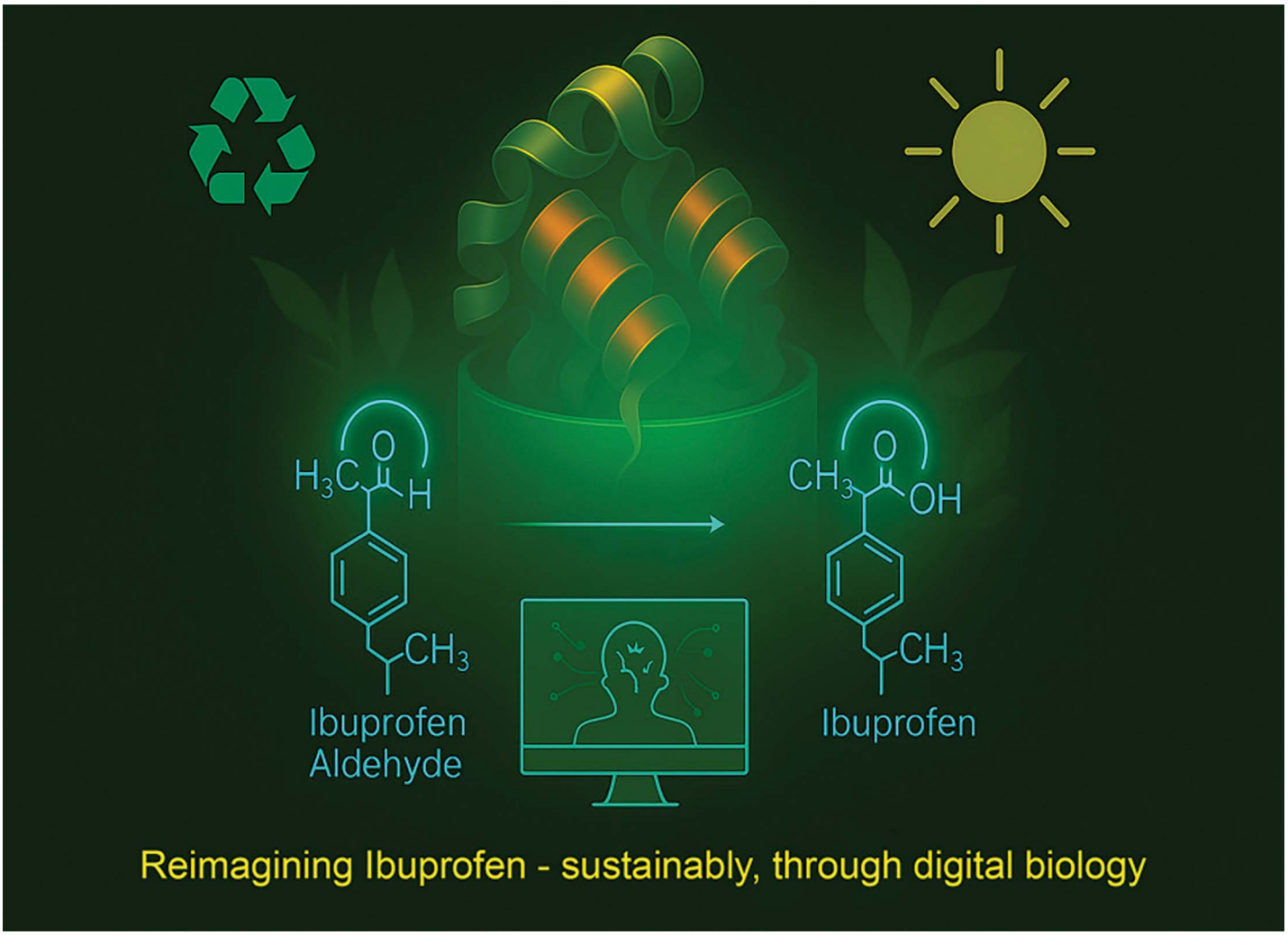

Showcasing research from Naveen Kulkarni's group at Quantumzyme LLP, Bangalore, India.

In silico enzyme engineering of aldehyde dehydrogenase for eco-friendly ibuprofen synthesis

This study demonstrates a major shift in ibuprofen production—from chemical synthesis to biocatalysis. Through in silico enzyme engineering, an aldehyde dehydrogenase is redesigned to convert ibuprofen aldehyde into ibuprofen under greener, milder conditions, symbolizing the fusion of computation and catalysis for sustainable pharmaceutical manufacturing.

Cover image generated with AI (Microsoft Copilot).

Image reproduced by permission of Naveen Kulkarni from *RSC Sustainability,* 2025, **3**, 5495.

As featured in:

See Naveen Kulkarni *et al.*, *RSC Sustainability,* 2025, **3**, 5495.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/rscsus

Registered charity number: 207890

RSC Sustainability

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Sustainability*, 2025, **3**, 5495

In silico enzyme engineering of aldehyde dehydrogenase for eco-friendly ibuprofen synthesis

Ankita Tripathi, † Anisha Ashokan, † Ipsita Basu, Sabhyata Gopal, Akash Ravandur, Shreya Shroff and Naveen Kulkarni *

Received 3rd February 2025
Accepted 10th September 2025

DOI: 10.1039/d5su00073d

rsc.li/rscsus

Ibuprofen is the most widely utilized nonsteroidal anti-inflammatory drug (NSAID) for managing pain and inflammation globally. However, traditional chemical synthesis methods for producing ibuprofen are not environmentally friendly, as they involve hazardous reagents, high energy consumption, and significant chemical waste, along with toxic effluents. The rising global demand for ibuprofen necessitates the exploration of alternative biocatalytic approaches. This study focuses on developing an enzyme through *in silico* driven rational enzyme discovery and engineering to convert ibuprofen aldehyde into ibuprofen. The current study demonstrates a unique biocatalytic synthesis of ibuprofen from ibuprofen aldehyde using aldehyde dehydrogenase enzyme (*bc*PADH), providing a greener alternative by reducing the number of reaction steps from six to four. Furthermore, the solvent tolerance of the enzyme, which is crucial for its practical application in industrial processes has been enhanced, using advanced molecular dynamics simulations and experimental validations. Among the engineered variants, *bc*PADH01 demonstrated the highest stability and conversion efficiency (>80%) in 30% DMSO. Our findings demonstrate that *in silico* strategies are fundamental for rational enzyme design which is required for advancing biocatalytic innovation.

Sustainability spotlight

Biocatalytic revolution: eco-friendly, sustainable synthesis of ibuprofen leveraging *in silico* strategies. The present study addresses the need for greener and more sustainable approaches to ibuprofen synthesis, moving away from traditional chemical processes that pose significant environmental challenges. Our study is the first to demonstrate the synthesis of ibuprofen from ibuprofen aldehyde using an enzyme, in contrast to the existing studies that focus on the synthesis of ibuprofen derivatives using enzymes. Our study represents a significant advancement in the field of green chemistry by being the first to establish a chemo-enzymatic pathway for the sustainable synthesis of ibuprofen. This novel approach eliminates the reliance on harsh, carcinogenic, and pyrophoric chemicals typically used in chemical synthesis methods, thereby reducing environmental and safety concerns. By utilizing *in silico* approaches, enzymes were identified and engineered for efficient catalysis of ibuprofen aldehyde to ibuprofen. The current study also has addressed the limitations of the wild-type enzyme under industrial conditions by employing *in silico* steered rational enzyme engineering, specifically targeting solvent stability. Furthermore, we have enhanced the solvent tolerance of the enzyme, a critical improvement that makes it suitable for large-scale industrial applications. These innovations collectively contribute to greener, safer, and more efficient manufacturing routes. This approach not only aligns with the principles of sustainability but also demonstrates the potential of integrating computational and experimental methodologies for enzyme engineering, paving the way for more eco-friendly pharmaceutical manufacturing processes. Our work emphasizes the importance of the following UN sustainable development goals: good health and well-being (SDG 3), clean water and sanitation (SDG 6), industry, innovation, and infrastructure (SDG 9), responsible consumption and production (SDG 12) and climate action (SDG 13).

1. Introduction

Ibuprofen, a widely used NSAID, was first developed in 1960 as an alternative to steroidal therapy.[1–3] It plays a pivotal role in managing pain due to different conditions, fever and inflammation. Its significance in the pharmaceutical industry cannot be overstated, given its broad range of applications in treating conditions such as arthritis, fever, and dysmenorrhea.[4,5] The global ibuprofen market size valued at USD 1.43 billion in 2023 is estimated to grow at a CAGR of 2.44% by 2032 due to several factors, including rising urbanization and increasing public awareness of chronic diseases.[6] Although crucial for healthcare, the industrial production of ibuprofen has a substantial impact on environmental pollution.[7,8] The widely used Boots process[1,3] is a six-step synthesis method that involves the use of harsh carcinogenic chemicals and generates substantial waste products.[9] This process employs harmful reagents like acetic anhydride and nitrile compounds, which lead to the production of

Quantumzyme LLP, Bangalore, Karnataka, 560004, India. E-mail: naveen.kulkarni@quantumzyme.com

† Equal contribution to this study.

toxic effluents, thereby causing severe environmental damage. This drawback has prompted extensive research aimed at developing alternative environmentally sustainable production methods. One significant alternative is the Hoechst–Celanese (BHC) synthesis route[10] which boasts of higher atom economy in comparison to the Boots process but still employs perilous substances.[11] Despite receiving the Presidential Green Chemistry Award, the BHC process involves the use of carbon monoxide, contributing to serious climate change issues, as well as highly corrosive hydrofluoric acid and pyrophoric Raney nickel, resulting in significant environmental waste and health concerns. Leading companies often face challenges due to the cost and complexity of effluent treatment. Despite process modernization and increased efficiency, their large production scale leaves a strong environmental footprint emphasizing on the urgent need to switch to more environmentally friendly and sustainable production techniques.

In the quest for greener alternatives, biocatalysis has emerged as a promising solution to conventional chemical synthesis.[12,13] Enzymes have been extensively used for the synthesis of active pharmaceutical ingredients due to their ability to catalyze complex reactions with high stereoselectivity and regioselectivity, resulting in highly pure products with lesser chemical wastage.[14,15] While nature offers a wide array of enzymes as biocatalysts, identifying a specific enzyme with high selectivity and specificity for catalyzing the desired reaction can be quite challenging. Additionally, enzyme stability and activity can get affected under different reaction conditions such as pH, temperature and solvents.[16] These factors can impede the cost-effectiveness and practical applications of biocatalysts in industrial settings. To tackle these problems, *in silico* rational enzyme discovery and engineering has proven to be the key in revolutionizing the field of biocatalysis.[17–19]

Prior enzymatic methodologies have been employed for the enantiomeric resolution of racemic ibuprofen to yield *S*-ibuprofen, which exhibits higher pharmacological efficacy compared to the racemate.[20] Nevertheless, due to the ease of production and the *in vivo* conversion of the *R*-enantiomer to the *S*-enantiomer, racemic ibuprofen remains the most extensively sold drug formulation globally. Various biocatalytic investigations have elucidated the synthesis of ibuprofen derivatives, such as monoglycerides and glucopyranoside methyl esters, using enzymes.[21–24] While extant enzymatic strategies predominantly focus on the generation of ibuprofen formulations, our current research outlines the first novel report of the biocatalytic synthesis of ibuprofen from ibuprofen aldehyde, thereby significantly advancing the state-of-the-art. This study introduces environmentally benign, green synthesis reagents, such as the aldehyde dehydrogenase enzyme, into the widely adopted industrial production methods. The reduction of the synthesis process from six to four steps and use of an enzyme catalyst further highlights the green potential of this approach compared to Boots and BHC processes. Enhancing enzyme solvent tolerance and overall stability is pivotal for the feasibility of biocatalytic processes on an industrial scale. Our research intends to address these gaps by the design of a comprehensive chemoenzymatic pathway for the synthesis of ibuprofen guided by *in silico* strategies. This work demonstrates ibuprofen synthesis using a sustainable, greener method and proves that *in silico* strategies form the cornerstone of modern biocatalytic innovation.

2. Methodology

2.1. Enzyme discovery

The oxidation of aldehydes can be catalyzed by several dehydrogenases, specifically aldehyde dehydrogenase family, as identified through literature survey. Among these, phenylacetaldehyde dehydrogenase was selected as the primary template due to its substrate structural similarity with ibuprofen aldehyde.

A dual *in silico* strategy was employed for enzyme screening. In the first approach, a BLAST[25] search using phenylacetaldehyde dehydrogenase (EC 1.2.1.39) from *Pseudomonas putida* (PDB ID: 4QYJ) yielded homologous sequences, which were aligned and filtered based on active site conservation. The second approach involved querying the BRENDA[26] database for oxidoreductases acting on aldehydes or oxo groups as donors with NAD^+/$NADP^+$ as acceptors. Enzymes with substrate profiles resembling ibuprofen aldehyde—particularly those with large, non-polar substrates—were prioritized. Binding pocket characteristics were evaluated using CAVER,[27] and 22 enzymes were shortlisted based on structural availability, pocket volume, and active site accessibility.

2.2. Molecular docking

The crystal structure of the enzymes was obtained from RCSB protein data bank. The missing amino acids were modelled using Modeller 10.[28,29] Modeller utilises the homology modelling algorithm which involves identifying and aligning homologous sequences from known templates to the target sequence (Table S1). Then the substrate (ibuprofen aldehyde) was prepared in Avogadro.[30] Molecular docking was performed using AutoDock vina[31] to get the final enzyme substrate complex. A grid box encompassing the entire active site was defined to ensure comprehensive sampling. Flexible docking was employed, allowing key active site residues to adapt during ligand binding, while the rest of the protein remained rigid. The conformational search was carried out using the Lamarckian Genetic Algorithm with default parameters to explore the optimal binding poses.

2.3. Gene expression

The genes were synthesized and cloned into the bacterial expression vector pET28a between the restriction sites NdeI and XhoI. The plasmids were transformed using *E. coli* BL21(DE3) competent cells as the expression host. A single microbial colony was inoculated in 10 mL fresh Luria–Bertani broth containing 50 μg mL^{-1} kanamycin and grown overnight at 37 °C with continuous shaking. The overnight grown culture was then transferred to 500 mL Terrific Broth containing 50 μg mL^{-1} kanamycin with continuous shaking and allowed to grow at 37 °C. Once the OD600 reaches 0.8, IPTG (0.6 mM) was added for protein induction and continued to grow for 16 hours at 37 °C.

The culture was harvested, and the cell pellet was resuspended in ice-cold lysis buffer[32] (50 mM Tris-Cl, pH 7, 200 mM NaCl). Cell lysis was carried out using a sonicator with 5 s 'ON' and 10 s 'OFF' cycle for 15 min, followed by centrifugation at 3800×*g* for 40 minutes to clarify the lysate and remove the cell debris. The total protein concentration of the lysate was estimated using Bradford Assay[33] with BSA as the standard, and the protein expression was analyzed using 12% SDS-PAGE. The variants were expressed using the same conditions as the wild type except that the expression host strain used was C41DE3.

2.4. *In vitro* biocatalytic synthesis of ibuprofen

The biocatalysis was conducted for the wildtype and variant enzymes using the ibuprofen aldehyde substrate (5–260 mM), 1–50% DMSO, 50 mM potassium phosphate buffer (pH 8), 10 mM NAD, and enzyme lysate (0.5–8.0 g L^{-1}). The reaction mixture was incubated at 40 °C for 24 hours with continuous stirring of 200–300 rpm. After 24 h, the reaction was stopped by adding an equal volume of ethyl acetate and mixed for 2–3 min. The mixture was then centrifuged at 6000 rpm for 5 min to separate the organic layer. The collected organic layer was concentrated and then analysed by RP-HPLC for estimation of product formation. HPLC analysis utilized a C18 column (Inertsil ODS-3V (4.6 × 250 mm), 5 μm) with the mobile phase consisting of acetonitrile and water (60 : 40, v/v) along with 0.1% orthophosphoric acid at a flow rate of 1.5 mL min^{-1} and detection at 214 nm. The standards and samples were resuspended in the mobile phase and filtered using a 0.2 μ syringe filter. ^{1}H NMR was used to confirm the product formed.

2.5. Enzyme engineering

Based on biocatalysis, *bc*PADH (PDB ID: 4O5H) enzyme was selected to engineer for enhanced solvent tolerance. For our studies, coarse-grained (CG) simulations were done using three solvents such as hexane, toluene and DMSO.

2.5.1. Molecular dynamic simulation. From the literature, it was determined that the active form of phenyl acetaldehyde dehydrogenase (*bc*PADH) exists as either a dimer or tetramer.[34,35] The CHARMM-GUI Martini Maker[36] was employed to generate the CG model of the protein using the Martini 3.0 force field (Fig. S1 and S2). The coarse-grained representation of the protein was then validated by calculating RMSD against the atomistic model, to ensure structural integrity and proper mapping of residues.

The solvent parameters were generated using the auto Martini approach, which ensures compatibility with the Martini force field. Different systems were prepared with varying concentrations of these solvents ranging from 10 to 30% (v/v). Each system was solvated accordingly and subjected to periodic boundary conditions[37] to mimic an infinite system and avoid edge effects. Energy minimization[38] was performed on each system to remove any steric hindrances or poor contacts. This was followed by an equilibration phase, where the systems were gradually heated to the target temperature of 310 K, using a Berendsen thermostat[39] to control the temperature. The pressure was maintained at 1 bar using a Parrinello–Rahman barostat.[40] Production molecular dynamics simulations were then carried out for 500 nanoseconds (ns) for each system. The simulations were conducted using GROMACS,[41,42] with a time step of 20 femtoseconds (fs), and trajectory data were collected every 100 picoseconds (ps). Analysis of the simulation trajectories was performed to evaluate the stability, conformational changes, and interactions of the protein dimer in different solvent environments. Key metrics such as root mean square deviation (RMSD),[43] solvent-accessible surface area (SASA),[44] radius of gyration (Rg)[45] and Molecular Mechanics Poisson–Boltzmann Surface Area (MM-PBSA)[46,47] analysis were calculated to assess the impact of the solvents on the protein structure.

3. Results and discussion

3.1. Enzyme discovery

Structural superposition of the shortlisted enzymes was performed to visualize conserved domain architecture and key functional features. All enzymes exhibited the characteristic aldehyde dehydrogenase fold comprising three primary domains: catalytic (wheat), cofactor binding (light green), and oligomerization (yellow) (Fig. 1a). Conserved catalytic residues—including the triad (Cys309, Glu275, Lys200), cofactor-interacting residues (Glu 203, Ilu 174, Gln 356, Glu 406, Phe 408), and substrate entry channel residues (Met126, Phe472)—were identified across the structures, supporting their functional relevance for ibuprofen aldehyde oxidation. While the specific residue numbers may vary among enzymes due to sequence differences, the type and spatial positioning of these functional residues remain conserved. The numbering provided corresponds to the top-performing enzyme, *bc*PADH (Fig. 1b).

Stemming from the enzyme discovery highlighted in the methodology, 22 enzymes were shortlisted as mentioned in Table S1. Molecular docking for these enzymes was performed using ibuprofen aldehyde as the substrate, and the reactive distance, *i.e.*, the donor–acceptor distance between the cofactor and the catalytic residue – cysteine (Cys) was calculated. According to the reactive distance and binding affinity (Table 1), the five enzymes that showed maximum binding energy were chosen for further *in vitro* validation. Fig. 1c–g represents the best pose of the shortlisted enzyme–substrate complexes.

3.2. Initial screening of wildtype enzymes

The catalytic activities of the five shortlisted enzymes with the best QZyme score were tested using the oxidation reaction of ibuprofen aldehyde to ibuprofen. The results of the enzymatic activity as derived from RP-HPLC analysis are provided. Conversion represents the percentage of product observed at the end of 24 hours in the reaction mixture. The enzymatic assay showed that *bc*PADH exhibited the maximum conversion of the ibuprofen aldehyde substrate to ibuprofen. The purity of the ibuprofen formed was confirmed as ~89% using HPLC analysis (Fig. S3).

3.3. Effect of pH and enzyme loading

To enhance the reaction efficiency, an investigation into the impact of pH and enzyme loading was conducted. The wild-type

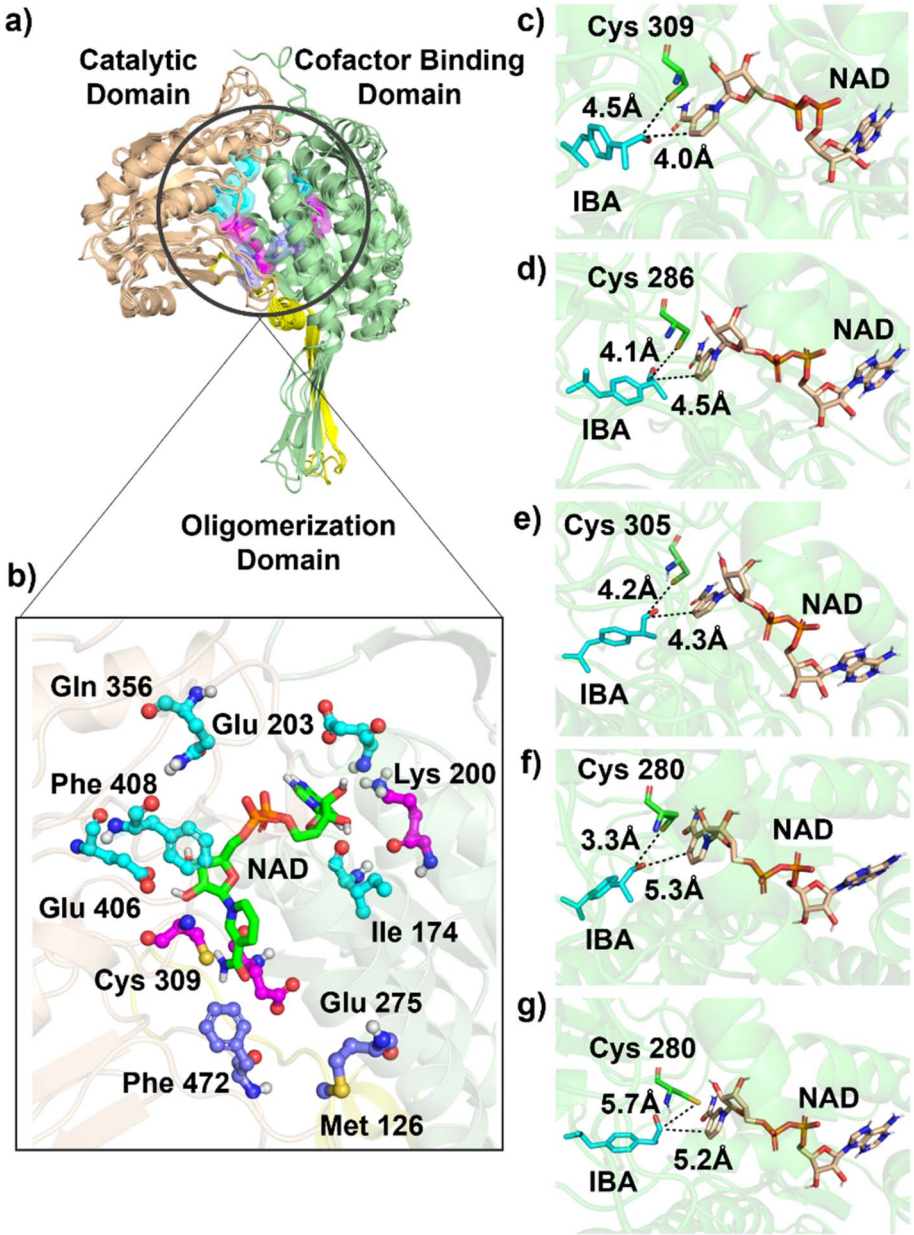

Fig. 1 Structural superposition of the top five shortlisted aldehyde dehydrogenases highlighting conserved domains and functional features. (a) Wheat: catalytic domain; light green: cofactor binding domain; yellow: oligomerization domain; (b) cyan: cofactor-binding residues; magenta: catalytic triad; light blue: substrate entry/exit channel; (c–g) docked poses of ibuprofen aldehyde in the active sites of the five shortlisted aldehyde dehydrogenases. The substrate, catalytic cysteine, and cofactor (NAD^+) are shown in stick representation with their respective catalytic distance. Enzymes include (c) *bc*PADH (PDB ID: 4O5H), (d) *Kp*BADH (PDB ID: 7SWK), (e) *tt*CHADH (PDB ID: 2D4E), (f) *sm*PADH (PDB ID: 4DAL), and (g) *sw*GADH (PDB ID: 6C43), selected based on binding energy and reactive distance criteria (see Table 1).

enzyme *bc*PADH exhibited peak activity at pH 8.0, indicating a preference for a slightly basic pH environment. The slight reduction in activity at pH 7.5 and 9.0 may be attributed to the potential structural instability of the enzyme (Fig. 2a). This pH dependency aligns with findings in related aldehyde dehydrogenases wherein the formation of thiolate anion of the conserved active site cysteine as a nucleophile in reductive catalysis occurs at this pH.

Furthermore, it was noted that elevating the enzyme loading up to 4 g L^{-1} resulted in increased biocatalytic conversion. However, beyond 4 g L^{-1}, the enzymatic reaction reached a saturation point indicating the substrate unavailability (Fig. 2b). All the data have been obtained by RP-HPLC analysis.

3.4. Assay with increased substrate loading

An experiment involving varying substrate loadings was conducted to assess the maximum substrate consumption of the wildtype enzyme. The reaction was carried out under previously optimized conditions, specifically with 4 g L^{-1} enzyme loading

Table 1 Summary of the binding energy and conversion obtained using different wildtype enzymes shortlisted for ibuprofen synthesis

Enzyme ID	PDB ID	Source organism	Enzyme type	Binding energy (kcal mol^{-1})	Score	Conversion (%)
*bc*PADH	4O5H	*Burkholderia cenocepacia*	Phenylacetaldehyde dehydrogenase	−7.48	99.5	85.32
*Kp*BADH	7SWK	*Klebsiella pneumoniae*	Betaine-aldehyde dehydrogenase	−7.60	83.0	18.84
*tt*CHADH	2D4E	*Thermus thermophilus*	5-Carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase	−6.70	94.3	18.48
*sm*PADH	4DAL	*Sinorhizobium meliloti*	Putative aldehyde dehydrogenase	−6.20	91.3	19.08
*sw*GADH	6C43	*Salmonella enterica*	Gamma-aminobutyraldehyde dehydrogenase	−6.08	94.3	18.48

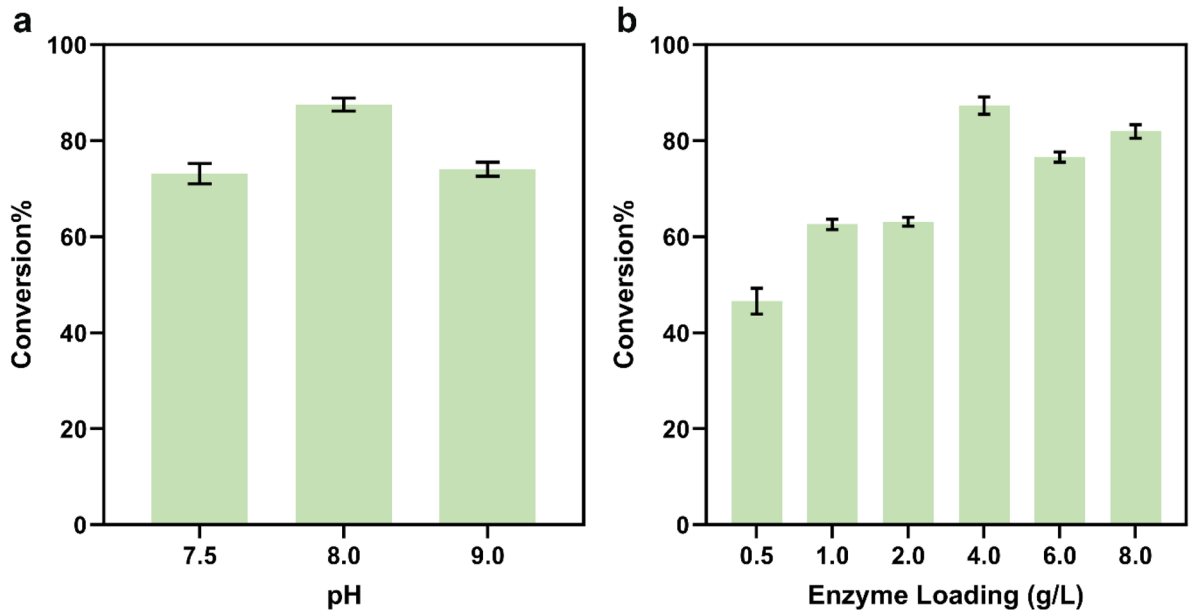

Fig. 2 (a) Effect of pH on enzyme activity. (b) Effect of enzyme loading on enzyme activity.

and at pH 8.0. When the substrate loading was set at 2 g L^{-1}, the wildtype enzyme exhibited peak activity, leaving only ~1.6% of the substrate unreacted. However, as the substrate loading increased, enzyme activity decreased, accompanied by the significant generation of impurities (Fig. 3a). It is plausible that these impurities resulted from the reaction between the unreacted substrate and other components or from the degradation of the substrate over time. The decline in activity with increasing substrate loading may also be attributed to reduced availability of the substrate to the enzyme. This emphasizes the significance of using larger solvent volumes in the reaction.

To confirm this hypothesis, we conducted an experiment by varying the solvent volumes (using DMSO solvent) in the reaction, ranging from 2% to 75%, while maintaining a substrate : enzyme loading ratio of 1 : 1 (4 g L^{-1}). As the solvent volume increased, there was a substantial decrease in the formation of the product, accompanied by a significant generation of impurities. This phenomenon could be attributed to the reduced stability and activity of the enzyme in the presence of higher solvent concentrations. This highlights the need to enhance the enzyme's tolerance to solvents in order to improve its overall activity.

3.5. Enzyme engineering

Following the enzyme biocatalysis assay, *bc*PADH showed the highest substrate conversion at a low concentration of 2 g L^{-1}. Increasing the organic solvent volume enhances substrate solubility and allows more substrate to reach the enzyme active site, potentially boosting product formation. However, *in vitro* results reveal that organic solvents negatively affect the enzyme stability by disrupting essential hydrophobic interactions and destabilizing its dimeric or tetrameric forms. To address these issues, enzyme engineering approaches were employed, including site-directed mutagenesis to enhance hydrophobic interactions, enzyme engineering to improve solvent tolerance and designing stable multimeric forms to maintain enzyme functionality.

With this approach, the aim was to enhance the stability of the enzyme in its tetrameric form in various organic solvents. To attain this, a coarse-grained model of the enzyme was employed. This approach played a pivotal role in identifying the key residues at the dimeric interface efficiently. The goal here was to strengthen the dimeric interactions to assure that the enzyme remains stable in the presence of organic solvents.

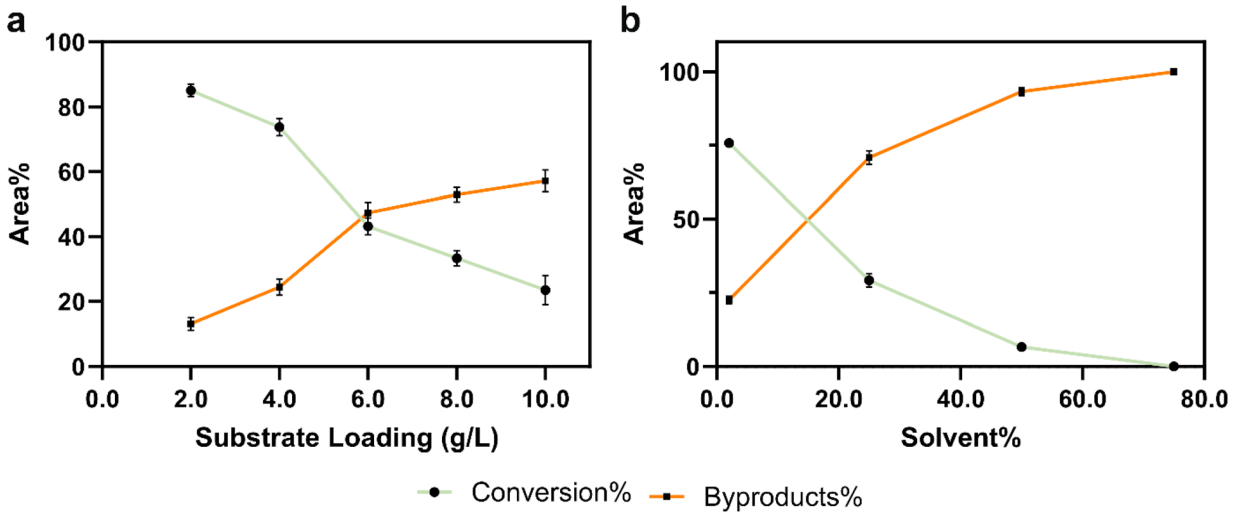

Fig. 3 (a) Summary of RP-HPLC analysis of experimentation with increased substrate loading; (b) effect of increased solvent volumes in the reaction on enzyme activity.

Two strategies were employed to locate hotspots in the dimeric interfacial area. The first approach necessitated identifying non-polar residues at the interface region and mutating these with charged residues. This strategy aimed to form salt bridges with nearby oppositely charged residues from the other monomeric unit, making it highly unlikely for organic solvents to disrupt the stable salt bridge interactions. The second approach focused on substituting aliphatic non-polar residues at the dimeric interface with aromatic residues, thereby increasing cation–π interactions with adjacent residues at the dimeric interface.

To systematically implement these strategies, it was crucial to identify key residues at the dimeric interface that could serve as potential hotspots for mutation. To achieve this, trajectory clustering analysis[48] was performed using the GROMOS method, with a root-mean-square deviation (RMSD) cutoff of 2 Å, to determine the most populated conformational state of the *bc*PADH wild-type enzyme. The resulting structure was then analyzed to identify potential hotspots. The analysis revealed that adjacent to V458 and L259 at the interface, there were positively charged K496 and K274 residues from the other monomer, while around M158, there was a negatively charged residue, D475 (Fig. 4). Monomer 1 is shown in cyan while Monomer 2 is in green. Consequently, the first set of mutations included V458E, V458W, M158K, L259D, L259W, and L259F (Table 2). The coarse-grained simulations for each mutant were conducted for 500 ns. To calculate the interfacial interaction energy between all possible units in the tetramer, Molecular Mechanics Poisson–Boltzmann Surface Area (MM-PBSA) analysis was performed. The MM-PBSA analysis (Table 3) revealed that among all the designed variants, M158K exhibited the lowest total interaction energy (-4211 kcal mol^{-1}), indicating the highest stability of the dimeric interface. The breakdown of individual monomer–monomer interaction energies further supports this, with significant stabilization observed in the A–D and B–C interactions compared to the wild type (WT). Moreover, energy decomposition (Table 4) of the hotspot residue demonstrated that the lysine substitution at position 158 (M158K) contributed the most to the total interaction energy (-345 kcal mol^{-1}), highlighting its key role in enhancing

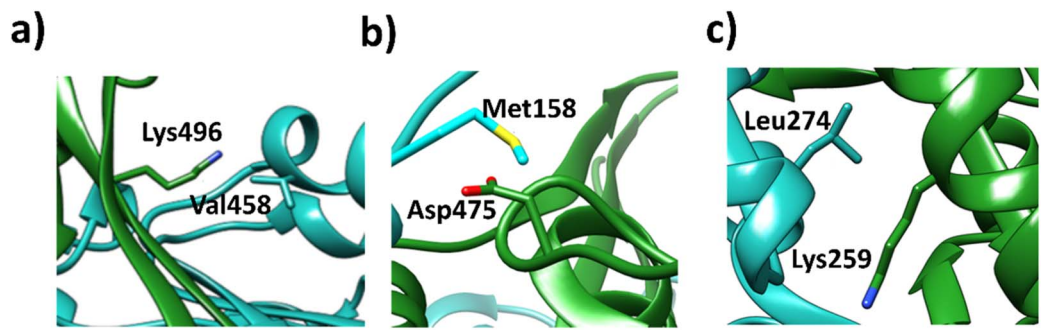

Fig. 4 The most populated structure of the *bc*PADH wild type (WT) enzyme. Monomer 1 is in cyan and Monomer 2 is in green. (a) The position of Val458 and Lys496 from Monomer 1 and 2 respectively. (b) The position of Met158 and Asp475 from Monomer 1 and 2 respectively. (c) The position of Leu259 and Lys274 from Monomer 1 and 2 respectively.

Table 2 Possible mutations for highlighted residues in Fig. 4

Residue pairs		Mutations to introduce salt–bridge interaction	Mutations to introduce cation–π interaction
Monomer 1	Monomer 2		
V458	K496	V458E	V458W
M158	D475	M158K	—
K274	L259	L259D	L259W, L259F

Table 3 Summary of interaction energies between each monomeric unit of the WT *bc*PADH tetramer structure and its respective variants from MM-PBSA analysis. Each monomeric unit is labeled as A, B, C and D. Interaction energies are represented in kcal mol^{-1}

	WT	V458E	V458W	M158K	L259D	L259W	L259F
A–B	−785	−82	−442	−446	−355	−404	−563
A–C	−521	−405	−677	−636	−106	−647	−656
A–D	−606	−365	−643	−1134	−467	−343	−547
B–C	−527	−587	−375	−1152	−54	−133	−548
B–D	−127	−194	−459	−385	−445	−248	−238
C–D	−564	−100	−691	−459	−407	−404	−585
Total	−3130	−1733	−3287	−4211	−1835	−2180	−3138

Table 4 Summary of total interaction energy and contributions of the hotspots to the interaction energy of the WT *bc*PADH tetramer structure and its respective variants from MM-PBSA analysis

Variants	Total interaction energy (kcal mol^{-1})	Contribution of the hotspot residues to interaction energy (kcal mol^{-1})		
		158	259	458
WT	−3130	−17 (Met)	−7 (Leu)	−6 (Val)
M158K	−4211	−345 (Lys)	NA	NA
L259D	−1835	NA	17 (Asp)	NA
L259W	−2180	NA	−33 (Trp)	NA
L259F	−3138	NA	−24 (Phe)	NA
V458E	−1733	NA	NA	67 (Glu)
V458W	−3287	NA	NA	−19

interfacial stability. In contrast, other variants, such as V458E and L259D, exhibited destabilizing effects, as reflected by their reduced interaction energies. The introduction of aromatic residues at position 259 (L259W, L259F) provided moderate stabilization, but their contributions were not as pronounced as that of M158K. Based on these findings, M158K was identified as the most stable variant, making it the best choice for improving dimeric interface interactions.

To further investigate the factors influencing multimeric stability, residue-wise energy decomposition analysis was conducted on the WT *bc*PADH enzyme. This analysis identified interfacial residues D93, D450, E491, D498, and E502 as contributing positive interaction energies, indicating their potential role in destabilizing the multimeric assembly. The presence of these negatively charged residues at the interface may lead to electrostatic repulsion or disruption of stabilizing interactions, thereby weakening the structural integrity of the complex. To mitigate these destabilizing effects, a rational mutagenesis approach was employed, wherein these residues were substituted with either oppositely charged arginine (Arg) or aromatic polar tyrosine (Tyr) residues. These substitutions were designed to enhance interfacial interactions through the formation of stabilizing salt bridges or hydrogen bonds. The proposed mutations, along with their corresponding interfacial residues, are summarized in Table S2.

All the target variants were modelled using Modeller 10, and coarse grain simulation for 500 ns was performed to calculate the interaction energies using MM-PBSA analysis (Table S3). The MM-PBSA analysis demonstrated that among the engineered variants, D93R/E491R exhibited the lowest total interaction energy (−6677 kcal mol^{-1}), indicating the highest stabilization of the multimeric assembly. Individual residue-wise energy decomposition (Table S4) further confirmed that the D93R and E491R substitutions contributed significantly to this stabilization, reducing the positive interaction energy observed in the wild-type enzyme. Based on the analysis of all mutants, those exhibiting the lowest binding energies are listed in Table 5 and have been selected for further engineering to improve solvent tolerance.

Notably, the coarse-grained model of the enzyme did not include the substrate, cofactor, or specific organic solvents in these simulations. To extend this analysis, the top-performing variants with the lowest binding energies, *bc*PADH01 and *bc*PADH04, from Table 5 were selected for further evaluation under different solvent conditions. Three organic solvents—DMSO, toluene, and hexane—were chosen due to their widespread industrial relevance, allowing for a systematic assessment of enzyme stability and performance in diverse environments.

For further understanding, a stability analysis was carried out by evaluating RMSD of the protein backbone structures of

Table 5 List of the six best variants with the lowest binding energy

Sl. No	Enzyme ID	Variant	ΔE (kcal mol^{-1}) (QZ-variant-WT)
1	*bc*PADH01	D93R/E491R	−3547
2	*bc*PADH02	D93R	−2313
3	*bc*PADH03	E491R	−1239
4	*bc*PADH04	D498R	−1225
5	*bc*PADH05	E502R	−1199
6	*bc*PADH06	M158K	−1081

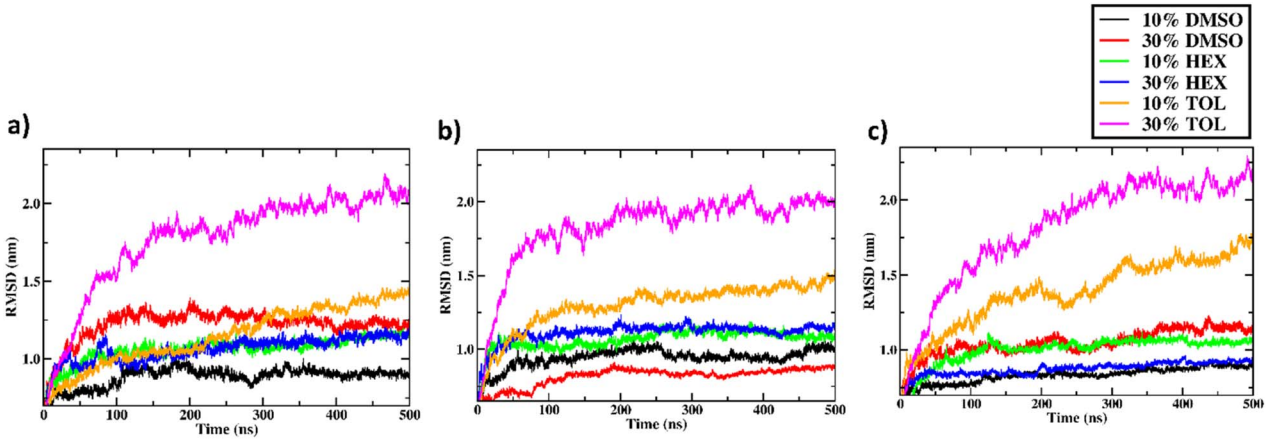

Fig. 5 This graph shows the RMSD of the backbone structures for (a) *bc*PADH, (b) *bc*PADH01, and (c) *bc*PADH04. While the black and red lines represent simulations with 10% and 30% DMSO, respectively, the green and blue lines show simulations with 10% and 30% hexane. The orange and magenta lines depict simulations with 10% and 30% toluene as labeled in the figure.

*bc*PADH01, and *bc*PADH04 relative to the crystal structure, *bc*PADH (PDB ID: 4O5H). The results for this analysis are shown in Fig. 5. A distinguishable trend was noticed, indicating that the selected variants (*bc*PADH, *bc*PADH01, and *bc*PADH04) demonstrated more pronounced deviations when exposed to toluene. In contrast, deviations were least pronounced in the presence of DMSO while deviations were comparatively more pronounced in hexane. Hence this result suggested that DMSO provides a comparatively more stable environment than the other solvents. Notably, *bc*PADH01 demonstrated significant stability compared to its counterparts.

When the DMSO concentration increased from 10% to 30%, *bc*PADH01 exhibited consistent deviations, indicating its stability, whereas *bc*PADH04 and *bc*PADH showed increased deviations, whereas in the presence of hexane, an increase in concentration from 10% to 30% resulted in stable deviations for all *bc*PADH, *bc*PADH01, and *bc*PADH04. In contrast, when the toluene concentration was elevated, *bc*PADH01 demonstrated the highest structural stability, whereas *bc*PADH04 and *bc*PADH showed greater deviations. These findings underscore the distinct influence of solvent composition on enzyme stability, highlighting *bc*PADH01 as the most stable variant under different solvent conditions.

Following this analysis, solvent accessible surface area (SASA) calculations were conducted to determine the extent to which protein sidechains were exposed to the solvents. Increased exposure of sidechains to solvents can potentially destabilize the protein structure. In 30% toluene simulation, the average SASA values for *bc*PADH, *bc*PADH01, and *bc*PADH04 were 857 nm^2, 855 nm^2, and 858 nm^2, respectively, as depicted in Fig. 6. In contrast, the average SASA values for the DMSO and hexane were below 600 nm^2. This elevated exposure of protein sidechains in the toluene environment suggested a higher degree of solvent accessibility, potentially leading to destabilization and decreased activity. Conversely, the lower SASA values observed in simulations with DMSO and hexane indicated

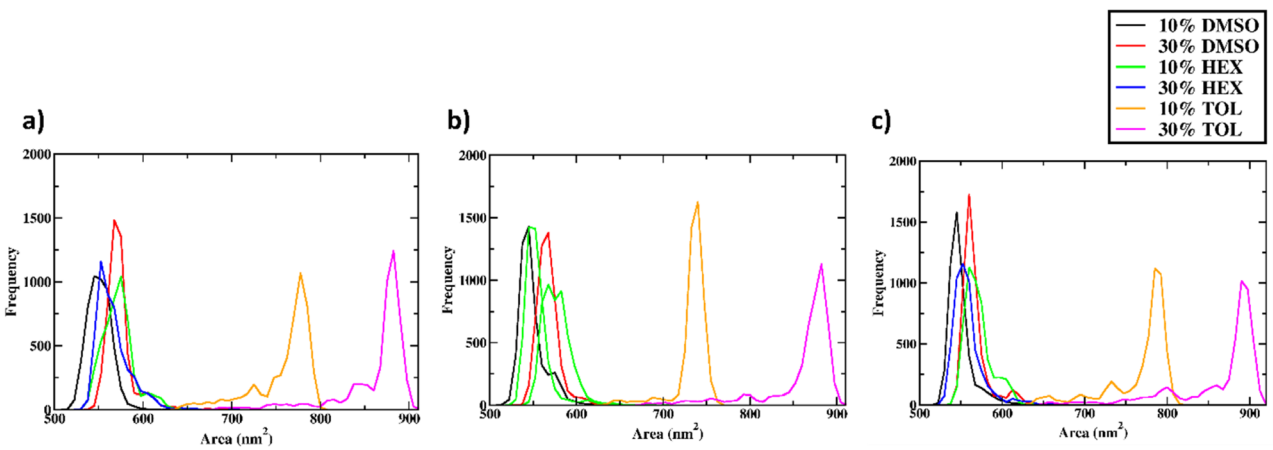

Fig. 6 Distribution graph of solvent accessible surface area for (a) *bc*PADH, (b) *bc*PADH01, and (c) *bc*PADH04. The black and red lines represent simulations with 10% and 30% DMSO, respectively, green and blue lines show simulations with 10% and 30% hexane, while orange and magenta lines depict simulations with 10% and 30% toluene as labeled in the figure.

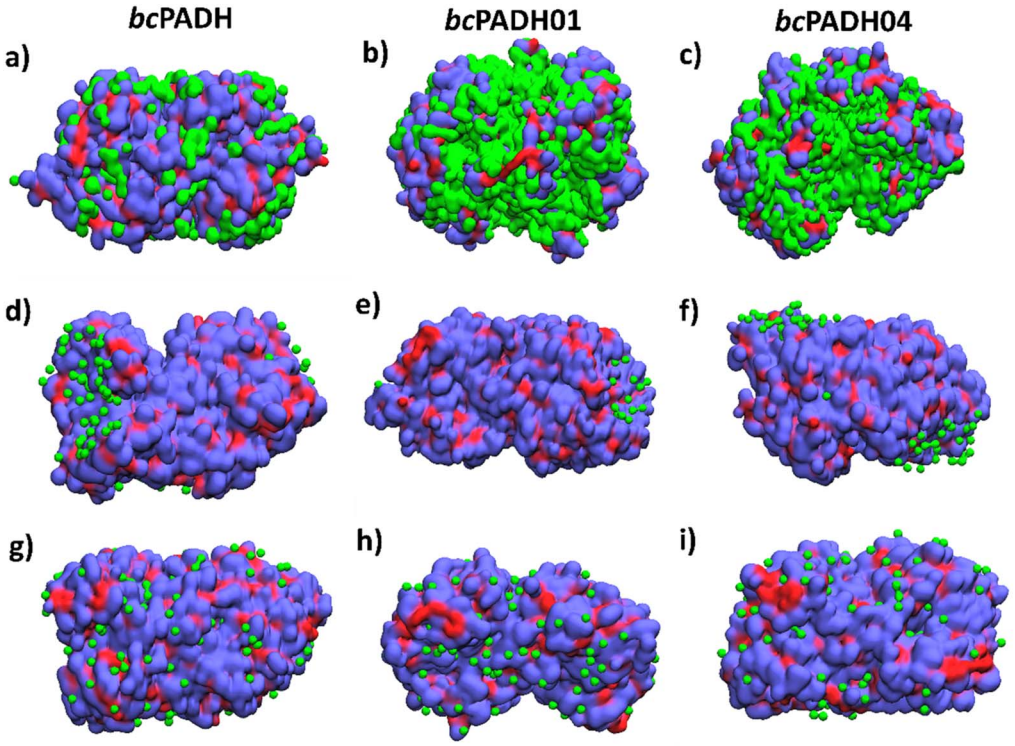

Fig. 7 Representative population in the simulation of solvent molecules within 5 Å of enzyme under 30% solvent conditions. (a–c) Toluene (d–f) hexane (g–i) DMSO. The green surface represents the respective solvent and mass coloring method of VMD for the enzyme.

reduced solvent accessibility, likely contributing to greater protein stability in those environments.

To further understand the structural implications of the solvent interactions, we analyzed the radius of gyration (RoG) of the enzymes under solvent conditions. The RoG is a measure of the protein overall compactness and can provide insights into conformational changes in response to different environments. By correlating the SASA and RoG results, we aimed to better understand the impact of solvent accessibility on the protein structural stability. As depicted in Fig. S4, the RoG remained similar under hexane and DMSO conditions, indicating a comparable level of structural compactness and stability in both these solvents. However, a recognizable increase in RoG was observed in the presence of toluene. This finding is consistent with the SASA results, where higher solvent accessibility in toluene correlated with the destabilization of the protein structure. The increase in RoG in toluene solvent suggests that the structure of protein altered due to increased solvent–protein interactions, which aligns with the higher SASA values.

Following the RoG analysis, we investigated the proximity of solvent molecules to the enzyme surface, particularly focusing on those within a 5 Å radius. In the presence of toluene, solvent molecules are highly aggregated near the enzyme surface, as illustrated in Fig. 7a–c. This extensive coverage contrasted with the hexane simulation, where solvent molecules are aggregated near the protein but did not fully envelop it, as observed in Fig. 7c–e. Interestingly, in the DMSO simulation, aggregation of solvent molecules was not observed, with only a few DMSO molecules bound to the enzyme as shown in Fig. 7f–h. For quantitative values refer to Fig. S5.

The reduced interaction between DMSO molecules and the enzyme surface may have contributed to the overall stability observed under DMSO solvent conditions compared to toluene and hexane. The minimal interaction in DMSO aligned with the lower SASA values and stable RoG, indicating more solvent tolerance and a comparatively less effect on the protein structure. In contrast, the extensive solvent–protein interactions in toluene, reflected in both the higher SASA and increased RoG, suggested that the protein is more exposed to the solvent and hence less solvent tolerance, and also a destabilizing effect due

Table 6 Effect of varying solvents on the activity of enzyme variants

Substrate loading	Solvent %	*bc*PADH01 conversion%	*bc*PADH04 conversion%
2	1%	84.02 ± 2.01	78.66 ± 5.13
2	5%	82.64 ± 2.12	82.88 ± 3.35
2	10%	76.06 ± 3.48	85.28 ± 5.37
2	20%	77.91 ± 5.69	83.96 ± 1.64
2	30%	65.49 ± 2.55	69.86 ± 4.54
3	5%	88.01 ± 1.04	77.61 ± 5.54
5	10%	79.71 ± 1.32	83.73 ± 5.65
10	20%	74.27 ± 1.56	52.84 ± 5.39
10	30%	80.65 ± 2.08	56.70 ± 1.42
10	40%	39.91 ± 2.95	20.27 ± 2.13

Table 7 Comparative life cycle assessment of ibuprofen production

Parameter	Chemical process	Biocatalytic process	Remarks
Energy consumption	High	Low	The enzymatic step operates under mild conditions, reducing overall energy demand
Hazardous reagents	Acetic anhydride, nitrile compounds, carbon monoxide, Raney nickel	Reduced	Transition to biocatalysis minimizes the use of toxic and corrosive reagents
Chemical waste	High	Reduced	The enzymatic step does not produce toxic effluents, thereby reducing waste treatment
Greenhouse gas emission	High	Nil	Eliminates the use of carbon monoxide
Overall environmental impact	High	Low	The chemoenzymatic process demonstrates a clear advantage in sustainability metrics

to the close proximity of toluene molecules to the enzyme structure. By integrating the SASA, RoG, and solvent proximity analyses, we provide a comprehensive understanding of the influence of different solvents on protein structure and stability, highlighting the distinct effects of toluene, hexane, and DMSO on enzyme activity.

3.6. Effect of *bc*PADH enzyme engineering on solvent tolerance

As described above, *bc*PADH01 and *bc*PADH04 were tested for their catalytic activity in the presence of varying organic solvents. Concurring with the *in silico* results that suggest that DMSO exudes a protective effect on the enzyme, *bc*PADH01 demonstrated the highest tolerance to DMSO at 30% with >80% conversion while *bc*PADH04 showed reduced activity with ~60% conversion. The substrate consumption of *bc*PADH01 also increased 5-fold (10 g L^{-1}). Non-polar solvents beyond 5% (hexane and toluene) continued to have a deteriorating effect on the enzyme catalytic activity (Table 6).

3.7. Life cycle impact assessment

Our current study focuses on sustainable production methods to minimize environmental impact and enhance biosafety. This study presents a comparative life cycle impact assessment (LCA) of two distinct routes for ibuprofen production: the traditional chemical process and the innovative biocatalytic process in Table 7. Key impact categories such as energy consumption, hazardous reagents, chemical waste, greenhouse gas emissions, and overall environmental impact have been evaluated. Overall, the chemical process has significantly higher environmental impacts across all categories, while the biocatalytic process demonstrates its potential as a sustainable method for ibuprofen production.

4. Conclusion

The present study addresses the need for greener and more sustainable approaches to ibuprofen synthesis, moving away from traditional chemical processes that pose significant environmental challenges. By utilizing *in silico* approaches, enzymes were identified and engineered for efficient catalysis of ibuprofen aldehyde to ibuprofen. Among the enzymes screened, *bc*PADH demonstrated the highest conversion efficiency, making it a promising candidate for biocatalytic applications. However, higher substrate concentration and increased solvent adversely affected enzyme performance, highlighting the limitations of the wild-type enzyme under industrial conditions. To overcome these challenges, we employed *in silico* rational enzyme engineering, specifically targeting solvent stability. Strategic mutations at the dimeric interface were designed and validated to enhance the stability of the enzyme in the presence of higher volumes of organic solvents. Overall, our study provides a comprehensive chemoenzymatic pathway for the sustainable synthesis of ibuprofen. This approach not only aligns with the principles of green chemistry but also demonstrates the potential of integrating computational and experimental methodologies for enzyme engineering, paving the way for more environmentally friendly pharmaceutical manufacturing processes.

Conflicts of interest

There are no conflicts of interest to declare.

Data availability

The data supporting this article have been included as part of the SI. Fig. S1: CG Mapping schemes for hexane, DMSO and toluene. Bead positions are determined by the center of mass of atoms in it Fig. S2: CG Mapping schemes for protein backbone as a dimer Fig. S3: A representative (A) RP-HPLC chromatogram and (B) NMR data demonstrating the purity of ibuprofen product formed by biocatalytic synthesis from ibuprofen aldehyde Fig. S4: Radius of gyration for (a) bcPADH, (b) bcPADH01, and (c) bcPADH04. The black and red lines represent simulations with 10% and 30% DMSO, respectively, green and blue lines show simulations with 10% and 30% hexane, while orange and magenta lines depict simulations with 10% and 30% toluene as labeled in figure Fig. S5: Number of solvent molecules within 5Å of enzyme for (a) bcPADH, (b) bcPADH01, and (c) bcPADH04. The black and red lines represent simulations with 10% and 30% DMSO, respectively, green and blue lines show simulations with 10% and 30% hexane, while orange and

magenta lines depict simulations with 10% and 30% toluene as labeled in figure Table S1: Binding energy of 22 enzymes along with cluster population and reactive distance between NAD-carbonyl carbon Table S2: List of residues in the interface of WT bcPADH that positively contributed to destabilizing the multimeric form Enzyme residues and its possible mutation Table S3: Summary of interaction energy between each monomeric unit of WT bcPADH tetramer structure and its respective variants from MM-PBSA analysis. Each monomeric unit is labeled as A, B, C and D. Energy is in kcal mol^{-1} Table S4: Summary of total interaction energy and contribution of the hotspot residues to the interaction energy of WT bcPADH tetramer structure and its respective variants from MM-PBSA analysis. See DOI: https://doi.org/10.1039/d5su00073d.

References

1 K. D. Rainsford, History and Development of Ibuprofen, in *Ibuprofen*, 2015, p. 1–21, DOI: 10.1002/9781118743614.ch1.
2 F. D. Hart and E. C. Huskisson, Non-Steroidal Anti-Inflammatory Drugs, *Drugs*, 1984, **27**(3), 232–255, DOI: 10.2165/00003495-198427030-00004.
3 K. D. Rainsford, Ibuprofen: from invention to an OTC therapeutic mainstay, *Int. J. Clin. Pract.*, 2013, **67**(s178), 9–20, DOI: 10.1111/ijcp.12055.
4 A. Upadhyay, A. Amanullah, V. Joshi, R. Dhiman, V. K. Prajapati, K. M. Poluri, *et al.*, Ibuprofen-based advanced therapeutics: breaking the inflammatory link in cancer, neurodegeneration, and diseases, *Drug Metab. Rev.*, 2021, **53**(1), 100–121, DOI: 10.1080/03602532.2021.1903488.
5 K. D. Rainsford, Ibuprofen: pharmacology, efficacy and safety, *Inflammopharmacology*, 2009, **17**(6), 275–342, DOI: 10.1007/s10787-009-0016-x.
6 Straits Research, *Ibuprofen Market*, 2024, available from: https://straitsresearch.com/report/ibuprofen-market/.
7 A. Rastogi, M. K. Tiwari and M. M. Ghangrekar, A review on environmental occurrence, toxicity and microbial degradation of Non-Steroidal Anti-Inflammatory Drugs (NSAIDs), *J. Environ. Manage.*, 2021, **300**, 113694.
8 E. Brillas, A critical review on ibuprofen removal from synthetic waters, natural waters, and real wastewaters by advanced oxidation processes, *Chemosphere*, 2022, **286**, 131849, DOI: 10.1016/j.chemosphere.2021.131849.
9 A. Ivanković, A. Dronjić, A. Martinović Bevanda and S. Talić, Review of 12 Principles of Green Chemistry in Practice, *Int. J. Sustain. Green Energy*, 2017, **6**, 39–48.
10 M. A. Murphy, Early Industrial Roots of Green Chemistry and the history of the BHC Ibuprofen process invention and its Quality connection, *Found. Chem.*, 2018, **20**(2), 121–165, DOI: 10.1007/s10698-017-9300-9.
11 F. Grimaldi, N. N. Tran, M. M. Sarafraz, P. Lettieri, O. M. Morales-Gonzalez and V. Hessel, Life Cycle Assessment of an Enzymatic Ibuprofen Production Process with Automatic Recycling and Purification, *ACS Sustain. Chem. Eng.*, 2021, **9**(39), 13135–13150, DOI: 10.1021/acssuschemeng.1c02309.
12 R. A. Sheldon and J. M. Woodley, Role of Biocatalysis in Sustainable Chemistry, *Chem. Rev.*, 2018, **118**(2), 801–838, DOI: 10.1021/acs.chemrev.7b00203.
13 M. T. Reetz, Biocatalysis in Organic Chemistry and Biotechnology: Past, Present, and Future, *J. Am. Chem. Soc.*, 2013, **135**(34), 12480–12496, DOI: 10.1021/ja405051f.
14 S. Kar, H. Sanderson, K. Roy, E. Benfenati and J. Leszczynski, Green Chemistry in the Synthesis of Pharmaceuticals, *Chem. Rev.*, 2022, **122**(3), 3637–3710, DOI: 10.1021/acs.chemrev.1c00631.
15 P. N. Devine, R. M. Howard, R. Kumar, M. P. Thompson, M. D. Truppo and N. J. Turner, Extending the application of biocatalysis to meet the challenges of drug development, *Nat. Rev. Chem.*, 2018, **2**(12), 409–421, DOI: 10.1038/s41570-018-0055-1.
16 P. V. Iyer and L. Ananthanarayan, Enzyme stability and stabilization—Aqueous and non-aqueous environment, *Process Biochem.*, 2008, **43**(10), 1019–1032, DOI: 10.1016/j.procbio.2008.06.004.
17 R. Chowdhury and C. D. Maranas, From directed evolution to computational enzyme engineering—A review, *AIChE J.*, 2020, **66**(3), e16847, DOI: 10.1002/aic.16847.
18 Q. Liu, G. Xun and Y. Feng, The state-of-the-art strategies of protein engineering for enzyme stabilization, *Biotechnol. Adv.*, 2019, **37**(4), 530–537, DOI: 10.1016/j.biotechadv.2018.10.011.
19 F. Rigoldi, S. Donini, A. Redaelli, E. Parisini and A. Gautieri, Review: Engineering of thermostable enzymes for industrial applications, *APL Bioeng.*, 2018, **2**(1), 11501, DOI: 10.1063/1.4997367.
20 A. M. Evans, Comparative Pharmacology of S(+)-Ibuprofen and (RS)-Ibuprofen, *Clin. Rheumatol.*, 2001, **20**(1), 9–14, DOI: 10.1007/BF03342662.
21 X. G. Zhao, D. Z. Wei and Q. X. Song, A facile enzymatic process for the preparation of ibuprofen ester prodrug in organic media, *J. Mol. Catal. B Enzym.*, 2005, **36**(1), 47–53, DOI: 10.1016/j.molcatb.2005.08.005.
22 M. Ravelo, M. E. Gallardo, M. Ladero and F. Garcia-Ochoa, Synthesis of Ibuprofen Monoglyceride Using Novozym®435: Biocatalyst Activation and Stabilization in Multiphasic Systems, *Catalysts*, 2022, **12**, 1531, DOI: 10.3390/catal12121531.
23 L. Marc, S. Guillemer, J. M. Schneider and G. Coquerel, Continuous chiral resolution of racemic Ibuprofen by diastereomeric salt formation in a Couette-Taylor crystallizer, *Chem. Eng. Res. Des.*, 2022, **178**, 95–110, DOI: 10.1016/j.cherd.2021.12.016.
24 W. S. Long, A. Kamaruddin and S. Bhatia, Chiral resolution of racemic ibuprofen ester in an enzymatic membrane reactor, *J. Membr. Sci.*, 2005, **247**(1), 185–200.
25 S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, *et al.*, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.*, 1997, **25**(17), 3389–3402, DOI: 10.1093/nar/25.17.3389.
26 I. Schomburg, A. Chang, C. Ebeling, M. Gremse, C. Heldt, G. Huhn, *et al.*, BRENDA, the enzyme database: updates

and major new developments, *Nucleic Acids Res.*, 2004, **32**(1), D431–D433, DOI: **10.1093/nar/gkh081**.

27 E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, *et al.*, CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures, *PLoS Comput. Biol.*, 2012, **8**(10), e1002708, DOI: **10.1371/journal.pcbi.1002708**.

28 N. Eswar, D. Eramian, B. Webb, M. Y. Shen and A. Sali, in *Protein Structure Modeling with MODELLER BT - Structural Proteomics: High-Throughput Methods*, ed. Kobe B., Guss M. and Huber T., Totowa, NJ, Humana Press, 2008, pp. 145–159, DOI: **10.1007/978-1-60327-058-8_8**.

29 B. Webb and A. Sali, Comparative Protein Structure Modeling Using MODELLER, *Curr. Protoc. Bioinf.*, 2016, **54**(1), 561–5637, DOI: **10.1002/cpbi.3**.

30 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, *J. Cheminf.*, 2012, **4**(1), 17, DOI: **10.1186/1758-2946-4-17**.

31 O. Trott and A. J. Olson, AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.*, 2010, **31**(2), 455–461, DOI: **10.1002/jcc.21334**.

32 A. Ashokan and G. K. Aradhyam, Chapter 9 - Effective purification of recombinant peptide ligands for GPCR research, in *G Protein-Coupled Receptors Part A*, ed. Shukla A. K., Academic Press, 2017, pp. 111–118, DOI: **10.1016/bs.mcb.2017.07.001**.

33 M. M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.*, 1976, **72**, 248–254.

34 K. Shortall, A. Djeghader, E. Magner and T. Soulimane, Insights into Aldehyde Dehydrogenase Enzymes: A Structural Perspective, *Front. Mol. Biosci.*, 2021, **8**, 659550.

35 J. S. Rodriguez-Zavala and H. Weiner, Structural Aspects of Aldehyde Dehydrogenase that Influence Dimer–Tetramer Formation, *Biochemistry*, 2002, **41**(26), 8229–8237, DOI: **10.1021/bi012081x**.

36 Y. Qi, H. I. Ingólfsson, X. Cheng, J. Lee, S. J. Marrink and W. Im, CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field, *J. Chem. Theory Comput.*, 2015, **11**(9), 4486–4494, DOI: **10.1021/acs.jctc.5b00513**.

37 T. Aoyagi, F. Sawa, T. Shoji, H. Fukunaga, J. I. Takimoto and M. Doi, A general-purpose coarse-grained molecular dynamics program, *Comput. Phys. Commun.*, 2002, **145**(2), 267–279, DOI: **10.1016/S0010-4655(02)00271-0**.

38 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem.*, 1983, **4**(2), 187–217, DOI: **10.1002/jcc.540040211**.

39 J. E. Basconi and M. R. Shirts, Effects of Temperature Control Algorithms on Transport Properties and Kinetics in Molecular Dynamics Simulations, *J. Chem. Theory Comput.*, 2013, **9**(7), 2887–2899, DOI: **10.1021/ct400109a**.

40 D. Quigley and M. I. J. Probert, Langevin dynamics in constant pressure extended systems, *J. Chem. Phys.*, 2004, **120**(24), 11432–11441, DOI: **10.1063/1.1755657**.

41 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, *et al.*, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX*, 2015, **1–2**, 19–25, DOI: **10.1016/j.softx.2015.06.001**.

42 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, GROMACS: Fast, flexible, and free, *J. Comput. Chem.*, 2005, **26**(16), 1701–1718, DOI: **10.1002/jcc.20291**.

43 V. N. Maiorov and G. M. Crippen, Significance of Root-Mean-Square Deviation in Comparing Three-dimensional Structures of Globular Proteins, *J. Mol. Biol.*, 1994, **235**(2), 625–634, DOI: **10.1006/jmbi.1994.1017**.

44 E. Durham, B. Dorr, N. Woetzel, R. Staritzbichler and J. Meiler, Solvent accessible surface area approximations for rapid and accurate protein structure prediction, *J. Mol. Model.*, 2009, **15**(9), 1093–1108, DOI: **10.1007/s00894-009-0454-9**.

45 M. Y. Lobanov, N. S. Bogatyreva and O. V. Galzitskaya, Radius of gyration as an indicator of protein structure compactness, *Mol. Biol.*, 2008, **42**(4), 623–628, DOI: **10.1134/S0026893308040195**.

46 R. Kumari, R. Kumar and A. Lynn, g_mmpbsa—A GROMACS Tool for High-Throughput MM-PBSA Calculations, *J. Chem. Inf. Model.*, 2014, **54**(7), 1951–1962, DOI: **10.1021/ci500020m**.

47 A. Tripathi and K. Dubey, Combined MD and QM/MM Calculations Reveal Allostery-Driven Promiscuity in Dipeptide Epimerases of Enolase Family, *Chem.–Asian J.*, 2022, **17**(16), e202200528, DOI: **10.1002/asia.202200528**.

48 P. Junhui, W. Wang, Y. Q. Yu, H. L. Gu and X. Huang, Clustering algorithms to analyze molecular dynamics simulation trajectories for complex chemical and biological systems, *Chin. J. Chem. Phys.*, 2018, **31**, 404–420.