


RESEARCH

Open Access



Atomevo: a web server combining protein modelling, docking, molecular dynamic simulation and MMPBSA analysis of *Candida antarctica* lipase B (CaLB) fusion protein

Jin-Heng Hao^{1,2†}, Dun-Jin Zheng^{1,2†}, Yu-Hao Ye^{3†}, Jie-Ting Yu^{1,2}, Xin-Yao Li^{1,2}, Mei-Jie Xiong¹, Wen-Hao Jiang¹, Kang-Ping He¹, Pei-Yu Li¹, Yong-Si Lv¹, Wei-Ming Gu¹, Lin-Hao Lai¹, Yi-Da Wu¹ and Shi-Lin Cao^{1,2,3*} 

Abstract

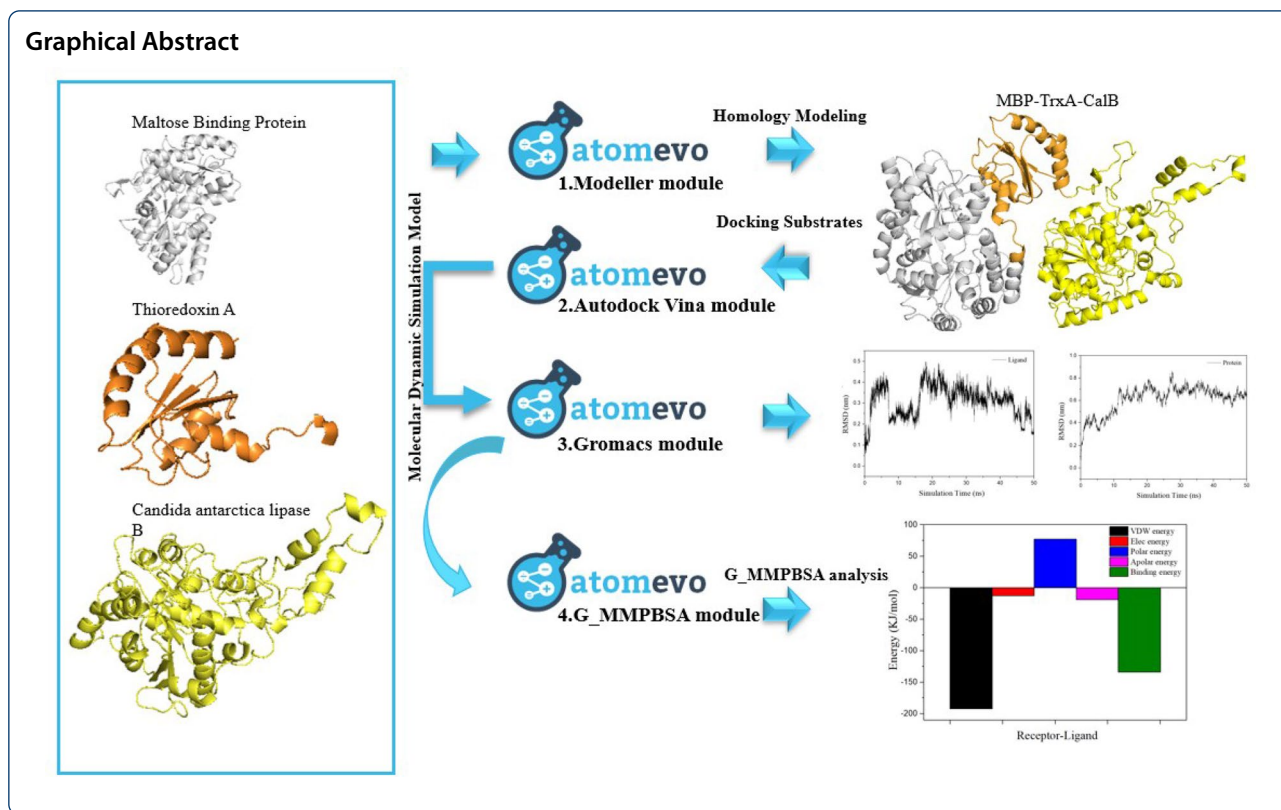
Although current computational biology software is available and has prompted the development of enzyme–substrates simulation, they are difficult to install and inconvenient to use. This makes the time-consuming and error-prone process. By far there is still a lack of a complete tool which can provide a one-stop service for the enzyme–substrates simulation process. Hence, in this study, several computational biology software was extended development and integrated as a website toolbox named Atomevo. The Atomevo is a free web server providing a user-friendly interface for enzyme–substrates simulation: (1) protein homologous modeling; (2) parallel docking module of Auto-dock Vina 1.2; (3) automatic modeling builder for Gromacs molecular dynamics simulation package; and (4) Molecular Mechanics/Poisson–Boltzmann Surface Area (MMPBSA) analysis module for receptor–ligand binding affinity analysis. We officially launched the web server and provided instructions through a case for the design and simulation of *Candida antarctica* lipase B (CaLB) fusion protein called Maltose Binding Protein—Thioredoxin A—*Candida antarctica* lipase B (MBP-TrxA-CaLB).

Keywords: Biofuel, Molecular dynamics simulation, Docking, MMPBSA, Web Server

[†]Jin-Heng Hao, Dun-Jin Zheng and Yu-Hao Ye are co-first author and contributed equally in this paper

*Correspondence: seelam.shilin.cao@gmail.com; shilin.cao@fosu.edu.cn

¹ Guangdong Key Laboratory of Food Intelligent Manufacturing, School of Food Science and Technology, Foshan University, Foshan 528000, China
Full list of author information is available at the end of the article



Introduction

Enzymes, such as lipase, exhibit great potential in bioenergy and biomedicine and are widely used for biofuel production (Fatma et al. 2021). However, there is an increasing demand for efficient enzymes for practical applications. Over the past decades, an increasing number of computational methods (Zhang et al. 2020; Xu et al. 2021) and software have been developed and exhibit great potential in the investigation of the structure and function relationship of enzymes (Rester 2008; Baek et al. 2021; Jumper et al. 2021) and rational and semirational design of enzyme catalysts (Liu et al. 2021). Homology modelling, molecular docking (Eberhardt et al. 2021) and molecular dynamics simulation (Phillips et al. 2020) are three important computational biology methods (McGregor et al. 2007).

For example, a typical investigation of enzyme-substrate interaction includes the following processes: (1) building the protein structure; (2) docking of candidate substrate molecules into the enzyme potential catalytic site and estimating and ranking the binding affinity between the substrate and enzyme to obtain potential drug candidates; (3) performing molecular dynamics (MD) simulation between the substrate and enzyme, according to the docking results; and (4) performing the Molecular Mechanics/Poisson-Boltzmann

Surface Area (MMPBSA) analysis to further confirm the enzyme/substrate interaction. Up to now, several molecular docking tools (AutoDock-Vina (Eberhardt et al. 2021), Ledock (Wang et al. 2016), SwissDock (Grosdidier et al. 2011), Schrödinger (Bhachoo and Beuming 2017), etc.) and molecular dynamics simulation (gromacs (Berendsen et al. 1995), NAMD (Nelson et al. 1996; Phillips et al. 2020), lammps (Thompson et al. 2022)) software have been developed. However, most software needs to be run on a Linux command-line system, which raises the bar for users' software debugging and installation capabilities. Although some software such as SwissDock (Grosdidier et al. 2011) can perform online docking tasks through web servers on the window, it can only provide a single molecular docking service and cannot complete a whole set of virtual screening. DS (Valentine et al. 2021) and Schrödinger (Bhachoo and Beuming 2017) can be used under a Windows operating system. They are convenient and simple in interface, but they are costly and thus not accessible to students or laboratories with insufficient funds. Thus, a free web server that provides molecular docking, molecular dynamics simulation, and relative data analysis is a great tool making enzyme rational and semirational design more accessible to the scientific community.

Herein, several academic free software for predicting the protein structure, enzyme–substrate molecular docking, and molecular dynamic simulation were extended development. The Atomevo web server is a free web server providing a user-friendly interface for (1) protein homologous (Modelling Module); (2) parallel docking module of Autodock Vina (Vina Module); (3) automatic modeling builder for Gromacs molecular dynamics simulation package (GMX Module); and (4) Molecular Mechanics Poisson–Boltzmann Surface Area (MMPBSA) analysis module for receptor–ligand binding affinity analysis (G_MMPBSA Module).

In this study, we officially launched the web server and provided instructions through a case description for (1) structure modelling of the *Candida antarctica* lipase B (CalB) fusion protein (MBP-TrxA-CalB), which fused with maltose binding protein and thioredoxin A (TrxA) to improve the CalB expression level in *Escherichia coli* (*E. coli*) by Atomevo-Modelling Module; (2) molecular docking between MBP-TrxA-CalB (MT-CalB) and fatty acid small molecules by Atomevo-Vina module; (3) molecular dynamic simulation of MT-CalB/fatty acid complex via Atomevo-GMX module; and (4) MMPBSA analysis between MT-CalB and fatty acid via Atomevo-G_MMPBSA module. Finally, the MT-CalB was expressed in *E. coli* and the enzyme properties were characterized. Generally speaking, this platform provides a user-friendly, comprehensive, and flexible tool for protein modelling, docking, molecular dynamic simulation to facilitate in the exploration of enzyme (such as lipase) structure and function relationships, as well as the rational and semi-rational design of enzyme catalysts.

Materials and methods

The program and modules used in Atomevo platform

The modules and software used in this platform are all open source projects or academic free software, and the modules are secondary-developed. The modules used mainly include: (1) molecular docking software: AutoDock Vina (Linux 1.2 version, open source project) (Eberhardt et al. 2021); molecular Docking scoring software XScore (academic free, non-open source project) (Wang et al. 2002) protein–ligand interaction analysis tool Protein–Ligand Interaction Profiler (academic free, open source project, <https://github.com/ssalentin/plip>) (Salentin et al. 2015); chemical toolbox openbabel (open source project, <https://github.com/openbabel/openbabel>) (O’Boyle et al. 2011); GROMACS2019.1. (open source project) (Berendsen et al. 1995).

Implementation of Atomevo platform

The Atomevo web server consists of the user interface, the local server and the job backend. The user interface is implemented in <https://atomevo.com>. This interface allows user to upload the input files and the configure parameters (or files), as well as download the output files. The normal server is deployed in Ali Cloud Server and the advance server is deployed in our local server with Intel(R) Xeon(R) CPU E5-2690 v3, MSI RTX 2080 with 8 GB of RAM, MD1200 storage devices 72 TB and 64 GB RAM.

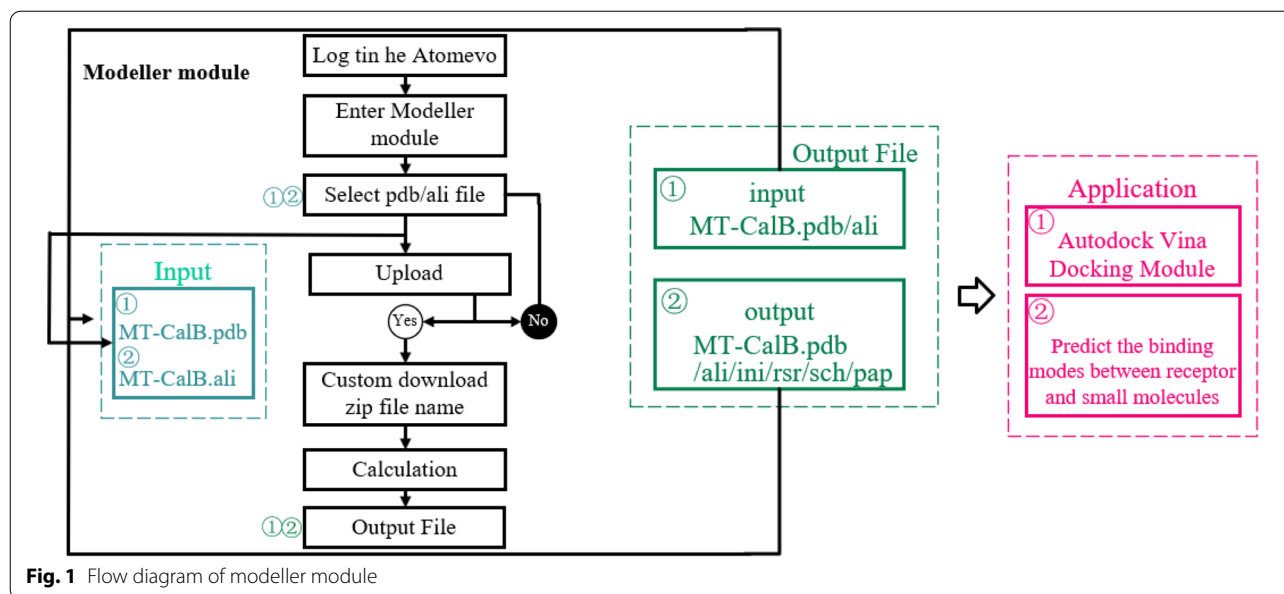
The job backend is developed using PHP technology. Thinkphp framework is used as a backend development framework and Medoo as a database framework. Axios is used to send the request to the background and Workerman is used to build the instant messaging framework. Vue.js is used to build progressive framework and Wepack module loader is used to convert the dependent module to static files which represents these packages. PhpSpreadsheet, a plug-in, is used to score statistics, and Handsontable, a tabular interactive plug-in in the footer, is used to organize, count, calculate and summarize the tables.

Implementation of protein homologous modeller module

We have created the Modeller module on the Atomevo platform to perform Homology Modeling online (<https://atomevo.com/modeller>).

The algorithm of the Modeller module includes several steps (Fig. 1): (1) input files (*.pdb, *.ali) for Modeller single-template modeling module calculation; (2) performing the Modeller calculation and obtaining the calculation results file (*.pdb, *.ali, *.ini, *.rsr, *.sch and *.pap files); and (3) output files and provide them to users for download.

The Modeller module is divided into three sub-modules: single-template modelling module, multi-template modelling module, and sequence alignment module. Its workflow includes the following steps (take a single-template modelling module as an example): (1) the user uploads the hydrogenated pdb file and the template ali file and submits a calculation task to the server; (2) after receiving the calculation request submitted by the user, the server creates a new calculation process and calls the calculation script of the Modeller single-template modelling module to calculation for protein homologous; and (3) when the calculation is completed, the server terminates the calculation process and moves the output file to a folder named after the task number, packs the folder, and sends the calculation file and task completion reminder mail to the user through the phpmailer module.



Implementation of the molecular docking module

To perform virtual screening of lipase by parallel module of Autodock Vina online, we have established the Autodock Vina module on the platform (<https://atomevo.com/autodock-vina>).

The Autodock Vina module includes two parts: Vina docking module and result analysis module.

The algorithm of the Autodock Vina module includes several steps (Fig. 2): (1) creating the vina configure file (conf.txt) by reading the parameters in the configure file (*.xlsx); (2) converting input file (*.pdb) to *.pdbqt; (3) making separate directories for AutoDock vina calculation; (4) running docking (vina -configure vina-dock.conf); and (5) analyzing the result.

Results analysis module is showed as follows:

(1) Vina result analysis: the multimodel pdbqt files (*_out.pdbqt) are split into individual models files (*_out*.pdbqt) using the vina_split command. The docking result of each individual model (including affinity (kcal/mol), dist from rmsd rmsd l.b., best mode rmsd u.b.) is exported through phpexcel plug-in into the vina-score file (vina-score.xlsx).

(2) X-Score analysis: first, the individual models files (*_out*.pdbqt) are converted to *_out*.mol2 by openbabel software; second, the ligand files (*_out*.mol2) and the receptor files (*.pdb) are submitted to the X-Score software (Wang et al. 2002). X-Score is a "Scoring function," which has its major applications to structure-based drug design studies. It computes a binding score for a given protein–ligand complex structure, and this binding score correlates to experimental binding constants well. Three individual

empirical scoring results, including HPScore, HMScore, and HSScore are exported through phpexcel plug-in into the ledock-score file (ledock-score.xlsx).

The computing platform is divided into Molecular preprocessing module, docking computing module, Grading module, and protein–ligand interaction analysis module.

The molecular preprocessing module mainly consists of the Openbabel module. Its workflow includes the following steps: (1) the user sends a compute request to the background through Axios from the client. Then the structure file needed to be converted should be uploaded to the server. And then, the user can select the output file type and output settings. (2) The Thinkphp framework is called the Openbabel software for calculation when the server receives the request. And finally, the server tidies up the compressed output file through Linux's compressed packaging commands.

The Docking computing module's workflow includes several steps: (1) the client is used to submit computational requests through the network interface and upload the relevant ligands, receptor structures, and docking condition profiles; (2) when the server receives the request, the server runs the modules of AutoDock Vina for docking calculation; and (3) the Linux operating system's file consolidation functions are used to settle the input file, output parameter file and pigeonhole to Input file, Output file, Parameter and Result separately.

The implementation method of molecular docking parallel computing: split the execution task of autodock_vina.php and create the Child process file named autodock_vina_chind.php. The Child process

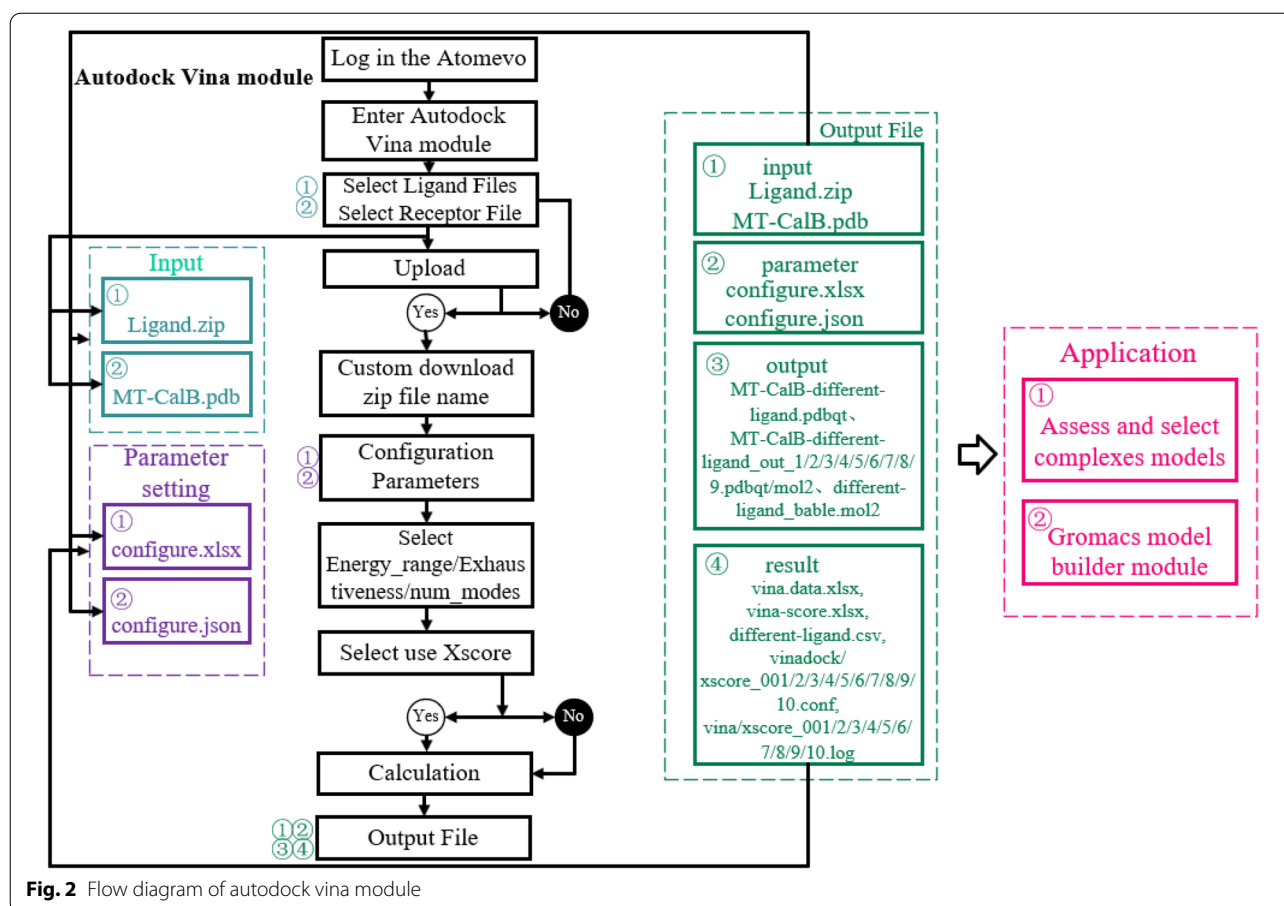


Fig. 2 Flow diagram of autodock vina module

is in charge of the calculating and scoring operations between a single ligand and receptor. The result will be stored in the Redis. The calculation between a ligand and a receptor is performed in this file. When a new task is received, autodock_vina.php will use Popen (a php function) to perform autodock_vinachind.php subprocess file and return a file pointer to Parent process autodock_vina.php. The while loop condition formulation is used to wait for the end of Child process, and Fget (a php function) is used to determine if it is false and Pclose (a php function) is used to close all the completed subprocess files. Then the Child process continues to open until all the tasks in the configure.xlsx are completed. When all the Child processes are completed, the Parent process reads the data in Redis module to perform a series of operations, such as summarizing the data.

The Grading module's workflow includes several steps: (1) the Grading module invokes the calculated results of the docking ligands (which requires mol2 format, otherwise use openbabel for conversion) and docking receptor file (which need pdb format, otherwise use openbabel for conversion) in the computing platform

automatically; (2) the XScore is invoked to calculate through the Thinkphp framework and grade the results; and (3) the PhpSpreadsheet and Handsontable plug-in are invoked to summarize, sort out and output the docking scoring results of each group of docking calculations which is from XScore score result and the corresponding calculation module.

Molecular dynamic simulation detail for MT-CalB using Gromacs module

To exam the conformational stability of the receptor–ligand complex, the Gromacs model builder module is deployed on the Atomevo platform (<https://atomevo.com/gromacs>).

The algorithm detail of the Gromacs_model_builder module includes several steps (Fig. 3): (1) creating the Gromacs_model_builder configure file by reading the parameters in the configure.file (*.xlsx); (2) inputting files (*.gro, *.top, *.itp and *.mdp) for Gromacs_model_builder calculation; (3) performing the define box and dolvate, add ions and energy minimization calculation; (4) generating NVT.mdp, NPT.mdp and MD.mpt file for NVT ensemble, NPT ensemble and MD simulation; and (5)

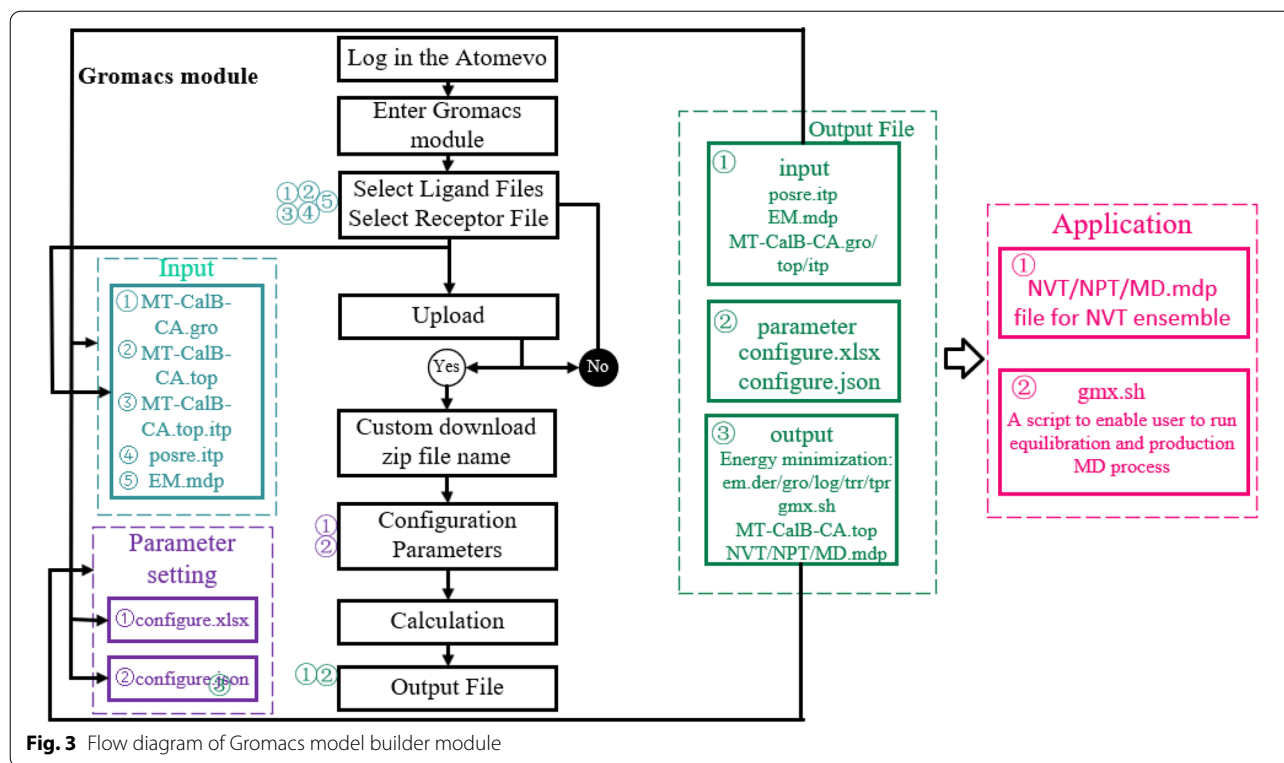


Fig. 3 Flow diagram of Gromacs model builder module

generating a shell script to enable user to run equilibration and production MD process in their own computer.

In this research, the MT-CalB with Chaulmoogric Acid (CA) ligands was placed into the center of a cubic, in which the MT-CalB–ligand complex was 1.5 nm apart from the box margin (with the size of 15.302 11.055 8.518 nm). Then, 15 Na⁺ were added to the box to balance the electrical neutrality of the simulation system.

First, the simulation box was submitted to 50,000 steps of energy minimization (EM) at 303.15 K using Particle-Mesh Ewald method. Then, position-restrained MD simulation was performed to equilibrate the solvent and ions around the ligand–receptor molecule via NVT ensemble (constant Number of particles, Volume, and Temperature), following by NPT ensemble (constant Number of particles, Pressure, and Temperature) and 50 ns production MD.

Implementation of the G_MMPBSA module

To execute *g_mmpbsa* calculations on the assembly model of receptor and ligands online, we have established the *g_mmpbsa* module on the Atomevo (<https://atomevo.com/g-mmpbsa>).

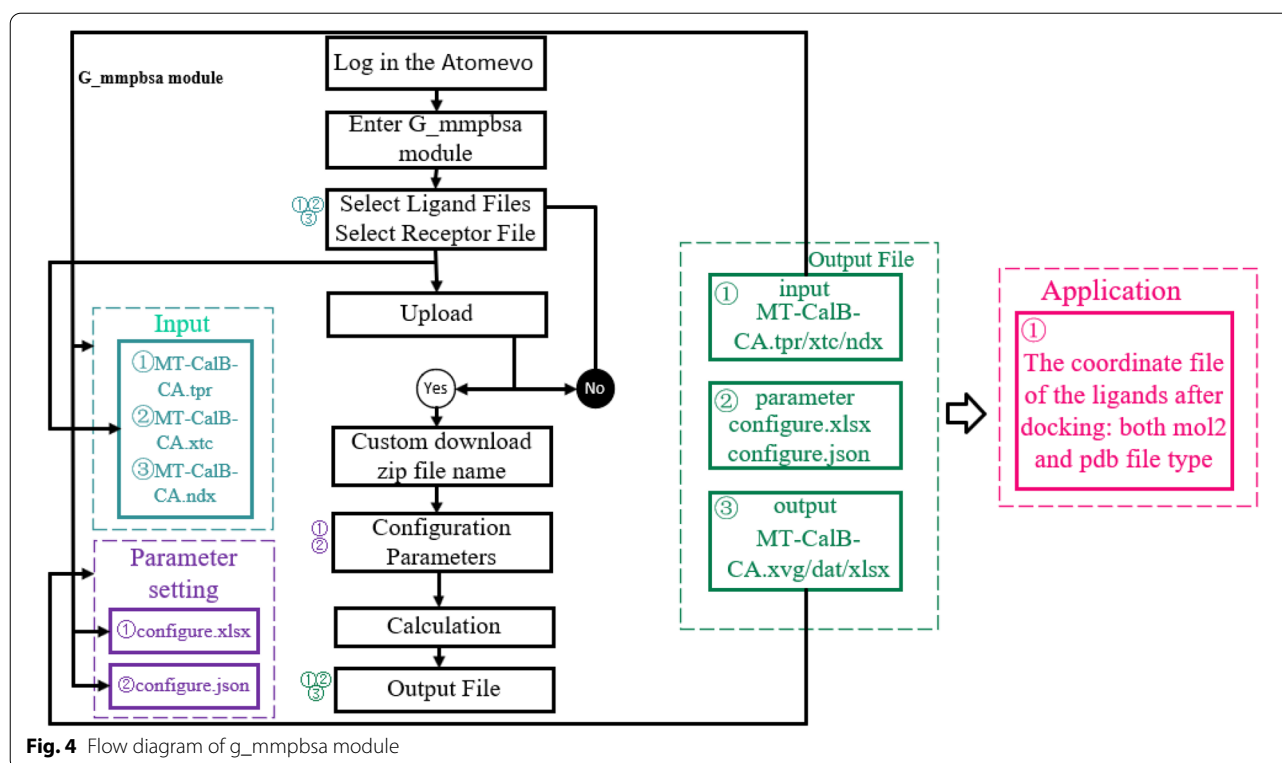
Algorithm details: *g_mmpbsa* module includes several steps (Fig. 4): (1) creating the *g_mmpbsa* configure.file by reading the parameters in the configure.file (*.xlsx); (2) inputting files (*.tpr, *.xtc, and *.ndx) for *g_mmpbsa*

calculation; (3) performing the *g_mmpbsa* calculation and obtaining the calculation results file (*.xvg and *.dat files); and (4) analysing the result *.xvg and *.dat files by performing with two python scripts (*MmPbSaStat.py* and *MmPbSaDecomp.py*) and the results are exported through phpxcel plug-in into the ledock-score file (*g_mmpbsa_analysis.xlsx*).

The Molecular mechanics Poisson–Boltzmann surface area (MMPBSA) analysis was performed by *g_mmpbsa* tool (Kumari et al. 2014). This paper analyzed the last 10 ns of the MT-CalB models that has been balanced.

Experimental verification of design results for MT-CalB

The DNA sequence which encoding the *Mycobacterium smegmatis* porin A (mSpA) signal peptide was fused with both Maltose Binding Protein (MBP) and TrxA domains at the N-terminal of CalB domain, and cloned into the expression vector pETDuet-1 simultaneously to form the recombinant plasmid mSpA-MT-CalB-pETDuet-1. The *E. coli* strain containing mSpA-MT-CalB-pETDuet-1 was inoculated in the Luria–Bertani (LB) fermentation medium and incubated at 37 °C with shaking at 180 rpm, then 0.3 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG) inducer was added, while OD reached 0.6. Then the fermentation was performed at 15 °C for 30 h. The MT-CalB protein was extracted from the bacterial periplasmic space. The lipase activity could be detected



as follows: a given amount of MT-CalB protein was first mixed with 2 mL phosphate buffered saline solution (50 mM, pH 7.0) at 30 °C, and then 1 mL *N*- α -benzoyl D, L-arginine-pyritroaniline solution (3.679 mM) was added to start the reaction. Finally, 5.3 mL anhydrous ethanol was added to terminate the reaction. The release amount of *p*-nitroaniline (pNA) can be determined by 405 nm spectrophotometry.

Results

Design of MT-CalB sequence

We fused the solution-promoting label MBP at the N-terminal for increasing the soluble expression of the lipase. In addition, the TrxA label, which greatly improved proper disulfide bond formation (Manta et al. 2019), was followed by located between the C-terminal of MBP label and N-terminal of the CalB domain.

Homology modeling of MT-CalB structure using Atomevo modeller module

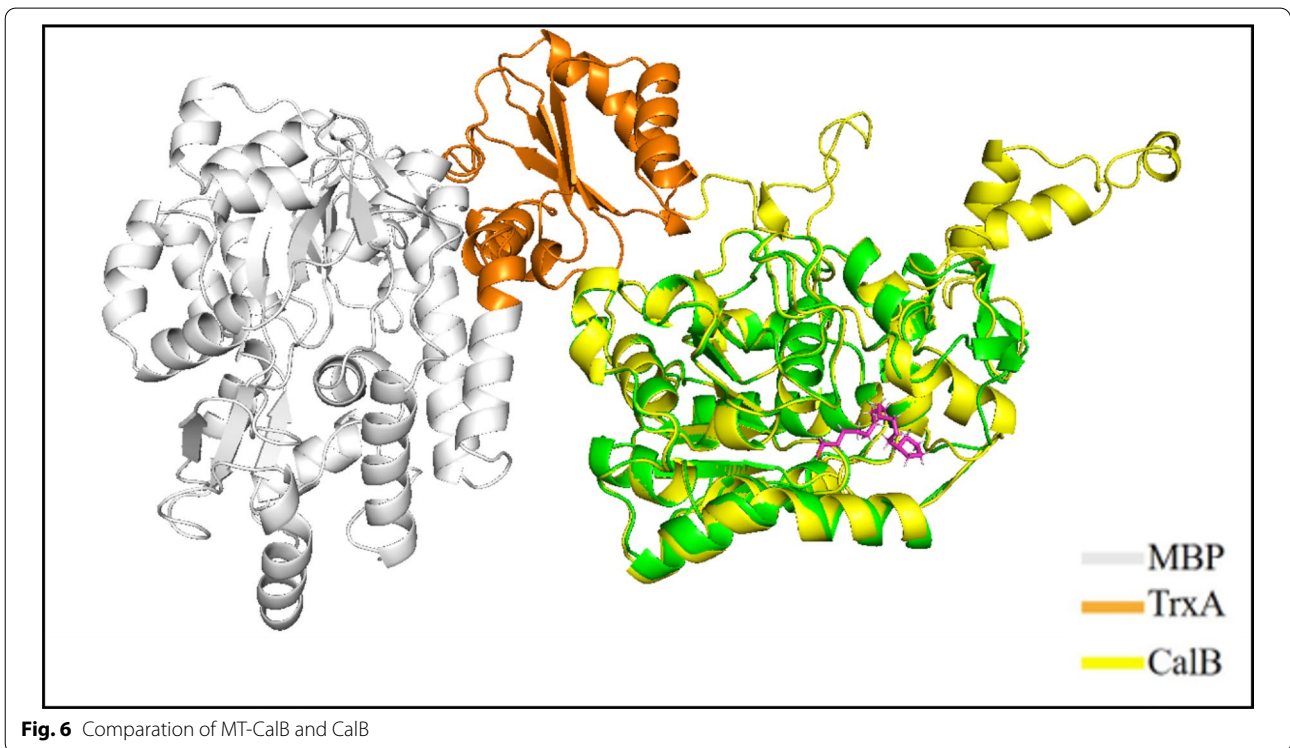
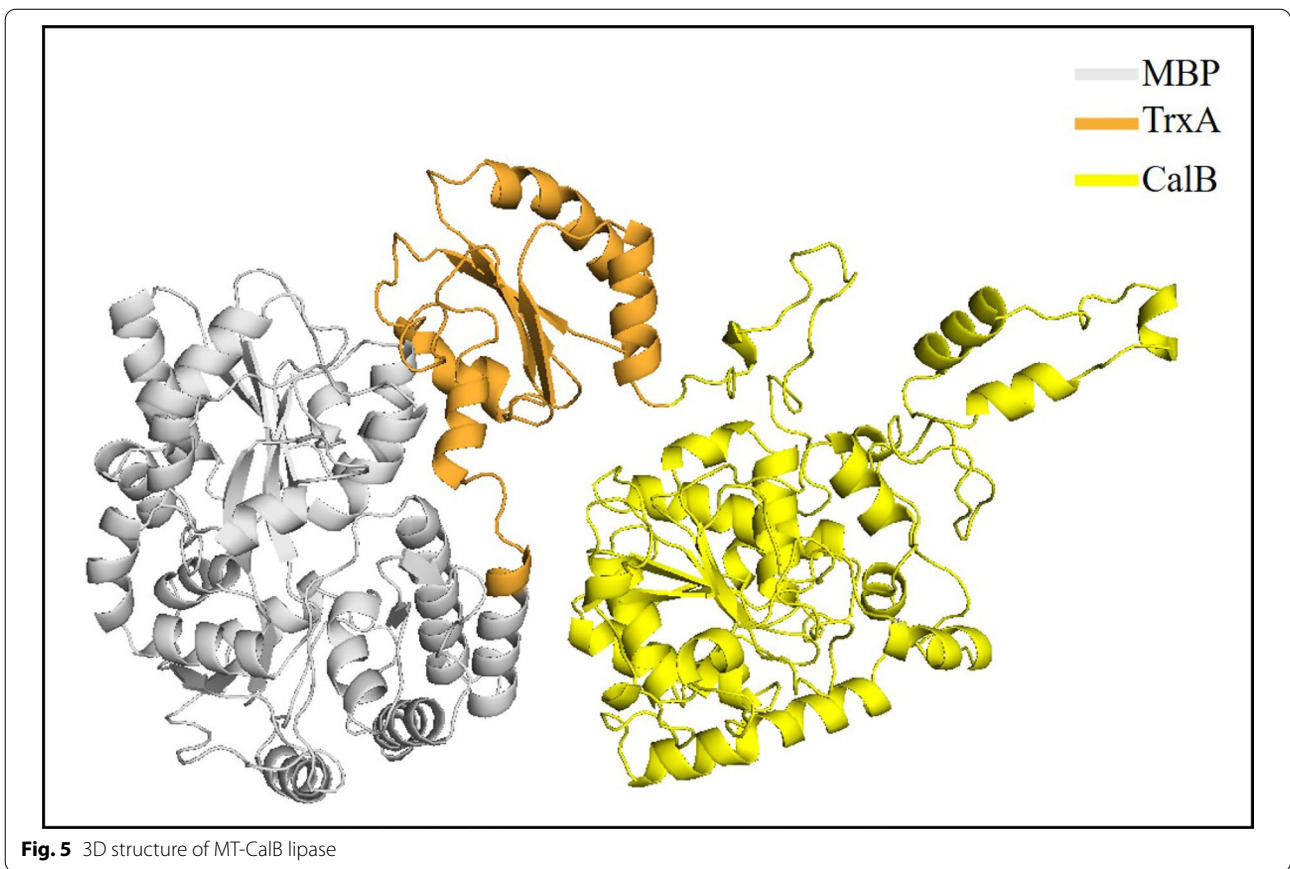
In this example, The MT-CalB structure was designed to increase the CalB expression level in *E. coli* using the Modeller module. The protein sequence of the whole lipase of MT-CalB is represented in Table S1. In addition, the 3D structure of MT-CalB is displayed in Fig. 5, and the MBP, TrxA, and CalB domain are colored in gray, orange, and yellow, respectively. By comparing

with homologous modeling CalB (Fig. 6), it was found that MT-CalB and CalB basically overlapped, the active pocket of MT-CalB was maintained, MBP and TrxA were far away from the active center of CalB, and the overall structure of CalB was not affected.

Docking of MT-CalB and substrates using Atomevo Autodock Vina module

In this example, the MT-CalB was docking with ten fatty acid ligands. Ligands database preparation for docking lipase (MT-CalB) was obtained from PubChem (Additional file 1: Table S2).

The input files pdb-ligands.zip (including *.pdb files for both receptor and ligands) can be submitted by uploading relative files to the system (Additional file 1: Fig. S1A), and the configure parameters can be submitted by uploading configure file (*.xlsx) or inputting in the configure parameter page. In the following configure parameter should be inputted: (1) box parameter (including center of the box($x(-18.1)$, $y(17.8)$, $z(1.1)$) and the size of the box ($x(23.3)$, $y(23.1)$, $z(21)$)) (Additional file 1: Fig. S1B). In addition, their molecular docking pocket is shown in Fig. 7 and (2) the name of the receptor and ligand. After clicking “Perform Calculation”, an ID number is generated for the submitted job. When the calculation is finished, a notification email will be sent to the user’s mailbox. The calculation results and output files



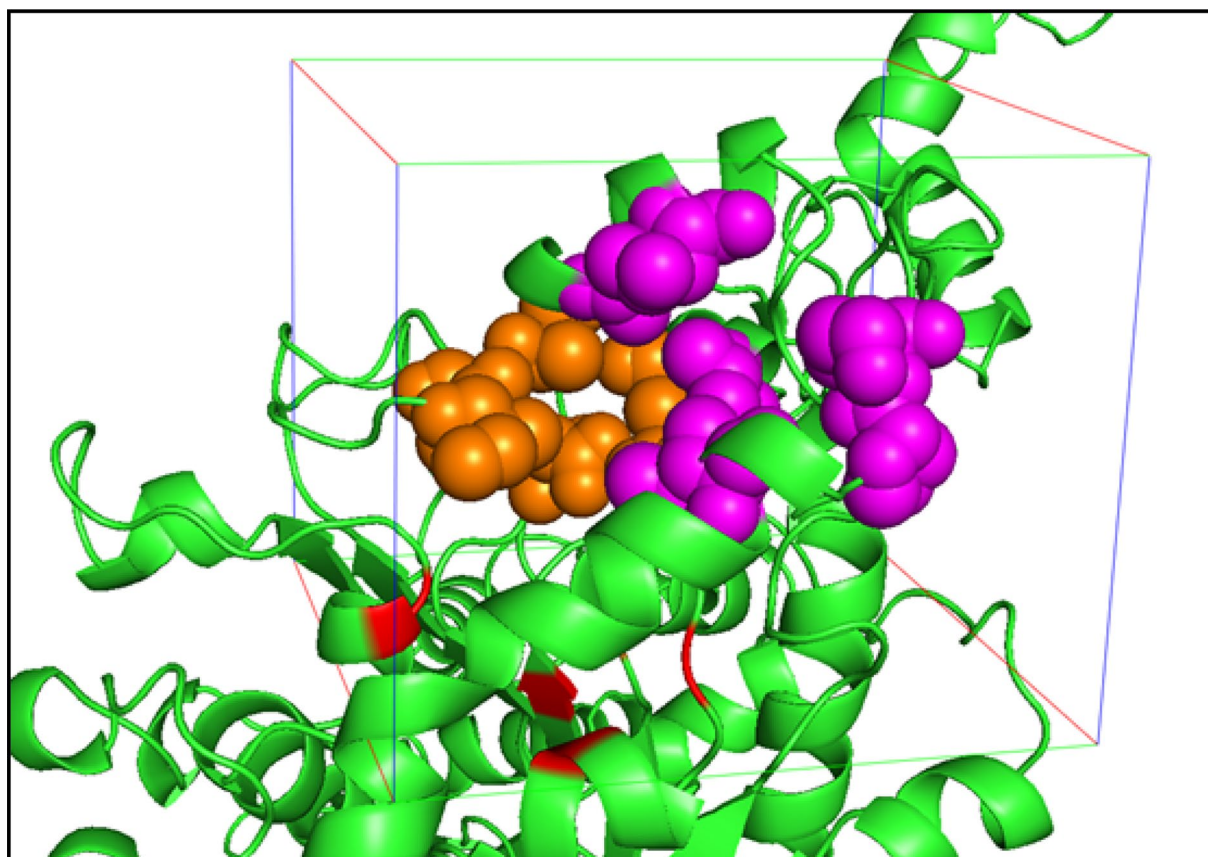


Fig. 7 Molecular docking pocket of MT-CalB

can be downloaded by clicking "Download" in the Auto-dock Vina model.

The output file is an MT-CalB-ligands-allFile.zip file which includes input_file, output_file, parameter file and result file: (1) the input_file includes files in pdb format for receptor and ligands; (2) the output_file includes rec.pdbqt, lig.pdb, lig.pdbqt, lig_out.pdbqt, lig_out*.pdbqt, lig_out*.mol2, lig_babel.mol2; (3) the parameter file includes configure.json and configure.xlsx; and (4) the result file includes vina_**.log, vina_**.conf, vina.xlsx, vina_score.xlsx.

The results are showed in Additional file 1: Table S3. The output docking result files record several data, including ligand name, the best binding affinity and so on. According to the best binding affinity and average binding affinity, Chaulmoogric acid possessed the lowest the best and average binding energy (-5.8 kcal/mol and -5.6 kcal/mol), followed by Docosahexaenoic acid (-5.8 kcal/mol and -5.43 kcal/mol), and Octadecatetraenoic acid (-5.8 kcal/mol and -5.39 kcal/mol). The protein–ligand interactions of MT-CalB with CA are represented in Fig. 8 (using Protein Ligand Interaction

Profiler (PLIP) server), highlighting the key residues involved in the interaction. PLIP analysis result shows that Thr560 Leu660 Ile709, Leu798, Ala801 of MT-CalB and CA exhibit hydrophobic interaction, while there are hydrogen bonding interactions between Thr560, Ser625, Gln626, Gln677 and CA, salt bridge between His744 and CA. This MT-CalB-CA complex model was used as initial structure for the following example about gromacs model builder module.

Building the molecular dynamic simulation model of MT-CalB-substrates model using Gromacs model builder module

The MT-CalB-CA complex is used as a model for gromacs calculation. For example, the input files gmx.zip (including: (1) MT-CalB-CA.gro and MT-CalB-CA.top; (2) posre.itp and posre-CA.itp; and (3) EM.mdp) can be submitted by uploading relative files to the system (Additional file 1: Fig. S2A), and the configure.parameters can be submitted by uploading configure.file (parameter.xlsx) or input in the parameter page (Additional file 1:

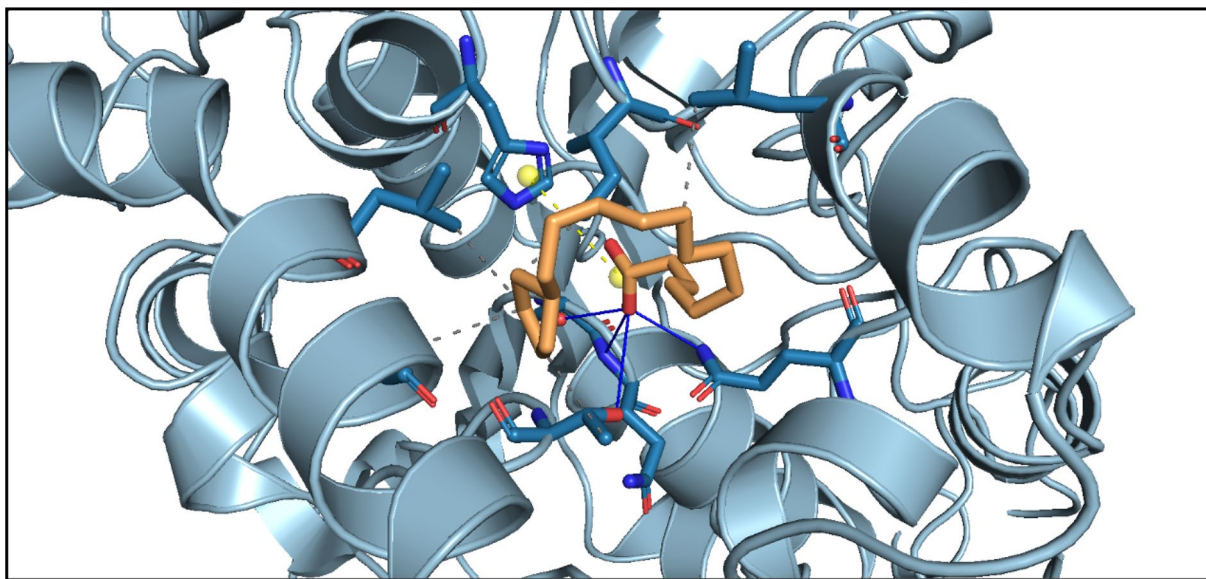


Fig. 8 Protein–ligand interactions of MT-CalB with Chaulmoogric acid

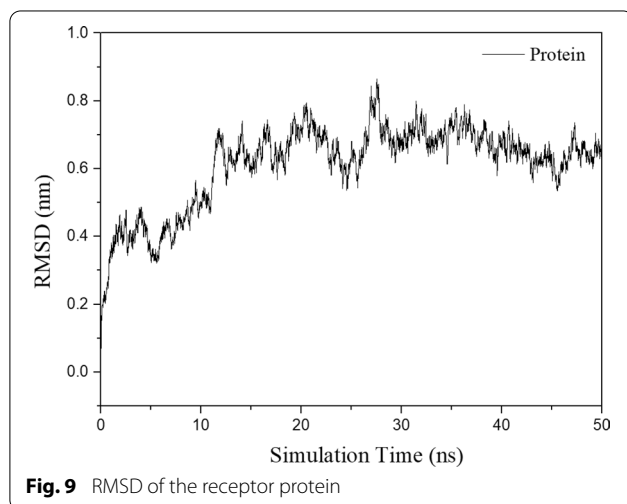


Fig. 9 RMSD of the receptor protein

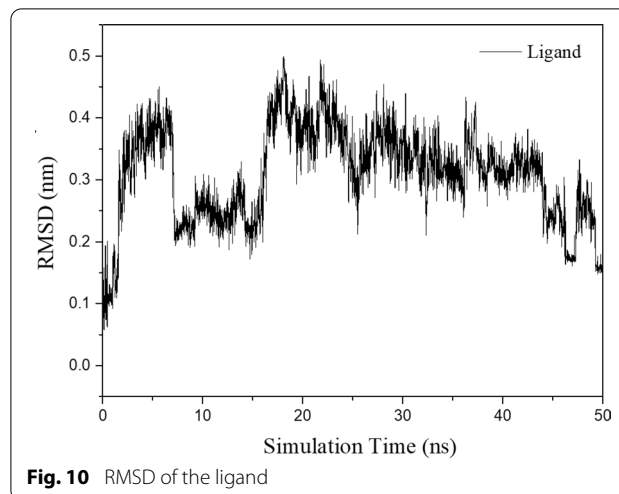


Fig. 10 RMSD of the ligand

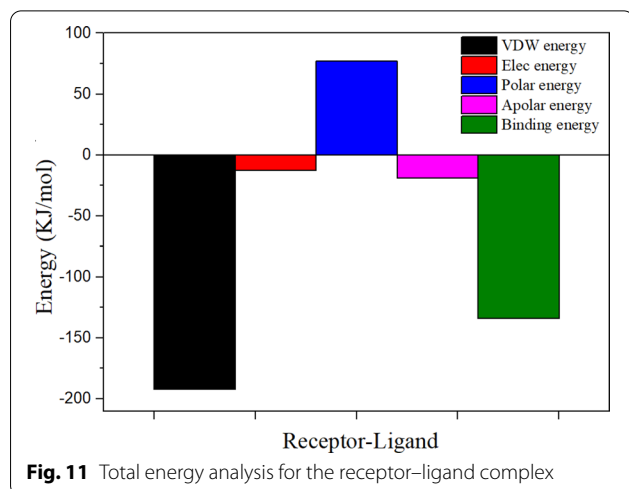
Fig. S2B). After clicking "Perform Calculation", an ID number is generated for the submitted job. When the calculation is finished, a notification email will be sent to the user's mailbox. The calculation results and output files can be downloaded by clicking "Download" in the gromacs module.

The output file is a *.zip file which includes: (1) files after Energy Minimization (em.edr, em.gro, em.log, em.tpr, em.trr, MT-CalB-CA.top); (2) the mdp file for equilibration (NVT.mdp, NPT.mdp) and production MD process (MD.mdp); and (3) a shell script (gmx.sh)

to enable user to run equilibration and production MD process on their own computer.

The RMSD of the receptor protein during the molecular dynamic simulation is shown in Fig. 9. In the complex model, RMSD of MT-CalB-CA are about 0.53–0.80 nm, indicating that protein structures were equilibrium during the simulation.

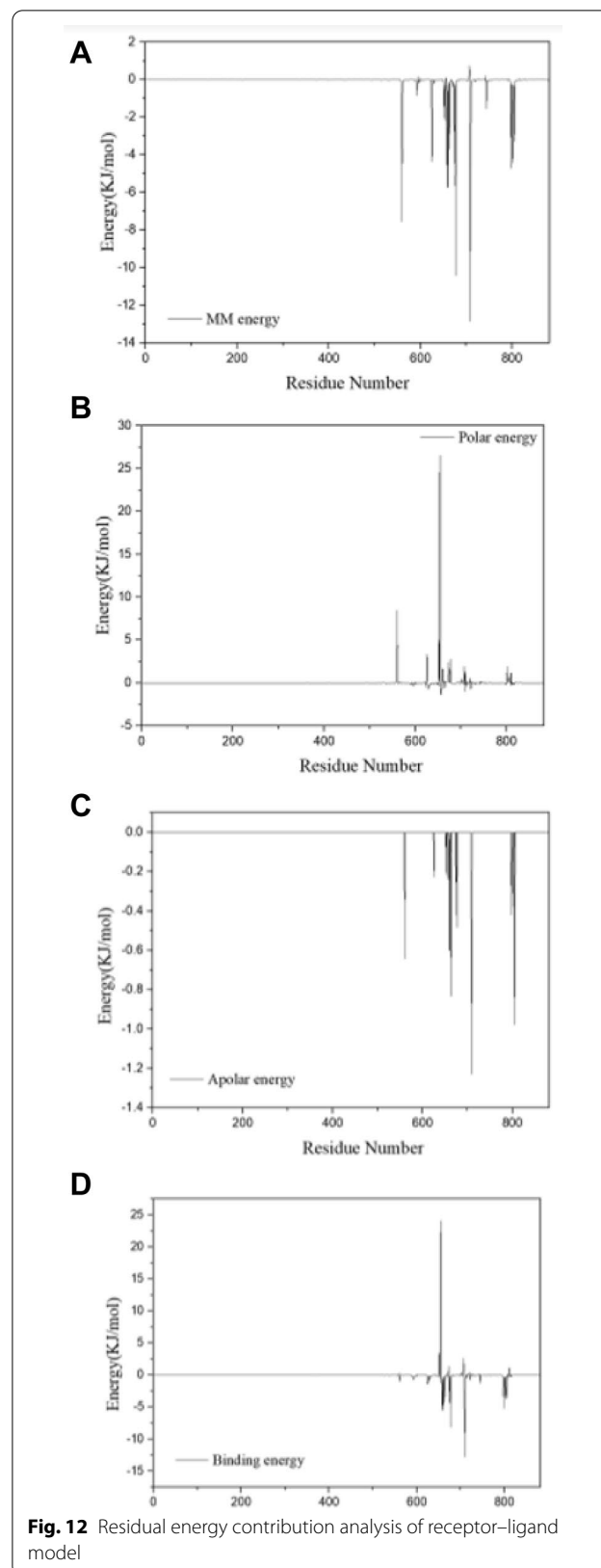
The RMSD of the CA-ligand during the molecular dynamic simulation is shown in Fig. 10. During the simulation time, RMSD of CA ranged from 0.18 to 0.44 nm, indicating that ligand structure was equilibrium during the simulation.



G_MMPBSA analysis of MT-CalB-substrates complex using G_MMPBSA module

The Gromacs trajectory of MT-CalB-CA is used as an example for `g_mmpbsa` module calculation. The input files `mmpbsa.zip` (including MT-CalB-CA.tpr, MT-CalB-CA.xtc and MT-CalB-CA.ndx) can be submitted by uploading relative files to the system (Additional file 1: Fig. S3A), and the configure parameters can be submitted by uploading configure file (parameter.xlsx) or inputting in the configure parameter page (Additional file 1: Fig. S3B). After clicking "Perform Calculation", an ID number is generated for the submitted job. When the calculation is finished, a notification email will be sent to the user's mailbox. The calculation results and output files can be downloaded by clicking "Download" in the `g_mmpbsa` model. The output file is an `allFile.zip` file which includes: `input_file`, `output_file`, `parameter file` and `result file`. The parameter file includes: `configure.xlsx` and `configure.json`. The output_file includes: `energy_MM.svg`, `polar.svg`, `apolar.svg`, `contrib_MM.dat`, `contrib_pol.dat`, `contrib_apol.dat`, `energymapin.dat`, `final_contrib_energy.dat`, `full_energy.dat` and `summary_energy.dat`. In addition, these above dates are all inserted into an excel file (`g_mmpbsa_analysis.xlsx`).

According to the MM-PBSA analysis results (Fig. 11), the Van Der Waals (VDW) interaction and electrostatic interaction are -179.618 kJ/mol and -12.618 kJ/mol, respectively. This indicates that VDW interaction plays the key role between MT-CalB with CA. Figure 12 ((A) MM energy contribution (B) Polar energy contribution (C) Apolar energy contribution (D) Binding energy contribution) shows the primary interaction residues between MT-CalB with CA. The polar interactions are observed between CA and Asp654, Thr560, Ala652, Ser625 of MT-CalB. The apolar interactions are observed



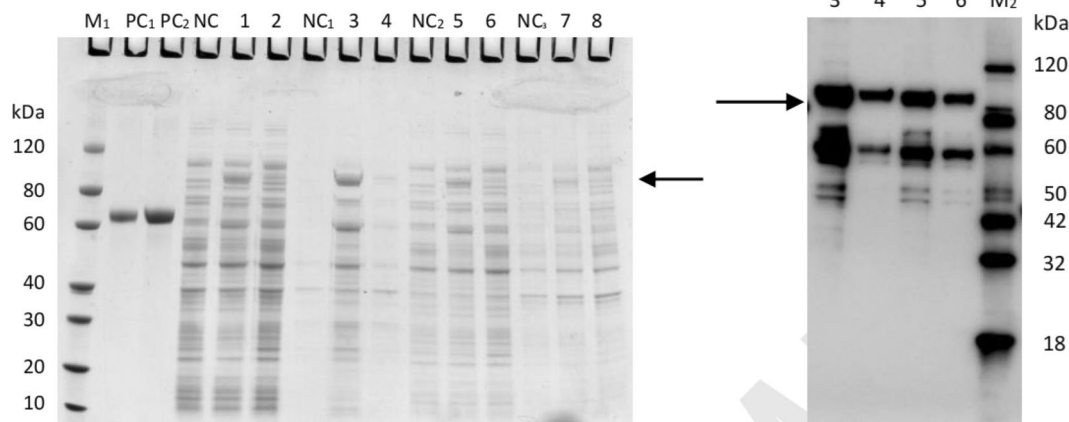


Fig. 13 SDS-PAGE result (left) and western blot (right) of MT-CalB. Lane M₁: protein marker; Lane M₂: western blot Protein marker; Lane PC₁: BSA (1 μg); Lane PC₂: BSA (2 μg); Lane NC: uninduced whole cell; Lane 1: whole cells were induced at 26 °C for 16 h; Lane 2: whole cells were induced at 15 °C for 16 h; Lane NC₁: uninduced pericytoplasm; Lane 3: periplasmic cells were induced at 26 °C for 16 h; Lane 4: pericytoplasm was induced at 15 °C for 16 h; Lane NC₂: uninduced supernatant of cell lysis; Lane 5: induced cell lysis supernatant at 26 °C for 16 h; Lane 6: cell lysis supernatant was induced at 15 °C for 16 h; Lane NC₃: uninduced cell lysis precipitation; Lane 7: 26 °C Induced cell lysis precipitation for 16 h; Lane 8: induced cell lysis precipitation for 16 h at 15 °C

between CA and Ile709, Ile805, Leu664, Thr560, Leu660 of MT-CalB. The VDW apolar interactions are observed between CA and Ile709, Gln677, Thr560, Leu660, Leu798, Thr658, Val674, Ala801, Ser625, Ala802 of MT-CalB. The results of hydrogen bond interaction analyzed by PLIP are consistent. Ile709, Gln677, Thr658, Leu798, Leu660, Leu664, Val674 are significant contributions to the binding energy residues. These residues are the main binding sites of MT-CalB and CA, and their binding with CA replace hydrogen bonds with water. The results of hydrophobicity are consistent with those of PLIP analysis.

Experimental verification of design results for MT-CalB

In this study, to increase the soluble expression of CalB in *E. coli*, a number of signal peptides, including mSpA, MBP-TrxA (MT), Carbohydrate Binding Domain (CBD), etc., were fused at N-terminal of CalB. Only MT fused CalB (MT-CalB) could be successfully expressed in *E. coli*. The yield of enzyme activity at optimal fermentation condition were 2085.57 U/L. The SDS-PAGE result is displayed in Fig. 13. The bands belonging to MT-CalB can be observed in the samples of periplasmic cells (lane 3), pericytoplasm (lane 4), induced cell lysis supernatant (lane 5) and cell lysis supernatant (lane 6) with the molecular weight at 93.2 kDa, indicating the successful expression of the MT-CalB.

Conclusions

In conclusion, an online server with a user-friendly interface has been established by our group, which is constructed for the following purposes: (1) predicting

the binding modes between receptor and small molecules using modeller and Auto dock Vina as docking engines; (2) building the gromacs model for dynamic simulation of receptor and ligand complex; and (3) performing the Molecular Mechanics/Poisson–Boltzmann Surface Area (G_MMPBSA) analysis module for receptor–ligand binding affinity analysis. MT-CalB-CA hopes to use it in the fields of biomedicine, energy, and chemical industry. Through this online server, the interaction between MT-CalB and CA, RMSD, the number of hydrogen bonds and Solvent Accessible Surface Area (SASA), and its stability in organic solvents are analyzed. The calculation results show that after MT-CalB interacts with CA, its structural rationality and stability are maintained. Finally, we also verifies through experiments that MT-CalB has better thermal stability and storage stability. The authors hope that the present web server will become a useful tool for enzyme design and computational biology.

Abbreviations

CA: Chaulmoogric acid; CalB: *Candida antarctica* Lipase B; CBD: Carbohydrate binding domain; *E. coli*: *Escherichia coli*; DS: Discovery studio; GMX: GROningen MACHine for Chemical Simulations; IPTG: Isopropyl β-D-1-thiogalactopyranoside; LB: Luria–Bertani; MBP: Maltose binding protein; MD: Molecule dynamics; MMPBSA: Molecular mechanics Poisson–Boltzmann surface area; mSpA: *Mycobacterium smegmatis* Porin A; MT-CalB: MBP-TrxA-CalB; NPT: Constant number of particles (N), pressure (P), and temperature (T); NAMD: NANoscale molecular dynamics; NVT: Constant number of particles (N), volume (V), and temperature (T); PLIP: Protein ligand interaction profiler; pNA: *p*-Nitroaniline; RMSD: Root mean square deviation; SASA: Solvent accessible surface area; SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis; TrxA: Thioredoxin A; VDW: Van der Waals.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40643-022-00546-y>.

Additional file 1: Fig. S1. Presentation of Vina module (A) Input file user interface; (B) Configure parameter user interface. **Fig. S2.** Presentation of Gromacs module (A) Input file user interface; (B) Configure parameter user interface. **Fig. S3.** Presentation of g_mmpbsa module (A) Input file user interface; (B) configure parameter user interface. **Fig. S4.** The quicklook of each module user what kind of file should be provided. **Table S1.** The whole protein sequence of MT-CalB. **Table S2.** The list of ligands from Pubchem. **Table S3.** The docking result of MT-CalB with ligands.

Acknowledgements

The authors want to thank Yun-Sheng Ye, Fei Wang and Rong-Huan Zhang, Biao-Bao Chen, Tian-Heng Zhou for their kindly help at deploying the modules onto the server. The authors want to thank Xu-Le Zou, Guan-Yuan Luo for their kindly help at deploying the python scripts.

Author contributions

J-HH and Y-HY are responsible for the website's structure and the development of Atomevo platform. Designed the algorithm: J-HH, Y-HY, W-HJ, M-JX, K-PH, and S-LC. Performed the Molecular Docking module M-JX. Performed the gromacs molecular simulation module: X-YL, W-HJ, J-TY. Performed the g_mmpbsa module: K-PH. Analyzed the data: X-YL, P-YL, Y-SL, W-MG, L-HL, Y-DW. Wrote the paper: J-HH, Y-HY, W-HJ, M-JX, K-PH and S-LC. D-JZ was in charge of experiment. All authors read and approved the final manuscript.

Funding

This research was funded by Guangdong Basic and Applied Basic Research Foundation (2019A1515110621), the Project of Department of Education of Guangdong Province (Young Creative Talents, Natural Science, No. 2017KQNCX217), the High-Level Talent Start-Up Research Project of Foshan University (GG07016), Academic Foundation for Students of Foshan University in 2021 (xsjj202108zra01, xsjj202108zrb08), the Special Funds for the Cultivation of Guangdong College Students Scientific and Technological Innovation (pdjh2022b0550 Climbing Program Special Funds). FOSU Open Funds of Laboratory Innovation (2022 No.005).

Availability of data and materials

All data generated or analyzed during this study are included in this article.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors consent to publishing the manuscript in *Bioresources and Bioprocessing*.

Competing interests

There is no conflict of interest for any of the authors regarding the submission of this manuscript.

Author details

¹Guangdong Key Laboratory of Food Intelligent Manufacturing, School of Food Science and Technology, Foshan University, Foshan 528000, China. ²School of Food Science and Technology, Foshan University, Foshan 528225, Guangdong, China. ³School of Food Science, South China Agricultural University, Guangzhou 510642, China.

Received: 13 December 2021 Accepted: 3 May 2022

Published online: 13 May 2022

References

- Baek M, DiMaio F, Anishchenko I et al (2021) Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 373:871–876. <https://doi.org/10.1126/science.abb8754>
- Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun* 91:43–56. [https://doi.org/10.1016/0010-4655\(95\)00042-E](https://doi.org/10.1016/0010-4655(95)00042-E)
- Bhachoo J, Beumung T (2017) Investigating protein–peptide interactions using the schrödinger computational suite. In: Schueler-Furman O, London N (eds) *Modeling peptide–protein interactions: methods and protocols*. Springer, New York
- Eberhardt J, Santos-Martins D, Tillack AF et al (2021) AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. *J Chem Inf Model* 61:3891–3898. <https://doi.org/10.1021/acs.jcim.1c00203>
- Fatma T, Zafar Z, Fatima S et al (2021) Computational assessment of *Botrytis cinerea* lipase for biofuel production. *Catalysts*. <https://doi.org/10.3390/catal11111319>
- Grosdidier A, Zoete V, Michielin O (2011) SwissDock, a protein–small molecule docking web service based on EADock DSS. *Nucleic Acids Res* 39:W270–W277. <https://doi.org/10.1093/nar/gkr366>
- Jumper J, Evans R, Pritzel A et al (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kumari R, Kumar R, Lynn A (2014) g_mmpbsa—a GROMACS tool for high-throughput MM-PBSA calculations. *J Chem Inf Model* 54:1951–1962. <https://doi.org/10.1021/ci500020m>
- Liu K, Chen X, Zhong Y et al (2021) Rational design of a highly efficient catalytic system for the production of PAPS from ATP and its application in the synthesis of chondroitin sulfate. *Biotechnol Bioeng* 118:4503–4515. <https://doi.org/10.1002/bit.27919>
- Manta B, Boyd D, Berkmen M et al (2019) Disulfide bond formation in the periplasm of *Escherichia coli*. *EcoSal plus*. <https://doi.org/10.1128/ecosalplus.ESP-0012-2018>
- McGregor MJ, Luo Z, Jiang X (2007) Virtual screening in drug discovery. In: Huang Z (ed) *Drug discovery research: new frontiers in the post-genomic era*. Wiley, New York
- Nelson MT, Humphrey W, Gursoy A et al (1996) NAMD: a parallel, object-oriented molecular dynamics program. *The Int J Supercomput Appl High Perform Comput* 10:251–268. <https://doi.org/10.1177/109434209601000401>
- O'Boyle NM, Banck M, James CA et al (2011) Open Babel: an open chemical toolbox. *J Cheminform*. <https://doi.org/10.1186/1758-2946-3-33>
- Phillips JC, Hardy DJ, Maia JDC et al (2020) Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* 10(1063/5):0014475
- Rester U (2008) From virtuality to reality—virtual screening in lead discovery and lead optimization: a medicinal chemistry perspective. *Curr Opin Drug Discov Devel* 11:559–568
- Salentin S, Schreiber S, Haupt VJ et al (2015) Current opinion in drug discovery & development PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Res* 43:W443–W447. <https://doi.org/10.1093/nar/gkv315>
- Thompson AP, Aktulga HM, Berger R et al (2022) LAMMPS—a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput Phys Commun*. <https://doi.org/10.1016/j.cpc.2021.108171>
- Valentine D, Zaslavsky I, Richard S, et al (2021) EarthCube data discovery studio: a gateway into geoscience data discovery and exploration with Jupyter notebooks. In: *Concurrency and computation: practice and experience*. <https://doi.org/10.1002/cpe.6086>
- Wang R, Lai L, Wang S (2002) Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J Comput Aided Mol Des* 16:11–26. <https://doi.org/10.1023/A:1016357811882>
- Wang Z, Sun H, Yao X et al (2016) Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys Chem Chem Phys* 18:12964–12975. <https://doi.org/10.1039/C6CP01555G>
- Xu J, Zhou H, Yu H et al (2021) Computational design of highly stable and soluble alcohol dehydrogenase for NADPH regeneration. *Bioresour Bioprocess* 8:12. <https://doi.org/10.1186/s40643-021-00362-w>

Zhang X, Peng Y, Zhao J et al (2020) Bacterial cytochrome P450-catalyzed regio- and stereoselective steroid hydroxylation enabled by directed evolution and rational design. *Bioresour Bioprocess* 7:2. <https://doi.org/10.1186/s40643-019-0290-4>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
