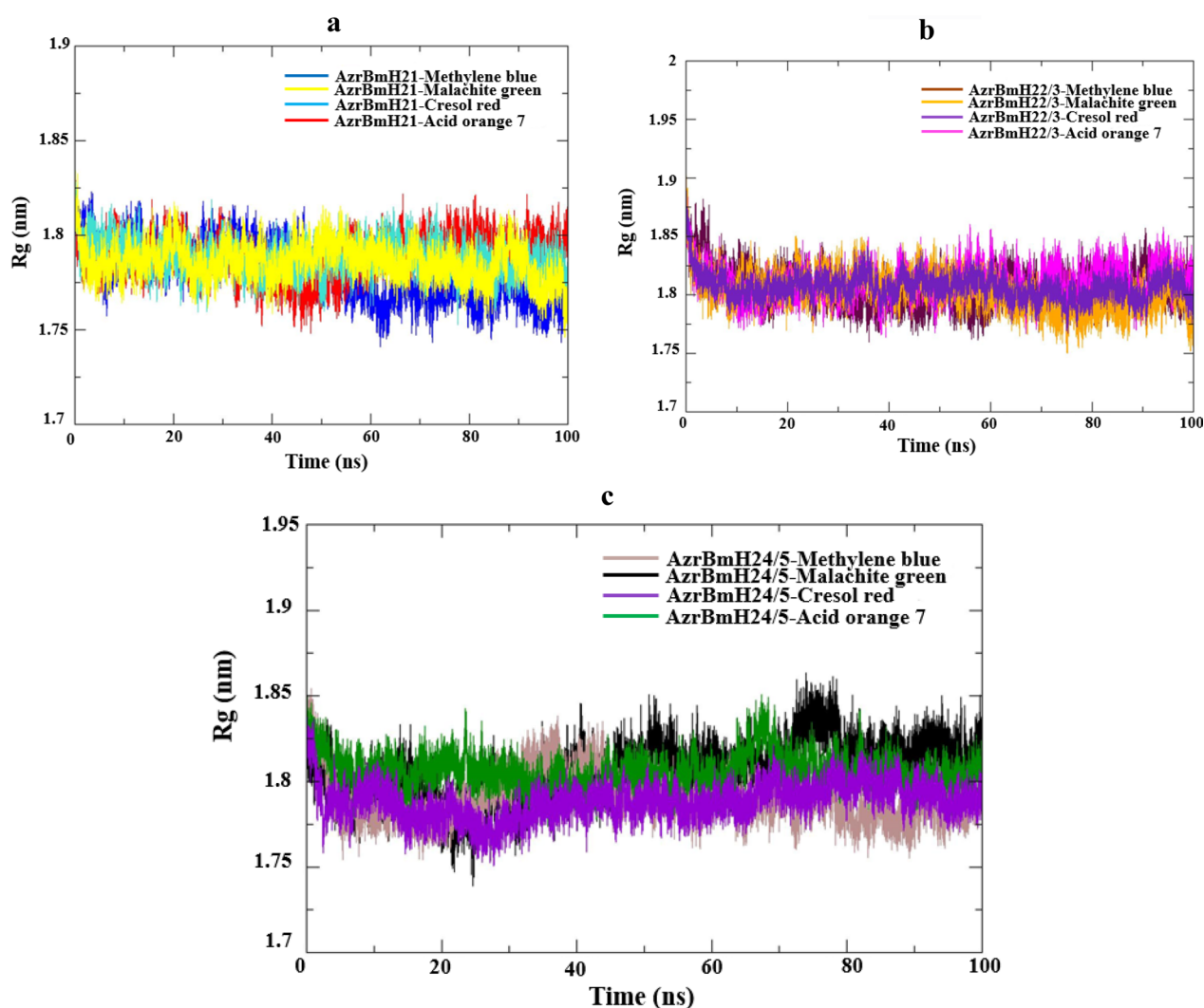
Fig. 8 Average is shown as a function of 100 ns simulation time at 300 K, showing AzrBmH2 against dyes as in **a** AzrBmH21 of methylene blue in blue, malachite green in yellow, acid orange in red, cresol red in turquoise; **b** AzrBmH22/3 of methylene blue in orange, malachite green in indigo, acid orange in maroon, cresol red in magenta; **c** AzrBmH24/5 of methylene blue in brown, malachite green in black, acid orange in green, cresol red in violet showing all radius of gyration (Rg) plotted together

the final 57 ns of the simulation (Fig. 8a). Conversely, AzrBmH21 with other ligands shows early stabilization with an average Rg of 1.79 nm from the beginning of the simulation (Fig. 8a). The compactness of AzrBmH22/3–ligand complexes during the simulation indicated that all ligands were more compact, with the AzrBmH22/3 stabilizing at 9 ns, with an average Rg of 1.82 nm (Fig. 8b). As for AzrBmH24/5 with cresol red, methylene blue, and acid orange 7, their Rg values stabilize at 37 ns, 58 ns, and 45 ns, respectively, with average Rg values of 1.77 nm, 1.79 nm, and 1.81 nm. However, a spike was observed at 78 ns for methylene blue (Fig. 8c). AzrBmH24/5–malachite green exhibits stability starting at 60 ns, with an average Rg of

1.84 nm, and spikes between 73 and 80 ns are observed (Fig. 8c).

This study's diverse range of Rg values indicates that the AzrBmH2–ligand complexes exhibited varying compactness and overall structural dimensions during the molecular dynamics (MD) simulation. It is important to highlight that a relatively constant Rg value indicates a stable folded structure and strong ligand interaction with the binding site. Conversely, an unfolded structure or a protein with looser packing is characterized by fluctuating or higher Rg values throughout the simulation [21, 25]. The Rg results indicate that the ligands successfully folded and bound to the predicted binding site of AzrBmH2, resulting in increased compactness of the

AzrBmH2–ligand complexes. The enhanced stability and binding of the complexes contribute to their compact nature.

Hydrogen bonds analysis

H-bonds are essential for the stability of biomolecular complexes. During the 100 ns MD simulations, the gmx hbond utility tool was employed to calculate the number of discrete hydrogen bonds formed between specific amino acid residues of the AzrBmH2 core proteins and dye atoms. Hydrogen bond interactions with the corresponding ligands were assessed using the evolution plot (Fig. 9). The present study showed a maximum

of six to zero H-bonds formed by the AzrBmH2–ligand complexes, and some of the H-bonds remained fluctuating till the end of the simulation (Fig. 9). As can be seen in Fig. 5a, the AzrBmH21–acid orange 7 revealed 6 H-bond, followed by AzrBmH21–cresol red and AzrBmH21–methylene blue complexes with 5 and 4 H-bonds, respectively. However, the AzrBmH21–malachite green complex showed no H-bond (Fig. 9a). Similar results were obtained for and AzrBmH22/3–ligand complexes (Fig. 9b). AzrBmH22/3–acid orange 7 complex exhibited 6 H-bonds while AzrBmH22/3–cresol red, AzrBmH22/3–methylene blue and AzrBmH22/3–malachite green showed fewer H-bond numbers with a
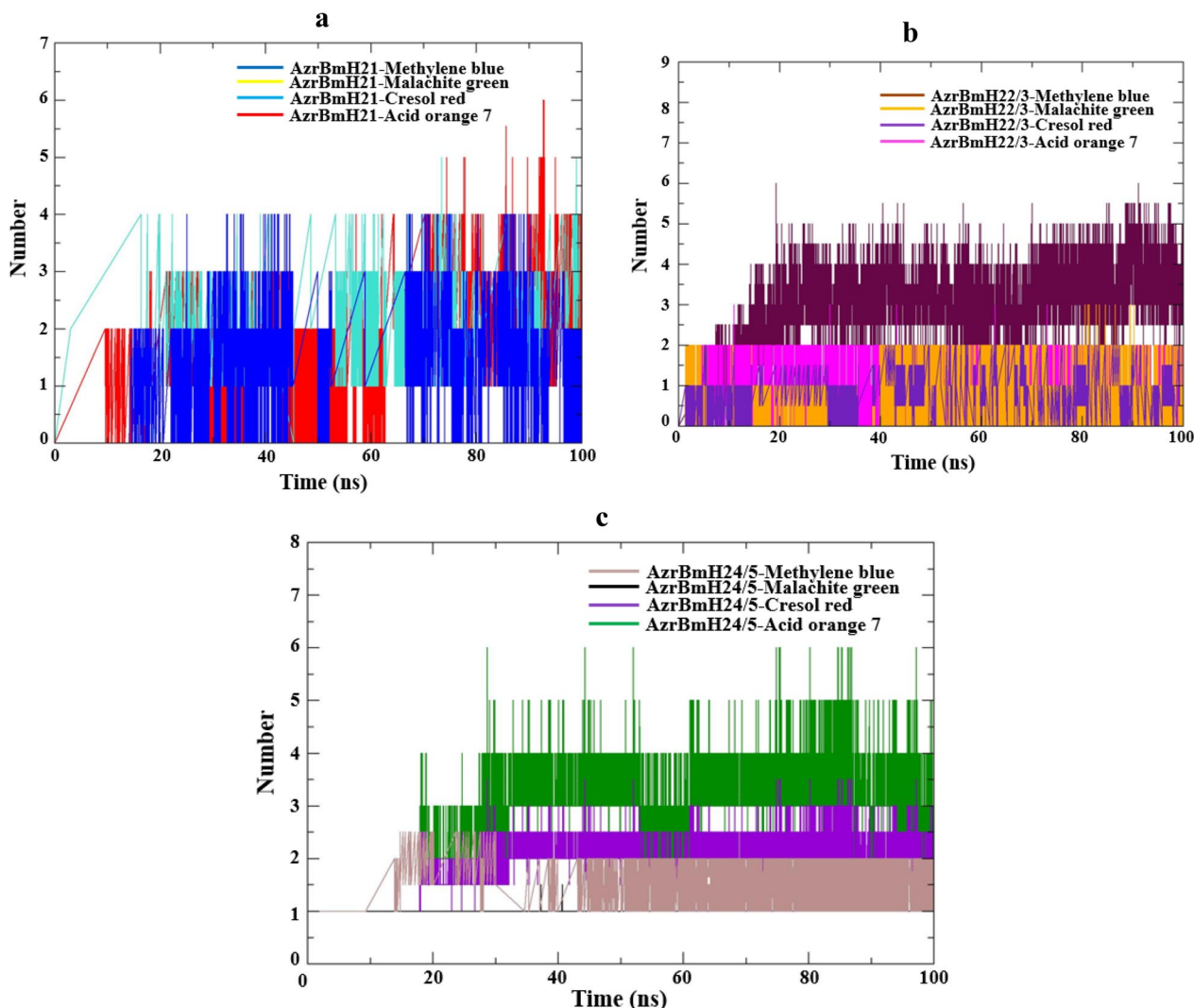
Fig. 9 Intermolecular hydrogen bonds along 100 ns simulation time in AzrBmH2–dyes when chemically bounded **a** AzrBmH21 of methylene blue in blue, malachite green in yellow, acid orange in red, cresol red in turquoise; **b** AzrBmH22/3 of methylene blue in orange, malachite green in indigo, acid orange in maroon, cresol red in magenta; **c** AzrBmH24/5 of methylene blue in brown, malachite green in black, acid orange in green, cresol red in violet showing all hydrogen bonds plotted together

Oyewusi *et al. Environmental Sciences Europe*   (2024) 36:31

Page 16 of 20

maximum of 3, 2 and 2 H-bonds, respectively (Fig. 9b). Comparatively, AzrBmH24/5−acid orange 7 formed a maximum of 5 H-bonds, followed by AzrBmH24/5−cresol red, AzrBmH24/5−methylene blue and AzrBmH24/5−malachite green complexes, corresponding to three, one and the absence of H-bonds (Fig. 9c).

The higher affinity of these molecules was likely attributed to the formation of hydrogen bonds, which supported better degradation of the dyes. The analysis revealed that the dyes formed 0–6 hydrogen bonds with the AzrBmH2 proteins (AzrBmH21, AzrBmH23/4, AzrBmH24/5) throughout the simulation period (Fig. 9). Notably, these findings are consistent with the molecular docking analysis of AzrBmH2−substrate interactions, as the MD simulation results exhibited a relatively higher number of hydrogen bonds than the molecular docking experiment. This suggests that the MD simulation of the AzrBmH2−dye complexes yielded structurally more compact and stable conformations, thus confirming the strong stability and binding of the four synthetic dyes with the active site of the AzrBmH2 proteins (AzrBmH21, AzrBmH23/4, AzrBmH24/5).

Solvent-accessible surface area (SASA)

Furthermore, the solvent-accessible surface area (SASA) refers to the portion of the protein sufficiently exposed to interact with surrounding solvent molecules. SASA plays a crucial role in studying protein stability and folding, and it is defined by the hypothetical center of the solvent sphere in contact with the molecule's surface through van der Waals interactions [18]. Solvent-accessible surface area (SASA) was also analyzed for the proteins in complex with different dyes. Figure 10 depicts the estimated surface area around the AzrBmH21−dyes, AzrBmH22/3−dyes, AzrBmH24/5−dyes complexes (Fig. 10). The total area of the AzrBmH21−dyes was ~ 118–124 nm^2 throughout the 100 ns (Fig. 10a). The surface area of AzrBmH22/3−dyes and AzrBmH24/5−dyes complexes were about 119–127 nm^2 and 115–125 nm^2, respectively (Fig. 10b, c). Generally, an increased protein SASA value during simulation indicates structural relaxation and reduced protein stability [5]. In summary, the comprehensive analysis of dynamic stability and hydrogen bonding in the molecular docking and MD simulation of AzrBmH2−dye complex binding demonstrates the good stability and capability of AzrBmH2 proteins to degrade the tested dyes. Hence, our findings collectively support the hypothesis that AzrBmH2 proteins exhibit favorable interactions and possess the potential for effective dye degradation.

MM-PBSA binding free energy calculation

To investigate the degradation and decolorization mechanisms of the tested dye molecules at the atomic level, the MM-PBSA method was employed to calculate the binding free energies between AzrBmH21, AzrBmH22/3, AzrBmH24/5, and the dyes (acid orange 7, cresol red, methylene blue, and malachite green). The components of the molecular mechanics computed by the MM-PBSA analysis are presented in Table 3. The binding free energies for AzrBmH21, AzrBmH22/3, and AzrBmH24/5 with the dyes ranged from − 300.22 ± 11.69 to − 35.07 ± 5.36, − 314.19 ± 6.58 to − 65.26 ± 8.81, and − 233.04 ± 6.39 to − 37.97 ± 2.94 kcal/mol, respectively. The energy components of the binding free energies obtained in this study are listed in Table 3. It can be observed that hydrogen bonds were the major favorable contributors to the degradation of the dyes, except for malachite green, where van der Waals and electrostatic energies played important roles in the binding process. Notably, the van der Waals energy in the AzrBmH2−malachite green complex was almost six times stronger than the hydrogen bond energy. Hence, hydrogen bond energies predominantly drove the binding of acid orange 7, cresol red, and methylene blue to AzrBmH2 proteins (Table 3), while van der Waals energies were primarily responsible for the binding the AzrBmH2 proteins to malachite green (Table 3).

The abovesaid result was further substantiated by calculating the binding free energy using the MM-PBSA method. The MM-PBSA analysis determined the binding free energy, including van der Waals, electrostatic, polar solvation, and non-polar solvation energies, to predict the strength of binding between the AzrBmH2 protein models and the tested dyes (as shown in Table 3). It is worth noting that hydrogen bond energy plays a crucial role in the degradation of acid orange 7, cresol red, and methylene blue, while van der Waals interactions contribute to malachite green degradation. These results are in alignment with the previous analyses, providing further support to the overall findings [2, 8, 33]. This result indicates that optimizing hydrogen bond interactions between acid orange 7, cresol red, and methylene blue and AzrBmH2 proteins, as well as the der Waals interactions between malachite green and AzrBmH2 proteins, augment AzrBmH2−dyes' degradation or decolorization.

Conclusion

Naturally occurring synthetic or commercially produced dyes are known to be toxic and carcinogenic, contributing to environmental pollution. Furthermore, these dyes are resistant to biodegradation and persist as pollutants in water bodies. The current study investigated
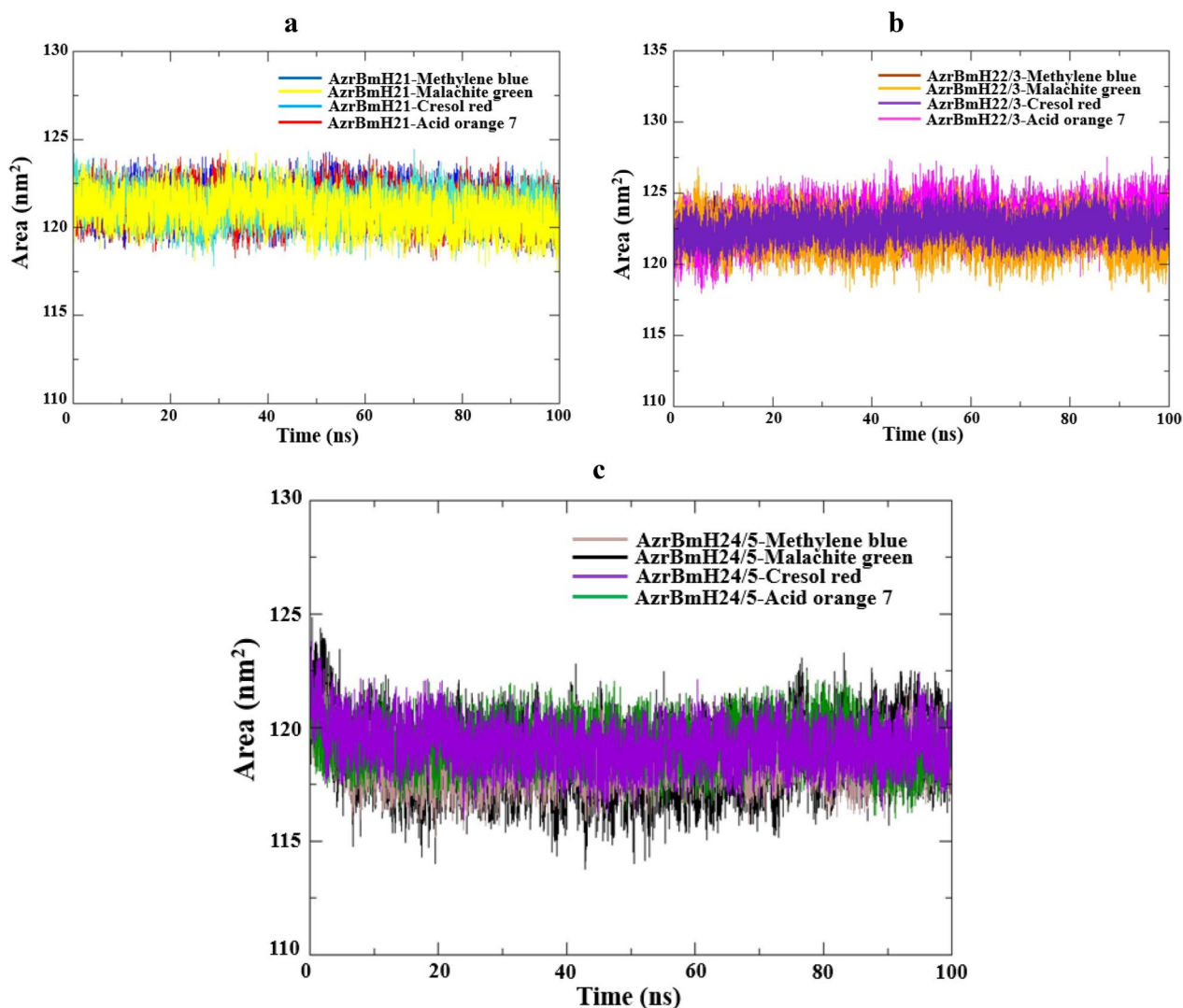
Fig. 10 Solvent-accessible surface area (SASA) analysis along 100 ns simulation time in AzrBmH2–dyes of **a** AzrBmH21 of methylene blue in blue, malachite green in yellow, acid orange in red, cresol red in turquoise; **b** AzrBmH22/3 of methylene blue in orange, malachite green in indigo, acid orange in maroon, cresol red in magenta; **c** AzrBmH24/5 of methylene blue in brown, malachite green in black, acid orange in green, cresol red in violet showing all SASA plotted together

the interaction between four commercial dyes by the *Bacillus megaterium* H2 azoreductases using in silico analysis, which utilized the ligand binding site modeling, molecular docking, and molecular dynamics simulation. The docking interactions revealed the potential relevance of azoreductases (AzrBmH2: AzrBmH21, AzrBmH22/3, and AzrBmH24/5) with the four dye compounds (acid orange 7, cresol red, methylene blue, and malachite green) in bioremediation. The binding affinity of the AzrBmH2–dye docking complexes followed the order: AzrBmH21–dye complexes (-9.4 to -5.5 kcal/mol) > AzrBmH22/3–dye complexes (-9.2 to -5.4 kcal/mol) > AzrBmH24/5-dye complexes (-9.4 to -5.5 kcal/

mol). Overall, the obtained binding affinity suggested that the stabilization of AzrBmH2–acid orange 7, AzrBmH2–cresol red, and AzrBmH2–methylene blue complexes is mainly driven by hydrogen interactions, with interatomic distances of interaction below 3.5 Å in all cases. In contrast, AzrBmH2–malachite green complexes are based on non-covalent hydrophobic interactions. Molecular dynamics simulations demonstrated the stability of the AzrBmH2–dye complexes, as evidenced by low root mean square deviation (RMSD) values ranging from 0.15 to 0.42 nm, root mean square fluctuation (RMSF) values ranging from 0.05 to 0.48 nm, and radius of gyration (Rg) values ranging from 1.75 to 1.88 nm. The predicted

Table 3 Binding free energies from MM-PBSA in kcal mol^{-1} for the AzrBmH2–substrates complexes

Energy components (kcal/mol)	AzrBmH21				AzrBmH22/3				AzrBmH24/5			
	Acid orange 7	Methylene blue	Cresol red	Malachite green	Acid orange 7	Methylene blue	Cresol red	Malachite green	Acid orange 7	Methylene blue	Cresol red	Malachite green
ΔE_{vdW}	−87.15±13.55	−107.94±14.49	−53.29±13.94	−272.59±5.97	−105.49±14.77	−133.84±17.27	−98.57±7.87	−418.92±29.45	−103.51±15.74	−93.33±9.42	−76.53±7.12	−268.25±12.25
ΔE_{ele}	−30.97±9.69	−73.73±19.06	−105.14±11.55	−186.10±15.39	−65.33±13.29	−68.49±12.83	−81.69±11.27	−212.06±20.93	−57.14±7.41	−51.01±6.32	−110.16±13.51	−164.13±10.11
ΔG_{polar}	38.5±7.56	94.79±7.47	63.76±8.50	92.91±7.19	27.99±2.94	95.66±12.22	107.97±4.46	109.62±27.02	64.33±9.46	54.12±8.41	38.74±5.28	91.87±15.62
$\Delta G_{nonpolar}$	−15.22±4.70	−12.08±1.12	−11.79±1.03	−10.27±2.08	−28.16±2.06	−14.25±0.86	−10.33±0.66	−9.89±1.39	−10.99±3.12	−9.21±1.98	−11.21±3.03	−7.95±1.68 s
$\Delta G_{binding}$	−300.22±11.69	−194.82±13.45	−262.49±14.86	−35.07±5.36	−314.19±6.58	−227.71±21.00	−294.36±2.08	−65.26±8.81	−233.04±6.39	−191.05±7.08	−209.51±14.29	−37.97±2.94

ligand binding-site residues further elucidated the inter-action mechanisms of the tested dyes with potential environmental impact. These findings suggest an efficient biodegradation mechanism for the four dyes studied. The hypothetical evidence presented in this study opens new perspectives for the empirical bioremediation of the methylene blue, malachite green, cresol red, and acid orange 7 dyes using the *Bacillus megaterium* H2 azore-ductases. Future work could progress toward implementing this strategy to identify novel enzymes for treating wastewater contaminated with synthetic dyes.

Availability of data and materials
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
There are no competing to declare.

Author details
[1]Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia. [2]Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia. [3]Enzyme Technology and Green Synthesis Research Group, Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia. [4]Department of Science Technology, Biochemistry unit, The Federal Polytechnic, P.M. B 5351, Ado Ekiti, Ekiti, Nigeria. [5]Department of Biology, College of Science, Princess Nourah Bint Abdulrahman University, 84428 Riyadh, Saudi Arabia. [6]Department of Zoology, College of Science, King Saud University, P.O. Box 2455, 11451 Riyadh, Saudi Arabia. [7]Department of Pharmaceutical Sciences, Pharmacy Program, Batterjee Medical College, P.O. Box 6231, 21442 Jeddah, Saudi Arabia. [8]Pharmacology Department, Faculty of Veterinary Medicine, Suez Canal University, Ismailia 41522, Egypt. [9]Phytomedicine and Molecular Toxicology Research Laboratory, Department of Biochemistry, Federal University Oye-Ekiti, Oye, Ekiti, Nigeria.

References
1. Ahlawat S, Singh D, Virdi JS, Sharma KK (2019) Molecular modeling and MD-simulation studies: fast and reliable tool to study the role of low-redox bacterial laccases in the decolorization of various commercial dyes. Environ Pollut 253:1056–1065
2. Alam R, Mahmood RA, Islam S, Ardiati FC, Solihat NN, Alam MB, Lee SH, Yanto DHY, Kim S (2023) Understanding the biodegradation pathways of azo dyes by immobilized white-rot fungus, *Trametes hirsuta* D7, using UPLC-PDA-FTICR MS supported by in silico simulations and toxicity assessment. Chemosphere 313:137505
3. Arefin A, Ema TI, Islam T, Hossen MS, Islam T, Al Azad S, Badal MNU, Islam MA, Biswas P, Alam NU, Islam E, Paul PK (2021) Target specificity of selective bioactive compounds in blocking α-dystroglycan receptor to suppress Lassa virus infection: an in silico approach. J Biomed Res 35(6):459
4. Bata Z, Molnar Z, Madaras E, Molnar B, Santa-Bell E, Varga A, Leveles I, Qian R, Hammerschmidt F, Paizs C, Vertessy BG, Poppe L (2021) Substrate tunnel engineering aided by X-ray crystallography and functional dynamics swaps the function of MIO-enzymes. ACS Catal 11(8):4538–4549
5. Boroujeni MB, Dastjerdeh MS, Shokrgozar M, Rahimi H, Omidinia E (2021) Computational driven molecular dynamics simulation of keratinocyte growth factor behavior at different pH conditions. Inform Med Unlocked 23:100514
6. Ferdausi N, Islam S, Rimti FH, Quayum ST, Arshad EM, Ibnat A, Islam T, Arefin A, Ema TI, Biswas P, Dey D, Al Azad S (2022) Point-specific interactions of isovitexin with the neighboring amino acid residues of the hACE2 receptor as a targeted therapeutic agent in suppressing the SARS-CoV-2 influx mechanism. J Adv Vet Animal Res 9(2):230
7. Goud BS, Cha HL, Koyyada G, Kim JH (2020) Augmented biodegradation of textile azo dye effluents by plant endophytes: a sustainable, eco-friendly alternative. Curr Microbiol 77:3240–3255
8. Hossain MS, Paul GK, Mahmud S, Saleh MA, Uddin MS, Dutta AK, Roy AK, Saha AK, Sheam MM, Ahmed S, Rahman MM, Biswas SK (2022) Mixed dye degradation by *Bacillus pseudomycoides* and *Acinetobacter haemolyticus* isolated from industrial effluents: a combined affirmation with wetlab and in silico studies. Arab J Chem 15(9):104078
9. Jabin A, Uddin MF, Al Azad S, Rahman A, Tabassum F, Sarker P, Morshed AH, Rahman S, Raisa FF, Sakib MR, Olive AH, Jafary M (2023) Target-specificity of different amyrin subunits in impeding HCV influx mechanism inside the human cells considering the quantum tunnel profiles and molecular strings of the CD81 receptor: a combined in silico and in vivo study. Silico Pharmacol 11(1):8
10. Jendele L, Krivak R, Skoda P, Novotny M, Hoksza D (2019) PrankWeb: a web server for ligand binding site prediction and visualization. Nucleic Acids Res 47(W1):W345–W349
11. Krivák R, Hoksza D (2018) P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. J Cheminform 10:1–12
12. Kumari A, Mittal L, Srivastava M, Pathak DP, Asthana S (2023) Deciphering the structural determinants of F.X.R .Partial agonism. J Phys Chem B. https://doi.org/10.1021/acs.jpcb.2c06325
13. Malla MA, Dubey A, Raj A, Kumar A, Upadhyay N, Yadav S (2022) Emerging frontiers in microbe mediated pesticide remediation: unveiling role of omics and in silico approaches in engineered environment. Environ Pollut. https://doi.org/10.1016/j.envpol.2022.118851
14. Mishra S, Maiti A (2018) The efficacy of bacterial species to decolourise reactive azo, anthroquinone and triphenylmethane dyes from wastewater: a review. Environ Sci Pollut Res 25:8286–8314
15. Morshed AH, Al Azad S, Mia MAR, Uddin MF, Ema TI, Yeasin RB, Srishti SA, Sarker P, Aurthi RY, Jamil F, Samia NSN, Nurunnahar. (2023) Oncoinformatic screening of the gene clusters involved in the HER2-positive breast cancer formation along with the in silico pharmacodynamic profiling of selective long-chain omega-3 fatty acids as the metastatic antagonists. Mol Divers 27(6):2651–2672
16. Nipun TS, Ema TI, Mia MAR, Hossen MS, Arshe FA, Ahmed SZ, Masud A, Taheya FF, Khan AA, Haque F, Al Azad S, Dey D (2021) Active site-specific quantum tunneling of hACE2 receptor to assess its complexing poses with selective bioactive compounds in co-suppressing SARS-CoV-2 influx and subsequent cardiac injury. J Adv Vet Animal Res 8(4):540
17. Oyewusi HA, Akinyede KA, Abdul Wahab R, Huyop F (2021) In silico analysis of a putative dehalogenase from the genome of halophilic bacterium *Halomonas smyrnensis* AAD6T. J Biomol Struct Dyn 41(1):319–335
18. Oyewusi HA, Akinyede KA, Wahab RA, Susanti E, Syed Yaacob SN, Huyop F (2023) Biological and molecular approaches of the degradation or decolorization potential of the hypersaline Lake Tuz *Bacillus megaterium*

Oyewusi *et al. Environmental Sciences Europe*   (2024) 36:31

Page 20 of 20

H2 isolate. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2023.2234040

19. Oyewusi HA, Huyop F, Wahab RA (2021) Genomic characterization of a dehalogenase-producing bacterium (*Bacillus megaterium* H2) isolated from hypersaline Lake Tuz (Turkey). Gene Rep 25:101381

20. Oyewusi HA, Huyop F, Wahab RA (2022) Molecular docking and molecular dynamics simulation of *Bacillus thuringiensis* dehalogenase against haloacids, haloacetates and chlorpyrifos. J Biomol Struct Dyn 40(5):1979–1994

21. Oyewusi HA, Huyop F, Wahab RA, Hamid AAA (2022) In silico assessment of dehalogenase from *Bacillus thuringiensis* H2 in relation to its salinity-stability and pollutants degradation. J Biomol Struct Dyn 40(19):9332–9346

22. Oyewusi HA, Wahab RA, Huyop F (2020) Dehalogenase-producing halophiles and their potential role in bioremediation. Mar Pollut Bull 160:111603

23. Oyewusi HA, Wahab RA, Huyop F (2021) Whole genome strategies and bioremediation insight into dehalogenase-producing bacteria. Mol Biol Rep 48:2687–2701

24. Oyewusi HA, Wahab RA, Kaya Y, Edbeib MF, Huyop F (2020) Alternative bioremediation agents against haloacids, haloacetates and chlorpyrifos using novel halogen-degrading bacterial isolates from the hypersaline lake Tuz. Catalysts 10(6):651

25. Oyewusi HA, Wu YS, Safi SZ, Wahab RA, Hatta MHM, Batumalaie K (2022) Molecular dynamics simulations reveal the inhibitory mechanism of withanolide a against α-glucosidase and α-amylase. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2022.2104375

26. Paul PK, Al Azad S, Rahman MH, Farjana M, Uddin MR, Dey D, Mahmud S, Ema TI, Biswas P, Anjum M, Akhi OJ, Ahmed SZ (2022) Catabolic profiling of selective enzymes in the saccharification of non-food lignocellulose parts of biomass into functional edible sugars and bioenergy: an in silico bioprospecting. J Adv Vet Animal Res 9(1):19

27. Rahman MH, Al Azad S, Uddin MF, Farzana M, Sharmeen IA, Kabbo KS, Jabin A, Rahman A, Jamil F, Srishti SA, Riya FH, Rahman MB (2023) WGS-based screening of the co-chaperone protein DjlA-induced curved DNA binding protein A (CbpA) from a new multidrug-resistant zoonotic mastitis-causing Klebsiella pneumoniae strain: a novel molecular target of selective flavonoids. Mol Divers. https://doi.org/10.1007/s11030-023-10731-6

28. Roda S, Terholsen H, Meyer JRH, Cañellas-Solé A, Guallar V, Bornscheuer U, Kazemi M (2023) A site design: a semi rational algorithm for an automated enzyme design. J Phys Chem B. https://doi.org/10.1021/acs.jpcb.2c07091

29. Romes NB, Abdul Wahab R, Abdul Hamid M, Oyewusi HA, Huda N, Kobun R (2021) Thermodynamic stability, in-vitro permeability, and in-silico molecular modeling of the optimal *Elaeis guineensis* leaves extract water-in-oil nanoemulsion. Sci Rep 11(1):20851

30. Salem SS, Mohamed A, El-Gamal M, Talat M, Fouda A (2019) Biological decolorization and degradation of azo dyes from textile wastewater effluent by *Aspergillus niger*. Egypt J Chem 62(10):1799–1813

31. Sarkar S, Banerjee A, Chakraborty N, Soren K, Chakraborty P, Bandopadhyay R (2020) Structural-functional analyses of textile dye degrading azoreductase, laccase and peroxidase: a comparative in silico study. Electron J Biotechnol 43:48–54

32. Sharma P, Bano A, Singh SP, Dubey NK, Chandra R, Iqbal HM (2022) Recent advancements in microbial-assisted remediation strategies for toxic contaminants. Clean Chem Eng. https://doi.org/10.1016/j.clce.2022.100020

33. Srinivasan S, Sadasivam SK (2018) Exploring docking and aerobic-microaerophilic biodegradation of textile azo dye by bacterial systems. J Water Process Eng 22:180–191

34. Tze Lin K, Mahat NA, Azman AR, Wahab RA, Oyewusi HA, Abdul Hamid AA (2023) Interaction of the nanobio-based reagent with sodium fluorescein and lipids via bioinformatics for forensic fingerprint visualisations. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2023.2186709

35. Vikrant K, Giri BS, Raza N, Roy K, Kim KH, Rai BN, Singh RS (2018) Recent advancements in bioremediation of dye: current status and challenges. Biores Technol 253:355–367

36. Wahhab BH, Oyewusi HA, Wahab RA, Mohammad Hood MH, Abdul Hamid AA, Al-Nimer MS, Edbeib MF, Kaya Y, Huyop F (2023) Comparative modeling and enzymatic affinity of novel haloacid dehalogenase from *Bacillus* megaterium strain BHS1 isolated from alkaline Blue Lake

in Turkey. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2023.2199870

37. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Lepore R (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46(W1):W296–W303

38. Wu Y, Tepper HL, Voth GA (2006) Flexible simple point-charge water model with improved liquid-state properties. J Chem Phys 124(2):024503

39. Yunus AAM, Wahab RA, Bahaman AH, Oyewusi HA, Yaacob SNS (2021) In-silico alanine scanning analysis on the catalytic residues of a novel β-glucosidase from *Trichoderma asperellum* UC1. J Teknol 83(3):61–73

40. Zakary S, Mashal H, Osmani AR, Oyewusi HA, Huyop F, Nasim MM (2022) In silico molecular characterization of a putative haloacid dehalogenase type II from genomic of mesorhizobium loti strain TONO: in silico molecular characterization of a putative haloacid dehalogenase type II. J Trop Life Sci 12(2):241–252

41. 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

Publisher's Note