6-Ethoxy-4- N-(2-morpholin-4-ylethyl) -2-N-propan-2-yl-1,3, 5-triazine-2, 4-diamine endows herbicidal activity against *Phalaris minor* a weed of wheat crop field: An *in-silico* and experimental approaches of herbicide discovery

Nandan Kumar¹ · Priyanka Rani¹ · Shikha Agarwal² · Durg Vijay Singh¹

Received: 17 June 2021 / Accepted: 14 December 2021 / Published online: 4 March 2022 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

**Abstract**

*Phalaris minor* is a major weed of wheat crop which has evolved resistance against herbicides. Isoproturon is the most accepted herbicide developed resistance in 1992. Later, introduced herbicides also developed resistance and cross-resistance to their respective binding sites. Isoproturon binds at the Qb binding site of the D1 protein of photosystem-II (PS-II), which blocks the electron transfer in photosynthesis. In this work, we have carried out a series of computational studies to prioritize the promising herbicides against D1 protein of *P. minor*. Through the computational studies, twenty-four lead molecules are prioritized which have shown a higher binding affinity and inhibition constant than the reference ligand molecule. The binding and conformational stability of docked complexes was evaluated by molecular dynamics simulations and binding free energy calculations i.e., MM/PBSA. A list of amino acids such as Ala225, Ser226, Phe227, and Asn229 present in the binding site of protein is obtained to be playing an important role in the stability of the protein-lead complex via hydrogen bond and π-π interactions. Binding free energy calculation revealed that the selected lead molecule binding is energetically favorable and driven by electrostatic interactions. Among 24 leads, computational results have uncovered eight promising compounds as potential herbicides which have shown comparable physiochemical profile, better docking scores, system stability, H-bond occupancy, and binding free energy than terbutryn, a reference molecule. These prioritized molecules were custom synthesized and evaluated for their herbicidal activity and specificity through whole plant assay under laboratory-controlled conditions. The lead molecule ELC5 (6-ethoxy-4-N-(2-morpholin-4-ylethyl)-2-N-propan-2-yl-1,3,5-triazine-2,4-diamine) has shown comparable activity to the reference herbicide(isoproturon) against *P. minor*.

**Keywords** *Phalaris minor* · Herbicide resistance · Photosystem-II D1 protein · *De novo* design · MD simulation · MM/PBSA · Whole plant assay

**Introduction**

*P. minor* is the most serious grass weed of wheat, widely distributed over the rice–wheat cropping belt of north-west India [1, 2]. The most threatened states of India are Punjab, Haryana, Himachal Pradesh, and also a serious cause of concern in Uttar Pradesh, Uttarakanch, Madhya Pradesh, and some parts of Bihar. Isoproturon has worked as the most effective herbicide to control *P. minor* growth, but it has developed resistance and thus withdrawn in 1992. Henceforth, some new alternate herbicides *viz.* clodinafop, fenoxaprop, sulfoxsulfuron, and tralkoxydim were introduced in the market in 1997–1998. Application of these alternate herbicides brought the *P. minor* infestation under control for some period and helped to resume wheat production. However, cross-resistance to these four alternate herbicides had been also reported on a few biotypes [3, 4]. Singh et al. has reported difference in RAPAD profiles of susceptible and resistant biotypes of *P. minor* [5]. Tripathi et al. have
reported four point mutations in the D1 protein of \textit{P. minor}, which indicate target site alteration \cite{6}.

Isoproturon herbicide binds at the \(Q_B\) site of the D1 protein of PS-II located in the thylakoid membrane of the chloroplast. PS-II reaction center is a heterodimer of D1 and D2 protein encoded by the \textit{psbA} and \textit{psbD} gene \cite{7}. PS-II complex binds with the reaction center pigment P680, and the primary acceptor pheophytin, as well as the secondary acceptors \(Q_A\) and \(Q_B\), as shown in Fig. 1. The D1 and D2 proteins of the reaction center together bind with two pheophytin molecules and four chlorophyll-\(\alpha\) molecules. The herbicide and \(Q_B\) binding niche are prone to mutation lie in the transmembrane helix IV, and V. Horovitz et al. have reported that due to mutation of serine at the 264th position, it is replaced by glycine in the mutant D1 protein \cite{8}. Because of this, the herbicide molecule is deprived of one hydrogen bond as it cannot form H bonds with glycine. This results in a considerable decrease in affinity with herbicide molecules towards the D1 niche and the normal \(PQ_B\) molecule easily replaces them from the niche; thus, the normal electron transport continues in the mutant even in the presence of herbicide \cite{9, 10}. In our earlier in silico studies, few analogs against ACCase protein of \textit{P. minor} were proposed and designed \cite{11} and tested against D1 protein, wherein triazol derivative of isoproturon was the most effective \cite{12, 13}. Likewise, other researchers have employed rational approaches for the identification of inhibitor/herbicides against the number of targets \cite{14–17}. Kishore and Singh have reported that D1 protein of \textit{P. minor} and wheat share high sequence similarity; therefore, designing selective herbicides against this weed imposes a huge challenge \cite{18}. Using rationale approach of herbicide discovery, our group is focused to design and develop new herbicide against the \textit{P. minor} weed of wheat crop \cite{11–13}. In this study, we have modeled D1 protein, and then high-throughput virtual screening, docking, de novo design of chemical molecules, molecular dynamic simulation, and binding free energy are carried out and compared with ref\_lig to identify the promising herbicides against the D1 protein.

By virtual screening, large datasets of retrieved small molecules have been screened to identify potential lead molecules and a de novo approach was followed up to design new analogs, against the D1 protein. The binding affinity of screened hit molecules has been tested by docking studies, and then docked conformations were evaluated with MD simulations followed by binding free energy calculations. The implemented computational methods in the field of drug design provide the finest blueprint for in vivo experimental validation. Later in this study, a whole plant assay \cite{19} has been performed under laboratory-controlled conditions, to test and determine the biological activity, efficacy, specificity, and sensitivity of new molecules.

**Materials and methods**

**Template selection**

**Amino acid sequence of susceptible** [gi|31,506,000|gb|AAP47827.1|psbA] and resistant D1 proteins [gi|30,413,115|gb|AAP33145.1|psbA] of \textit{P. minor} have been retrieved from the NCBI protein sequence database (http://www.ncbi.nlm.nih.gov/protein/). PDB BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was performed to identify the suitable template for modeling D1 protein. The Cyanobacterial PS-II (pdb|3PRQ|A) complexes with terbutryn (triazine herbicide) at the \(Q_B\) binding site, resolution 3.20 Å, was selected as a template for building the model. The PDB ID: 3PRQ is now superseded with PDB ID: 4V82 (3PRQ file and 3PRR have been consolidated into a 4V82 for complete representation and none of the coordinates have been changed; https://www.wwpdb.org/pdb?id=pdb_00003prq). The co-crystallized terbutryn molecule was taken as a reference ligand (ref\_lig) to design and discover new herbicides.

**Homology modeling**

The structure of the D1 protein was modeled by Modeller v9.12, a homology modeling program \cite{20}. The modeling and validation approach is being described briefly as follows: (i) target template alignment, (ii) derivation of the model by MODELLER v9.12, (iii) structure refinement, and (iv) validation of the model. The spatial restraint method has been

---

**Fig. 1** Architecture of D1 and D2 protein of photosystem-II and \(Q_A\) site on D2 protein showed in a black color circle
employed to construct the model which satisfies restraints derived from the template structure. The restraints, including bond lengths and bond angles, van der Waals contact distances, and dihedral angles, are then mapped on the target structure based on the alignment. Derived models were evaluated based on their discrete optimized protein energy (DOPE) scores and root mean square deviation (RMSDs). Steepest descent energy minimization of modeled D1 protein was carried out for 1000 steps and, subsequently, conjugate gradient energy minimization was done for 500 steps using Swiss PDB Viewer 4.10 [21]. Harmonic constraints were used for energy minimization in SPDBV. Further, modeled structure was validated using structural analysis and verification server (http://nihserver.mbi.ucla.edu/SAVES/) inbuilt tools such as PROCHECK, WHAT_CHECK, WHATIF, PROSA, ERRAT, PROVE, and VERIFY-3D. Superimposition between modeled D1 protein and template structure was performed using TM-align server (http://zhanglab.ccmb.med.umich.edu/TM-align/) [22–24].

**Lead identification**

**Preparation of virtual screening dataset**

Herbicides, known for PS-II inhibition, were retrieved from the international survey of herbicide-resistant weed database (http://www.weedscience.com/summary/home.aspx). The retrieved molecules were classified into three different classes i.e., C1 (phenyl-carbamates, pyridazinones, triazines, triazinones, triazolinone, and uracils), C2 (ureas and amides), and C3 (nitriles and Phenyl-pyridazines). To maintain the possible diversity of the selected molecules of the aforementioned classes, all possible analog shaving 70% 2D structural similarity or above were retrieved from ZINC [25] and PubChem [26] databases.

**Virtual high-throughput screening**

Virtual-high throughputscreening has been performed by Molegro [27], a method to identify new pharmacologically active compounds on massive trials. The virtual high-throughput screening (vHTS) has been employed to predict the putative binding affinities of small molecules with D1 protein of PS-II at the Q_B binding site. The bond length, bond order, explicit hydrogen, and flexible torsion have been checked and assigned in both the protein and ligand data set. To screen the prepared ligand dataset, the following parameters were taken into consideration i.e., number of runs 50, population size 50, crossover rate 0.9, scaling factor 0.5, maximum iteration 5,000, and grid resolution 0.30 Å. The prepared data set was taken for screening for both rigid and flexible approaches at the center of ref_lig i.e., X: 20.63; Y: 65.50; and Z: 35.64 with 10-Å diameter from the center of terbutryn binding. A total of 43 amino acid residues were kept flexible, which lies under the 10-Å diameter of the grid to carry out the flexible docking. The flexible amino acids were Leu172, Phe173, Ser174, Ala175, Met176, HIS177, Ser179, Leu180, Val181, Ser183, Ser184, Tyr208, Asn209, Ile210, Val211, Ala212, Ala213, His214, Tyr216, Phe217, Arg219, Leu220, Ile221, Phe222, Glu223, Tyr224, Ala225, Ser226, Phe227, Asn228, Asn229, Ser230, Arg231, Leu233, His234, Phe235, Phe236, and Leu237. In both rigid and flexible approaches, molecules have been selected based on re-ranking, H bond, and MolDock scores.

**De novo design of inhibitors**

De novo design of new analogs has been performed by LigBuilder v2 package [28], a tool for structure-based drug discovery. It has a CAVITY and BUILD module for pharmacophore modeling, de novo synthesis of ligands, and processing of synthesized molecules. The ligand detection mode of the CAVITY module was used to find the key interaction residues at the binding site with the minimal distance between any two pharmacophore features (not to be less than 3 Å) and the binding site was selected based on ligandability less than or equal to 6.0Kd (equivalent to 1 µM). The EXTRACTION submodule was used to design seed molecules to ensure the binding affinity that decides the manner in which the ligand would grow, and growing mode of the BUILD module was used to construct molecules in the binding pocket. The whole construction process was controlled by a genetic algorithm. Protein–ligand binding affinity was evaluated by using an empirical scoring function: chemscore [28]. Besides binding affinity, the synthesizability of the ligands was also taken into account by applying certain sets of chemical rules (LigBuilder, http://www.ligbuilder.org/).

**Clustering of small molecule**

Ligands obtained from vHTS and de novo design have been clustered separately in a hierarchical fashion by the ChemAxon’sJKlustor Suite. The MCS tool of JKlustor has been used to search for the maximum common substructures (MCSs) of a compound library hierarchically. Clustering in a hierarchical fashion identifies the largest substructure shared by several molecular structures (JKlustorhttps://www.chema xon.com/products/jklustor/). All the vHTS and de novo synthesized ligands were clustered into 13 and 11 groups respectively. These molecules were clustered to retain at least one typical molecule of each class of herbicides. For further discussions, vHTS and de novo clusters are named as existing ligand clusters (ELCs) and de novo synthesized ligand clusters (DLCs) respectively.
Prioritization of leads

Molecular docking was performed for both ELC and DLC lead molecules at the binding site (terbutryn binding site) of D1 protein using AutoDock [29]. AutoDockTools (ADT) was used to select the parameters and to prepare the ligand and protein files. The screened molecules (screened through vHTS) were subjected to rigid docking followed by flexible docking. The grid parameter file was prepared by setting the dimensions of the grids for resistant D1 proteins, 40:40:40 Å, and spacing 0.375 Å with the center coordinates (20.63, 65.50, and 35.64). Grid point spacing was calculated for each of the atom types of protein and ligand (protein: A (aromatic C), C, HD, N, NA, OA, and SA; ligand: C, A, HD, N, NA, e (electrostatic), and d (desolvation)). AutoDock-run parameters were set as maximum iterations of energy evaluations, iterations of genetic algorithm, and number of GA per run were set to 2,500,000: 27, 000; and 100 respectively. The Lamarckian Genetic Algorithm (LGA) was selected for the ligand conformational search. The best molecule of each ELC and DLC cluster was selected based on binding energy, ligand efficiency, and inhibition constant. Finally, twenty-four leads were selected (13 from ELC and 11 from DLC).

Lead validation through molecular dynamics simulation

Molecular dynamics (MD) simulation was performed using Gromacs4.6.5 [30] to validate the 24 prioritized leads. The stability of D1 protein, complexed with ref_lig (docked complex), was evaluated in the lipid bilayer. MD simulations for the prioritized 24 lead complexes were evaluated in two sets, such as, in the first set, all 13 ELC, and, in the second set, all 11 DLC complexes were evaluated. The topology file for ELCs and DLCs was generated using PRODRG server [31] in the framework of GROMOS 53A6 force field. In order to impersonate the real environment, simulation of protein-lead complexes was carried out in DPPC (dipalmitylophosphatidylcholine) lipid bilayer environment [32]. The desired orientation of the D1 protein–ligand complex in lipid bilayer was taken from the Orientations of Proteins in Membranes (OPM) database (http://opm.phar.umich.edu). The bilayer was then expanded by translating lipid molecules and scaling the lateral dimensions of the box size using a scaling factor of 4 and subsequently shrinking step was performed to scale down the lipids by a factor of 0.95 to avoid the overlapping of the atoms in the system. During the scaling of the lipid bilayer, 22–23 steps of shrinking and energy minimization were performed to reach an area per lipid of ~71 Å² (close to the experimental value of ~62 Å²). Lipid parameters and topology files were obtained from Tieleman’s website (http://moose.bio.ucalgary.ca) [33]. Afterward, the DPPC head groups were hydrated by adding sufficient solvent molecules on either side of the lipids (unit cell dimension 8.30626 Å × 8.33877 Å × 6.59650 Å). The resulting simulation system contained 125 DPPC molecules (63 molecules in the upper leaflet and 62 in the lower leaflet) solvated on both sides with SPC water. The complete system contains ~21,000 atoms. A periodic boundary condition was employed to avoid edge effects. At physiological pH, the protein was having a negative charge; therefore, to make the system electrically neutral, 10 Na+ ions were added to the simulation box. The system was relaxed by energy minimization (10,000 steps of the steepest descent and 10,000 steps of the conjugated gradient) where protein and lipids were harmonically restrained to their original position with an isotropic force constant of 1000 kJ mol⁻¹ nm⁻². The above restraints allow for further equilibration of the bilayer and solvent while keeping the conformation of protein and ligand unchanged. The bilayer, solvent, and protein–ligand were separately coupled to the temperature and pressure bath. Temperature and pressure were maintained constant at 300 K and 100 kPa using weak coupling technique with coupling and relaxation parameters of 0.1 and the 1.0 Ps respectively [34]. The resulting systems were subjected to the production run for 240 ns (10 ns for each complex) using GROMOS 53A6 force field [35]. Linear constraint solver (LINCS) algorithm [36] was employed for covalent bond constraints. Results were analyzed by extracting information from trajectory files obtained from simulation. The root mean square deviation (RMSD), root mean square fluctuation (RMSF), and hydrogen bonds were analyzed using the Gromacs modules and graphs and images were produced using Xmgrace and Visual Molecular Dynamics [37].

Binding free energy calculation

Binding energy (ΔGbind) of D1 protein complexes of ELCs and DLCs leads was estimated from the trajectory file of each complex obtained in MD simulation using molecular mechanics-Possion-Boltzmann surface area (MM-PBSA) method [38]. The polar solvation energy values were calculated using the implicit solvation Poisson–Boltzmann model (PB) and nonpolar solvation energy values were calculated based on the solvent-accessible surface area (SASA). Finally, these energy components were combined to obtain ΔGbind value as shown in Eq. 1.

\[
\Delta G_{\text{bind}} = \Delta E_{\text{vdw}} + \Delta E_{\text{elec}} + \Delta G_{\text{PB}} + \Delta G_{\text{sur}} + -T\Delta S
\]

where ΔE_vdw and ΔE_elec represent van der Waals interactions and coulombic interactions respectively. ΔG_PB and ΔG_sur represent polar and nonpolar solvation energies, respectively. The entropic term TΔS was computed in the normal mode approximation [39].
Theoretical study of physiochemical properties of known herbicides of similar group

Most of the molecules fail during the experimental testing due to their poor physiochemical properties. Therefore, the properties of existing herbicides were investigated using Drug likeness Tool, DruLiTo, (http://www.niper.gov.in/pi_dev_tools/DruLiToWeb/DruLiTo_index.html).

Whole plant bioassay for the herbicidal activity of proposed molecules

Testing the bioactivity of test molecules

The proposed prioritized molecules were synthesized and characterized by Alba Nova (Alba Nova Life Science Private Limited, Hyderabad). The molecules were high in purity which was directly employed for testing. Some of the prioritized molecules were excluded due to the high-cost index of synthesis. The efficacy and specificity of custom synthesized molecules were evaluated through whole plant assay. The experiment was designed in triplicates under laboratory-controlled conditions. The test was conducted in plastic Petri plates (9-cm diameter). Ten seeds of P. minor, as well as wheat, were arranged on the moist double-layered Whatman filter paper in each plate. The prepared plates were placed in the plant growth chamber for germination of seeds, at 17 °C in dark conditions, and relative humidity of 45% for 6 days. An aliquot of distilled water was continuously supplied to Petri plates to keep the germinating seeds hydrated. The proposed herbicides were evaluated on P. minor at a concentration of 600 μM, along with control, isoproturon treated, and untreated control in triplicates. Isoproturon herbicide was taken as control/reference for bioactivity testing since terbutryn has an adverse effect (non-selective) over wheat plantlets: a non-selective herbicide. After 6 days (post germination of seeds), the seedlings in Petri plates were subjected to test as accordingly mentioned factors. The seedlings were provided with added 2 mL of respective solutions (distilled water in control, isoproturon in treated control, and test molecules in test-treated population) at an interval of 2 days, until a total 10 mL of the solution was supplied to each plate, followed by adding an aliquot amount of distilled water for hydration. Petri plates have again been incubated in plant growth chambers under the same optimal conditions but with altering light conditions (8/16 h day/night cycle) for the next 15/16 days. Afterward, observations regarding the effect of the compound upon plantlet growth, biomass accumulation, and other physical changes that occurred in plantlets due to treatment have been noted. Further data of wet and dry weight of plantlets have also been tabulated to compare the effect of compounds through biomass accumulation. Later, the molecule with the most effective efficacy and specificity towards the said weed was determined by comparing the results with the reference molecule isoproturon.

Testing the efficacy and selectivity of proposed molecule in laboratory controlled condition

The different subjects considered in this test were the same as earlier mentioned (untreated control population, treated control population with isoproturon as a reference, and treated populations with synthesized lead molecules). Further, the effective molecule was tested in triplicate to determine the concentration at which the growth rate of weed plantlets gets reduced by 50% (GR50). The testing was performed on a concentration of 0–600 μM at a regular interval of 40 μM in order to determine the meticulous GR50 dose of the new effective herbicide molecule.

Results and discussion

Homology modeling

Resistance D1 protein (gi|30,413,115|gb|AAP33145.1) was taken as a query sequence to find out the most suitable template and it has shown significant alignment with chain A of the photosynthetic oxygen evolving center (1SSL) having a score: 503; query coverage: 99%; identity: 90%; and E-value: 1e-178 and resolution 3.5 Å. Cyanobacterial D1 protein (pdb|3PRQ) complex with terbutryn at the QB binding site has also shown significant alignment i.e., query coverage of 100%, identity 90%, similarity 95.2%, and E-value 1e-178 without any gap in the alignment and resolution 3.2 Å (Fig. 2). The crystal structure of Cyanobacterial PS-II consists of 20 polypeptides along with 88 ligands and 10 cofactors [40] which was selected as a template protein to model the D1 protein. A total of 10 models were built and the best model was selected based on their discrete optimized protein energy (DOPE) scores and root mean square deviation (RMSDs). The selected model was subjected to a short 1000 and 500 steps of steepest descent and conjugate gradient energy minimization respectively and then subjected to the structural validation using SAVES server and other approaches.

The Ramachandran plot showed phi and psi torsion angle of 93.6% amino acid residues in core regions and 6.4% amino acid residues in the allowed region for the modeled structure whereas 82.3% amino acid residues in the core region and 17.0% amino acid residues in the allowed region for template structure (Table S1 and Figure S1). The quality of the model was also validated by the ProSA server, a web server for protein structural analysis. The ProSA Z-score was –4.0 which indicates the overall quality and
measures the deviation from the total energy of the structure with respect to an energy distribution derived from random conformations. A negative PROSA score indicates the correctness of the modeled protein. Verify 3D was used for validation of the modeled protein and it showed an average 3D-1D profile score of -0.09 and 0.52. TM-align server was used to further validate the modeled protein which showed TM score 0.99125 (0.5 < TM Score > 1, indicates about the same fold), suggesting the modeled protein to be of good quality, as shown in Fig. 3. Thus, results suggest that the comparative modeled structure of D1 protein is of good quality, and could be used for further studies. The homology modeling was performed to get the blueprint of the binding mode of herbicide terbutryn complexed with D1 protein (PDB ID 3PRQ/4V82) which is precisely located at the QB site of D1 protein for the design and development of herbicides. The target template identity is ~90% without any gap so the model structure could be considered equivalent to crystallographic structure except for a few side chains [41].

In a previous study, our group has built a chimeric homology model of the PS-II reaction center of P. minor. The D1 protein is truncated at both N and C terminals, so both missing ends were concatenated with ref sequence of T. aestivum’s D1 protein. Sequence for D2 protein was taken from its nearest homolog i.e., T. aestivum (from Ref Seq Database). The build reaction center was having one β-carotene, one bicarbonate ion, five chlorophyll-a, two pheophytin-a, two plastoquinone and cofactors (iron and OEC) along with D1 and D2 protein. To check the integrity and stability of the build chimeric model, it was subjected for MD simulation, and we reported that the complete reaction center deviates by 1.7175 Å RMSD from its equivalent bacterial reaction center i.e., template; PDB ID 2AXT [13].

Virtual high-throughput screening

The stringent criteria were applied in vHTS to narrow down the number of molecules as shown in Fig. 4. vHTS for all retrieved 33,016 ligands (Table S2) have been performed at the ref_lig binding site of modeled D1 protein and 7250 ligands were selected based on rerank, H bond, and MolDock score either more or equal to ref_lig molecule score
To enrich the screening, selected ligands were passed through three times flexible docking, and at least one molecule from each class and subclass of the C1, C2, and C3 groups was retained. Best screened ligands were selected based on the rerank score (best 5% ligands), H bond, and MolDock score which were having more or equal to ref_lig score.

**Designing of new analogs at the binding site of D1 protein**

The de novo drug design program Ligbuilder v2 was used to design the new analogs of selected ligands iteratively using an inbuilt library of organic fragments. Overall, 28 ligands *i.e.*, 16, 7, and 5 ligands from class C1, C2, and C3 respectively, were set for the de novo synthesis (Table S3). Ligand detection mode of the CAVITY module was used to detect the binding site of the protein and the best binding site was decided based on ligandability (Fig. 5). These were compared with ref_lig binding sites in D1 protein. The automatic extractor module was used which has extracted 71, 29, and 19 seeds from C1, C2, and C3 classes respectively (Table S3). Extracted seed molecules were initially used for growing to design new ligand molecules using BUILD module. BUILD module uses a pocket and grid file generated by the CAVITY module which acts as pharmacophore for de novo synthesis. Here, molecules were synthesized considering the Lipinski rule of five, chemscore, binding affinity, synthesizability, and toxicity parameter for all molecules as well as the molecules were
represented with the route of synthesis with cost index. As a result, 282 ligands were synthesized which belong to C1 (169), C2 (72), and C3 (41) classes.

**Clustering**

The hierarchical clustering of both vHTS ligands (230) and de novo synthesized ligands (282) was performed to get 13 and 11 clusters respectively (Table S3). The vHTS and de novo clusters were named as existing ligand clusters (ELCs) and de novo synthesized ligand clusters (DLCs) respectively. Screened leads from individual clusters of ELC and DLC (Table S3) were selected for the docking study.

**Docking**

The ELC and DLC hits were docked at the center coordinates ($X$: 20.63; $Y$: 65.50; and $Z$: 35.64) of the binding site of modeled D1 protein using AutoDock4.2 as well as the Molegro tool. The criteria chosen to select the best pose of each hit were AutoDock term (binding energy, ligand efficiency, and inhibition constant) and Molegro term (MolDock score, H bond score, and rerank score). One hundred poses per ligand were generated and clustered with an RMSD tolerance limit of 1.5 Å. Binding poses of all hits of the ELC and DLC were selected from the most populated cluster and a summary of these hits are tabulated in Table S5 and Table S6. Thirteen hits from ELC and 11 hits from DLC were selected from each cluster of the ELC and DLC respectively after employing the docking score filter. Three amino acids *i.e.*, His177, Phe227, and Phe236, in binding sites were observed to be playing a crucial role in interaction with hits. Docking results have shown that ref_lig has two H bonds and one π-π interaction in the binding site of protein with ligand efficiency of $-0.25$, and inhibition constant of 1.08 mM (Table S4 and Figure S2).

The ref_lig docking scores were set as criteria to select the best hits from each cluster, as a result, 13 and 11 hits were selected from ELC and DLC clusters. In addition to ref_lig-based criteria, two or more H bond and one π-π interaction with better binding energy, ligand efficiency, and inhibition constant were employed as selection criteria to select the molecule (Table S5-S6 and Figure S3-S6).

**Molecular dynamics simulations**

To find the binding affinity and binding mode of ligand, rigid docking followed by flexible docking was applied, but it does not properly sample the conformational space that could lead to a variety of different binding modes of the ligands [42]. The goals of these simulations were to observe the structural behavior of protein (flexibility of
the D1 protein with ref_lig, ELC, and DLC leads) and ligands (the behavior of leads inside the binding site of D1 protein). MD simulations were performed for each protein-lead complex for 10 ns using Gromacs 4.6.5 in an environment of a lipid bilayer embedded in a periodic box solvated with water, shown in Fig. 6. PS-II complex has 20 peptides and 95 co-factors (PDB ID 4V82). D1 protein complexes were with at least one β-carotene, one pheophytin, two α-chlorophylls, two plastoquinone, at least two Fe^{2+} ions, and one bicarbonate. The PS-II reaction center protein and associated molecules of the complex were in a tight-knit providing stability to each other, and having interlinked function. In our previous attempt, a chimeric D1/D2 protein was constructed from its nearest homolog and subjected to MD simulation. The chimeric D1/D2 protein has shown RMSD of 1.7175 Å from its template protein [13]. Therefore, we have refrained from doing any long simulation with the D1 subunit in this study. RMSD curves of the protein backbone, energy profile, and H bond profile of each protein-lead complexes were calculated and compared with the ref_lig protein complex. Variation in the total energy plot of the ref_lig protein complex has shown that the overall energy tends to stabilize after 3 ns of simulation (Fig. 6B), which conveys the average energy remains constant, thus suggesting structural stabilization. The total energy patterns were similar for all ELC and DLC lead protein complexes (data are not shown).

Throughout the simulation period, no significant fluctuations were observed in the backbone of protein with ref_lig. It implies that the binding of ref_lig at the binding site of protein is not only stable and strong but also does not perturb the protein backbone stability. The binding stability of the ELC leads (ELC2, ELC3, ELC4, ELC5, ELC6, ELC10, ELC11, and ELC12) and DLC leads (DLC4, DLC6, DLC7, and DLC8) in the binding site of the protein have shown comparable fluctuation with ref_lig protein complex throughout the simulation in the size of the noticeable window of 0.30–0.40 nm (Fig. 7). It ensures and suggests that these leads are well accommodated inside the binding site throughout the simulation. H bonds at the binding site of the ref_lig protein complex were determined which revealed that it encompassed four

---

Fig. 6 A Snapshot of the simulation system after solvation. Water molecules are shown in red; phospholipids groups are shown in gray; protein is shown in orange and ligand is shown in blue. B Total energy variation for the ref_lig protein complex throughout the simulation

Fig. 7 Time dependence root means square deviation (RMSD) of the backbone of protein throughout the simulation. A RMSD of protein with ELC leads and compared with RMSD of protein with ref_lig. B RMSD of protein with DLC leads is compared with an RMSD of protein with ref_lig
H bonds throughout the simulation period. The amino acid residues involved in the key H-bond formation at the binding site were Ala225, Ser226, and Phe227. The succession of frames for ELC and DLC leads inside the binding site of the protein has shown that ELC leads ELC2, ELC3, ELC5, ELC7, ELC8, ELC11, and ELC13 and DLC leads DLC3, DLC4, DLC6, DLC8, and DLC11 comprise three to four H bonds throughout the simulation as shown in Fig. 8. The amino acid residues such as His177, Ala225, Ser226, Phe227, Asn229, and Ser232 were observed to be holding the lead molecules at the binding site via H bonds. These interacting patterns advocate that leads are also interacting in a similar fashion as ref_lig. Hence, MD results confirm that these leads are accommodating well into the binding site of the D1 protein.

**Amino acid fluctuation analysis**

The root mean square fluctuation (RMSF) analysis revealed the possible movements of amino acid residues Phe201, Phe222, Tyr224, and Phe227 (inside the binding pocket) and Arg26, Gln127, and Arg187 (out the binding site) in the ref_lig ligand–protein complex (Fig. 9). In the case of selected ELC and DLC lead complexes, RMSF appears to be almost similar to the ref_lig except for the other small fluctuations. This signified that protein-ELC and protein-DLC lead complexes are also stable like the ref_lig-protein complex. The amino acid residues inside the binding sites have shown slightly less fluctuation because of the formation of a π-π interaction between the leads and the aromatic ring of Tyr224 and Phe227. Moreover, residues of the binding site showed lower fluctuation, revealing to be a more stable protein–ligand system.

**Analysis of hydrogen bond occupancy**

The MD trajectories were analyzed for the H-bond distance and the H-bond occupancy. H-bond distance is determined by the bond length (< 3.5 Å) and the H-bond occupancy is defined as the time percent existence of a
specific H bond during the simulation. The H bonds predicted for the ref_lig, ELC, and DLC lead and amino acid residues of the binding site of protein were observed consistently for the major part of trajectory with high occupancy for all the proposed lead molecules. H-bond interactions were categorized as strong (>60%), medium (40–60%), and weak (<40%) based on H-bond occupancy with ref_lig and the binding site of the protein. Four H bonds were found between ref_lig and protein i.e. O@Ala225…HAG@ref_lig, HG@Ser226…NAL@ref_lig, O@Phe227…HAG@ref_lig, and H@Phe227…NAL@ref_lig with occupancy 63.0%, 59.9%, 55.2%, and 60.3%, respectively, and the distance of H bonds were fluctuating in windows of 1 Å to 3.5 Å. Occupancies of the H bonds between protein, ELC, and DLC lead molecules showed a diverse range of fluctuation which is comparable to ref_lig. ELC leads ELC2, ELC3, and ELC8 have shown four H bonds, where two H bonds are having occupancy >70% and another two H bonds having occupancy 70 to 30%. ELC4 and ELC7 showed three stable H bonds with significant occupancy >40%. DLC leads DLC3, DLC6, and DLC9 showed four H bond occupancy >40%. DLC4 and DLC8 ligands also have shown two stable H bonds with significant occupancy. Other ELC and DLC lead molecules (except the above-described ELC and DLC lead molecules) have not shown a significant stable H bond as compared to ref_lig. The occupancies of the H-bond analysis also showed that amino acid residues His177, Ala225, Ser226, Phe227, and Asn229 were involved in the H bonding with ref_lig, ELC, and DLC lead molecules. However, the residues Ser226 and Phe227 were obtained as key residues to hold the molecules at the binding site as these residues showed higher occupancy in most of the ELC and DLC lead molecules (Fig. 10).

**Free energy calculations**

Binding free energy of D1 protein complexed with ref_lig, ELC, and DLC molecules was determined by the MM/PBSA method. In the case of ref_lig, van der Waals interactions ($\Delta G_{vdw}$), showed the most significant contributions among the van der Waals interactions ($\Delta G_{vdw}$), the electrostatic energy ($\Delta G_{ele}$), and the nonpolar solvation free energy ($\Delta G_{nonpol\_sol}$). The results reveal that leads from ELC class (ELC1, ELC2, ELC4, ELC5, ELC8, ELC11, ELC12, and ELC13) and lead from the DLC class (DLC1, DLC3, DLC4, and DLC6) displayed higher binding free energy compared to ref_lig binding free energy ($-104.627 \pm 10.792$ kJ/Mol). The order of binding free energy of ELC’s and DLC’s is ELC4 > ELC11 > ELC8 > ELC5 > ELC13 > ELC1 > ELC2 and DLC6 > DLC4 > DLC3 > DLC1 > DLC3 respectively. The binding energies of the ELC and DLC lead molecules are shown in Table 1, which reveals that $\Delta G_{vdw}$ has the most significant contribution to the binding energy.

From Table 1, it can be seen that ELC1, ELC13, DLC1, and DLC3 showed better binding affinity than ref_lig. However, these lead molecules are discarded in the broad-spectrum purpose as the number of H bonds was decreased from the initial structure, and occupancy was lower than the other lead molecules. Therefore, the lead molecules ELC2, ELC4, ELC5, ELC8, ELC11, ELC12, DLC4, and DLC6 were selected. These selected ELC and DLC lead molecules and

![Fig. 10](http://example.com/fig10.png)
the other known herbicides for D1 protein of PS-II protein were undertaken to predict the physiochemical properties (molecular weight, logP, AlogP, number of hydrogen donors and acceptors, total polar surface area (TPSA), number of rotatable bonds, number of acidic groups, number of rigid bonds, number of aromatic rings, and number of hydrogen bonds) (Fig. 11, Table S7).

From the data presented in Fig. 11 and Table S7, it can be observed that the properties of all the proposed molecules are comparable to known PS-II herbicides and have no toxic functional groups. Therefore, proposed molecules (ELC2, ELC4, ELC5, ELC8, ELC11, ELC12, DLC4, and DLC6) are considered potential lead molecules for experimental testing and further authentication. The 2D structures of these selected molecules are shown in Fig. 12, and ZINC, PubChem ID, and IUPAC name are shown in Table S8. However, only four molecules (ELC8, ELC5, ELC11, and ELC12) were synthesized among the proposed eight molecules, as shown in Table 2. Other molecules were not synthesizable due to the high polarity of intermediates, thereby increasing the cost of synthesis which would not be viable as an herbicide.

**Whole-plant bioassay**

**Testing the bioactivity of synthesized molecules**

The tabulated data of untreated control and treated control with isoproturon (ref, and treated with synthesized test molecules) (Table 2) were compared to interpret the effect of test molecules against *P. minor* and wheat with respect to the solvent (2% DMSO) and distilled water. Non-treated seedlings have a healthy growth rate as compared to treated seedlings. Besides late germination in treated seeds, yellowing of leaflets of treated plantlets has been observed after 1 week of treatment. Progression in the yellowing of...
leaflets has been observed in the treated plantlets, reflecting the visible chlorosis in treated plantlets which could be an indirect reflection of hindrance in photosynthesis thereby loss in biomass accumulation. At the end of the experiment, biomass accumulation in plantlets has been tabulated. Biomass accumulation has been compared among the test molecules and isoproturon (ref), to determine the most efficient molecule among the four test molecules graph that has been plotted (Fig. 13, Table S9). From the graph in Fig. 13, it has been recorded that there is no effect of either control isoproturon or test molecules upon wheat plantlets. However, it has been observed that there is a selective effect of test molecules upon *P. minor*, with different activity range i.e., \( ELC_5 > ELC_{12} = ELC_8 > ELC_{11} \). The efficacy and selectivity of ELC5 (6-ethoxy-4-N-(2-morpholin-4-ylethyl)-2-N-propan-2-yl-1,3,5-triazine-2,4-diamine) is comparable as isoproturon. Thus, it has been further tested and evaluated to determine the GR\(_{50}\) of ELC5 molecule.

Testing the efficacy and sensitivity of ELC5 molecule in laboratory-controlled condition

The ELC5 has been further tested to evaluate its activity at different concentrations (0–600 μM at a regular interval of 40 μM), to determine its GR\(_{50}\). The plot (Table S10) decipher that there is no effect of either isoproturon or test molecule ELC5 upon wheat at any concentration, thus promising the selective activity of ELC5 against *P. minor* only. The tabulated data of biomass has been statistically analyzed using a descriptive statistics (confidence level: 95%) approach to verify the truthfulness of the results. Upon further evaluation of tabulated data (Table S10), it has been interpreted that upon the treatment of plantlets with ELC5, there is \(~50\%\) reduction in biomass of *P. minor* plantlets at the concentration of \(~40\) μM (Fig. 14).

ELC5 (6-ethoxy-4-N-(2-morpholin-4-ylethyl)-2-N-propan-2-yl-1,3,5-triazine-2,4-diamine) molecule has shown
Table 2  Synthesized molecules for laboratory testing

| SI. no | Prioritized small molecules                                                                 | Code   | PubChem/ZINC ID                           |
|--------|---------------------------------------------------------------------------------------------|--------|-------------------------------------------|
| 1      | 4-Chloro-2,6-bis[(5-chloro-2-hydroxyphenyl)methyl]phenol                                      | ELC11  | PubChem ID.-62,616 ZINC ID.-1,679,491      |
| 2      | 1-Benzyl-3-(5-tert-butyl-1,2-oxazol-3-yl)-1-methylurea                                        | ELC8   | PubChem ID.-20,330,566 ZINC ID.-59,961,084|
| 3      | 6-Ethoxy-4-N-(2-morpholin-4-ylethyl)-2-N-propan-2-yl-1,3,5-triazine-2,4-diamine               | ELC5   | PubChem ID.-767,097 ZINC ID.-19,520,234    |
| 4      | 2-[[4-Chloro-6-(2-hydroxyethylamino)-1,3,5-triazin-2-yl]amino]-2-methylpropan-1-ol            | ELC12  | PubChem ID.-22,333,149 ZINC ID.-6,721,886  |

Fig. 12  2D structure of proposed eight best lead molecules (six from ELC and two from DLC) as inhibitors of D1 protein
comparable activity to isoproturon a reference molecule against *P. minor* under laboratory-controlled conditions. It could be further taken for pot assay and field assay (pilot-scale synthesis and testing for various parameters) to test its effectiveness, efficacy, selectivity, and sensitivity in a natural uncontrolled environment. The outcome of the assays could suggest whether the molecule is ready to use or need some chemical modification and optimization.

**Conclusion**

*P. minor* resistance to herbicides is one of the major causes of reduced wheat production, in the wheat cropping belts in India as well as in other countries. Due to the development of resistance to existing herbicides, it is of utmost importance to find a novel herbicide to restore wheat production. In silico studies *i.e.*, molecular modeling, virtual screening, de novo synthesis, and docking as well as molecular dynamics simulation and MM/PBSA calculation, were performed to find promising molecules with a high binding affinity than ref_lig (terbutryn; complexed with D1 protein PDB:3PRQ) at the QB site of D1 protein. The screened small molecules from large datasets and de novo synthesized molecules were taken for docking studies, molecular dynamics simulation, and MM/PBSA calculation. From the result obtained, ELC2, ELC4, ELC5, ELC8, ELC11, ELC12, DLC4, and DLC6 have shown better features in all aspects in comparison to ref_lig. These molecules were selected based on protein backbone stability, the number of H bond and their occupancy, non-covalent interactions, and binding free energy. These theoretical results suggested that the selected molecules accommodate well inside

![Image](image_url)
the binding site throughout the MD simulation, which reveals the stability of the system, later confirmed by the RMSD analysis. The strength of binding of selected molecules has been evaluated by H-bond occupancy, and binding energy, and then physiochemical profiles were also checked which has shown comparable to the other PS-II herbicides. Finally, four prioritized molecules were synthesized and tested in whole plant assay in laboratory-controlled conditions. The whole plant assay showed that ELC5 (6-ethoxy-4-N-(2-morpholin-4-ylethyl)-2-N-propan-2-yl-1,3,5-triazine-2,4-diamine) is equally effective as isoproturon (GR50 at 40uM) and does not have an adverse effect over wheat plantlets. The newly discovered lead molecule mentioned formerly is proposed for pilot-scale testing in pot and crop fields.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00894-021-05006-6.

Acknowledgements The authors of this manuscript acknowledge Prof. Krishna Mishra, Indian Institute of Information Technology Allahabad, and Prof. R. S. Rathore HOD, Dept. of Bioinformatics, for extending all possible support in the completion of this research work.

Author contribution D. V. Singh conceived and supervised the study. N. K. and P. R. performed the calculations and experiments. N. K. and S. A. interpreted the results and writing of the manuscript is done by all the authors.

Funding The authors acknowledge the funding supported by the Science and Engineering Research Board, India file No.YSS/2015/001662.

Availability of data and material Supplementary information is available for this paper.

Code availability Not applicable.

Declarations

Competing interests The authors declare no competing interests.
7. References

1. Chhoker RS, Sharma RK (2008) Multiple herbicide resistant in little seed canary grass (Phalaris minor): a threat production in India. Weed Biol Manag 8:112–123
2. Chhoker RS, Sharma RK, Sharma I (2012) Weed management strategies in wheat-a review. J Wheat Res 4:1–21
3. Yadav A, Malik RK (2005) Botanical characteristics, biology and distribution in herbicide resistant Phalaris minor in wheat - a sustainability issue. Hisar, India Department of Agronomy and Directorate of Extension Education, CCS Haryana Agricultural University, 5–15
4. Yadav A, Malik RK (2012) Alternate herbicides against isoproturon resistant Phalaris minor in herbicide resistant Phalaris minor in wheat - a sustainability issue. Hisar, India, Department of Agronomy and Directorate of Extension Education CCS Haryana Agricultural University, 89–98
5. Singh DV, Gaur AK, Mishra DP (2004) Biochemical and molecular mechanisms of resistance against isoproturon in Phalaris minor: variations in protein and RAPD profiles of isoproturon resistant and sensitive Phalaris minor biotypes. Indian J Weed Sci 36:256–259
6. Tripathi MK, Yadav MK, Gaur AK, Mishra DP (2005) PCR based isolation of psbA (herbicide binding protein encoding) gene using chloroplast and genomic DNA from Phalaris minor biotype(s). Physiol Mol Biol Plants 11:161–163
7. Marder JB, Chapman DJ, Telfer A, Nixon PJ, Barber J (1987) Identification of psbA and psbD gene products, D1 and D2, as reaction centre proteins of photosystem 2. Plant Mol Biol 9:325–333
8. Horovitz A, Ohad N, Hirschberg J (1989) Predicted effects on herbicide binding of amino acid substitution in the D1 protein of photosystem II. FEBS Lett 234:161–164
9. Hirschberg J, Bleckeis A, Kyle DJ, McIntosh L, Arntzen C (1984) The molecular basis of triazine resistance in higher plant chloroplasts. Natur Forch 39:412–419
10. Hirschberg J, Yehuda AB, Pecker I, Ohad N (1987) Mutations resistant to photosystem II herbicides. NATO ASI series. Series A: Life Sciences 140:357–66
11. Rani P, Kumari J, Agarwal S, Singh DV (2019) Binding mode of aryloxophenoxypropionate (FOP) and cyclohexane (DPOP) groups of herbicides at the carbonyl transferase (CT) domain of Acetyl-CoA carboxylase of Phalaris minor. Nettw Model Anal Health Inform Bioinforma 8:1–14
12. Singh DV, Adeppa K, Misra K (2011) Mechanism of isoproturon resistance in Phalaris minor: in silico design, synthesis and testing of novel herbicides for regaining sensitivity. J Mol Model 18:1431–1445
13. Singh DV, Agarwal S, Kesharwani RK, Misra K (2012) Molecular modelling and computational simulation of the photosystem-II reaction center to address isoproturon resistance in Phalaris minor. J Mol Model 18:3903–3913
14. Zhang H, Yao Y, Yang H, Wang X, Kang Z, Li Y, Li G, Wang Y (2012) Molecular dynamics and free energy studies on the carboxypeptidases complexed with peptide/small molecular inhibitor: mechanism for drug resistance. Insect Biochem Mol Biol 42:583–595
15. Liu J, Liu M, Yao Y, Wang J, Li Y, Li G, Wang Y (2012) Identification of novel potential β-N-Acetyl-D-hexosaminidase inhibitors by virtual screening, molecular dynamics simulation and MM-PBSA calculations. Int J Mol Sci 13:4545–4563
16. Kang T, Gao S, Zhao L, Zhai Y, Ye F, Fu Y (2021) Design, synthesis, and SAR of Novel 1,3-disubstituted imidazolidine or hexahydropyrimidine derivatives as herbicide safeners. J Agric Food Chem 69:45–54
17. Fu Y, Wang M, Zhao L, Zhang S, Liu Y, Guo Y, Zhang D, Gao S, Ye F (2021) Design, synthesis, herbicidal activity and CoMFA of ary1-formyl piperidinone HPPD inhibitors. Pestico Biochem Physiol 174:104811
18. Kishore G, Singh DV (2018) Isoproturon tolerance and resistance in Phalaris minor: sequence and structural similarity in Phalaris minor and wheat D1 protein. Pestic Res J 30:246–250
19. Burgos NR (2015) Whole-plant and seed bioassays for resistance confirmation. Weed Sci 63:152–165
20. Fiser A, Salì A (2003) Modeller: generation and refinement of homology based protein structure models. Methods Enzymol 374:461–491
21. Guez N, Peitsch MC (1997) SWISS-MODEL and Swiss-Pdb Viewer: an environment for comparative protein modelling. Electrophoresis 18:2714–2723
22. Laskowski RA, MacArthur MW, Moss DS, Thornton J (1993) PROCHECK: a program to check the stereo chemical quality of protein structures. J Appl Crystallogr 26:283–291
23. Wiederstein M, Sippel MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:407–410
24. Zhang Y, Skolnick J (2005) TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic Acids Res 33:2302–2309
25. Irwin JJ, Shoichet BK (2005) ZINC – a free database of commercially available compounds for virtual screening. J Chem Inf Model 45:177–182
26. Kim S, Thiesen PA, Bolton EE, Chen J, Fu G, Gindulute A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2015) PubChem substance and compound databases. Nucleic Acids Res 44:202–213
27. Thomasen R, Christensen MH (2006) MolDock: a new technique for high-accuracy molecular docking. J Med Chem 49:3315–3321
28. Yuan Y, Pei J, Lai L (2011) LigBuilder 2: a practical de novo drug design approach. J Chem Inf Model 51:1083–1091
29. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 16:2785–2791
30. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4:435–447
31. Schultekopf AW, Aalten MPV (2004) PRODRG: a tool for high throughput crystallography of protein-ligand complexes. Acta Crystallogr 60:1355–1363
32. Marrink SJ, Berger O, Tieleman DP, Jahnig F (1998) Adhesion forces of lipids in a phospholipid membrane studied by molecular dynamics simulations. Biophys J 74:931–943
33. Tieleman DP, MacCallum JL, Ash WL, Kandt C, Xu Z, Monticelli L (2006) Membrane protein simulations with a united-atom lipid and all-atom protein model: lipid–protein interactions, side chain transfer free energies and model proteins. J Phys: Condens Matter 18:1221
34. Berendsen HJC, Postma JPM, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690
35. Oostenbrink C, Villa A, Mark AE, Gunsteren WV (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 25:1656–1676
36. Amiri S, Sansom MS, Biggin PC (2007) Molecular dynamics studies of AChBP with nicotine and carbamylcholine: the role of water in the binding pocket. Protein Eng Des Sel 20:353–359
37. Humphrey W, Dalke A, Schulten K (1996) VMD - visual molecular dynamics. J Mol Graph 14:33–38
38. Kumari R, Kumar R, Open Source Drug Discovery Consortium, Lynn A (2014) g_mmpbsa-A GROMACS tool for high-throughput MM-PBSA calculations. J Chem Inf Model 54:1951–62
39. McQuarrie DA (1973) Statistical thermodynamics, 2nd edn. Harper and Row, New York
40. Broser M, Glockner C, Zouni A (2011) Structural basis of cyanobacterial photosystem II inhibition by the herbicides terbutryn. J Biol Chem 286:15964–15972
41. Sippl MJ (1993) Recognition of errors in three dimensional structures of proteins. Proteins 17:355–362
42. Lu Q, Wang DS, Chen CS, Hu YD, Chen CS (2005) Structure-based optimization of phenylbutyrate-derived histone deacetylase inhibitors. J Med Chem 48:5530–5553

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