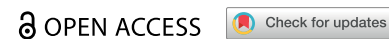


RESEARCH ARTICLE



## Investigation of the inhibitory properties of azo-dyes on chorismate synthase from *Paracoccidioides brasiliensis*

Katharina Fuchs\*, Massimo G. Totaro\*, Marina Toplak, Aleksandar Bijelic and Peter Macheroux

Institute of Biochemistry, Graz University of Technology, Graz, Austria

### ABSTRACT

The efficient inhibition of 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) by the broad-spectrum herbicide glyphosate validates the shikimate pathway as a promising target for developing antimicrobial, fungicidal and herbicidal agents. The last enzyme of this pathway, chorismate synthase (CS), catalyses an unusual reaction, making it an attractive target for novel inhibitors. Therefore, we tested a series of azo-dyes for their inhibitory potential against CS from the pathogenic fungus *Paracoccidioides brasiliensis* (PbCS) and identified the azo-dye PH011669 that exhibits a dissociation ( $K_d$ ) and 50% inhibitory constant ( $IC_{50}$ ) of  $1.1 \pm 0.1$  and  $10 \pm 1 \mu\text{M}$ , respectively. Molecular docking and MD simulations provided insight into the mode of inhibition, showing that PH011669 binds to the enzyme's active site primarily through electrostatic interactions. Thus, our study is the first to integrate structural and computational methods to guide future efforts towards designing the next generation of CS inhibitors.

### ARTICLE HISTORY

Received 13 May 2024  
Revised 24 September 2024  
Accepted 2 November 2024

### KEYWORDS

Azo-dyes; computational chemistry; drug design; flavoenzymes; inhibitor

### Introduction

Chorismate synthase (CS, EC 4.6.1.4) catalyses the seventh and last common step in the shikimate pathway, which is responsible for the biosynthesis of crucial aromatic compounds, such as the aromatic amino acids phenylalanine, tyrosine and tryptophan, as well as a host of other essential aromatic biomolecules (e.g. alkaloids, folates, lignin, ubiquinones, siderophores and plant secondary compounds). The enzyme-catalysed conversion of 5-enolpyruvylshikimate-3-phosphate (EPSP) to chorismate involves an anti-1,4-elimination of the phosphate and the C-(6*proR*)-hydrogen (Figure 1).

The shikimate pathway is essential for prokaryotes, protozoa, fungi and plants. The sixth enzyme of the common pathway, EPSP synthase (EPSPS), is the target of the prominent herbicide glyphosate (Figure 1), and thus, blocking the pathway is clearly deleterious for organisms that depend on a functioning shikimate pathway for the biosynthesis of essential aromatic compounds.

At first, the absolute requirement for reduced FMN in the CS-catalysed reaction was puzzling since the 1,4-anti elimination of a phosphate group and a hydrogen atom is redox neutral and thus does not require the assistance of a redox cofactor. However, it was also realised that abstraction of the 6*proR*-hydrogen is challenging owing to its high p*K*<sub>a</sub> value, estimated to be  $\approx 30^1$ . Detailed mechanistic studies involving substrate derivatives and rapid reaction kinetics, as well as density functional calculations, have provided evidence that the reduced flavin acts as a transient electron donor, thereby facilitating the bond-breaking steps<sup>2–5</sup>. The crystal structure of CS from *Streptococcus pneumoniae* (SpCS) in complex with (oxidised) FMN and its substrate EPSP agrees with such a


mechanism and rationalises the requirement of a redox cofactor in a redox neutral elimination reaction<sup>6</sup>.

In view of the success of glyphosate as one of the leading herbicides worldwide, CS appeared to be a very attractive drug target<sup>7,8</sup>. First attempts to identify an inhibitor of CS were already initiated in the 1990s before the first crystal structure was reported by Mclean and Ali, and thus mainly relied on the structure of EPSP as a lead to synthesise substrate analogs with potential inhibitory properties<sup>6</sup>. As CS has a strict stereospecificity for eliminating the *proR*-hydrogen, fluorinated substrate analogs were synthesised, i.e. 6(S)- and 6(R)-6-fluoro-EPSP<sup>9</sup>. It was shown that 6(S)-6-fluoro-EPSP is a very slow substrate exhibiting a 200-fold slower turnover than EPSP<sup>10</sup>. This result suggested that the reported antimicrobial properties of 6-fluoro-shikimate are probably not due to the inhibition of CS but the inhibition of 4-aminobenzoic acid synthesis, as proposed earlier<sup>11</sup>. On the other hand, 6(R)-6-fluoro-EPSP lacks the (6*proR*)-hydrogen and is therefore not a substrate of CS. This compound was found to be a competitive inhibitor of CS from *Neurospora crassa*, exhibiting a  $K_i$  of  $3.0 \pm 0.3 \mu\text{M}$ <sup>9</sup>. However, it is still unclear whether the reported antimicrobial activity of (6R)-6-fluoro-shikimate is due to CS inhibition.

In 2003, the use of an extensive library of diverse small molecules to screen for an inhibitor of SpCS was reported<sup>12</sup>. Based on the initial hit, a benzofuran-3[2H]-one (BF1) derivative with a 50% inhibitory constant ( $IC_{50}$ ) of  $8 \mu\text{M}$ , modification of the alkoxy chain length of the phenyl ring (BF2) resulted in a further improvement of the inhibitory properties, as shown in Figure 2. Unfortunately, further studies to characterise the potency and applicability of these compounds were not reported.

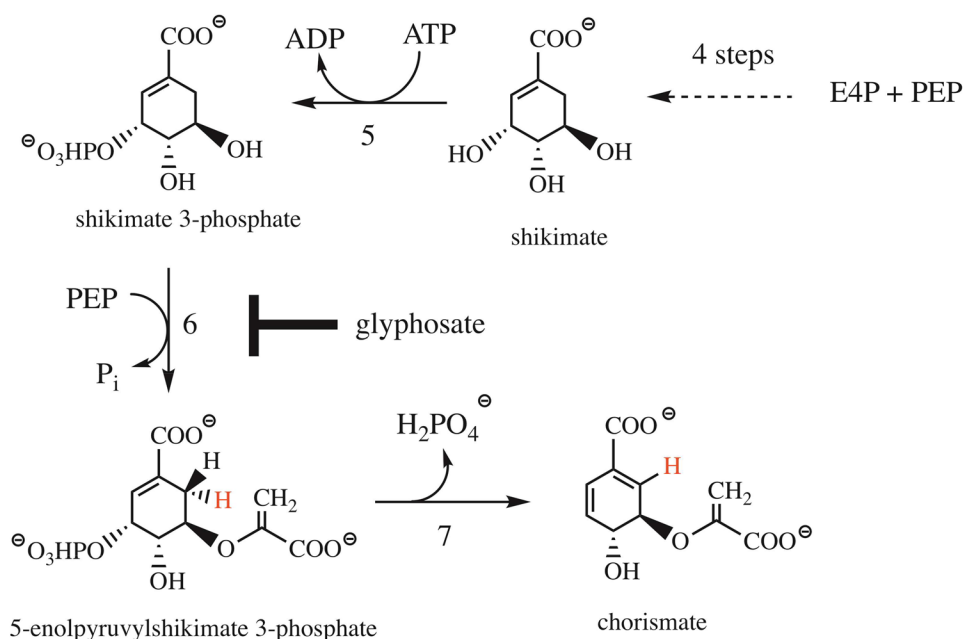
**CONTACT** Peter Macheroux  [peter.macheroux@tugraz.at](mailto:peter.macheroux@tugraz.at)  Institute of Biochemistry, Graz University of Technology, Petersgasse 12/II, Graz A-8010, Austria

\*These authors contributed equally

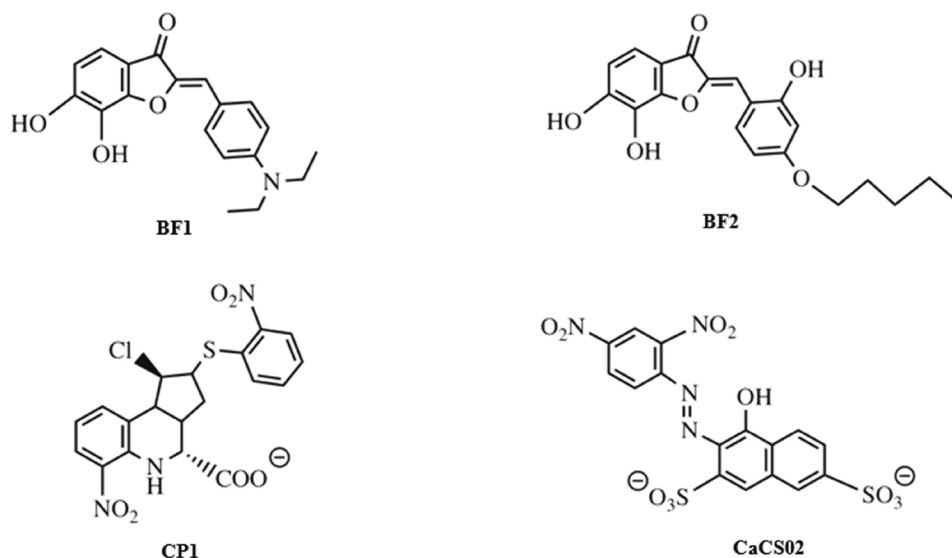
 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/14756366.2024.2427175>.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.



**Figure 1.** The last steps of the shikimate pathway. Erythrose 4-phosphate (E4P) and phosphoenolpyruvate (PEP) are used as building blocks to synthesise shikimate (top right) in four enzymatic steps. Shikimate is phosphorylated by shikimate kinase at the expense of ATP to shikimate-3-phosphate (step 5). In step six, 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) catalyses the conversion of shikimate-3-phosphate to 5-enolpyruvylshikimate-3-phosphate (EPSP, left). EPSPS is the target of the broad-band herbicide glyphosate. The seventh step of the common pathway is catalysed by chorismate synthase involving the stereospecific *anti*-1,4-elimination of the 3-phosphate and the  $\delta$ *proR*-hydrogen (bottom).



**Figure 2.** CS inhibitors reported by Thomas<sup>12</sup> and Rodrigues-Vendramini<sup>13</sup>. Top panel: initial hit (left) and the improved inhibitor (right) reported by Thomas and colleagues<sup>12</sup>. Bottom panel: inhibitors found by virtual screening by Rodrigues-Vendramini and colleagues<sup>13</sup>.  $\text{IC}_{50}$  values for compounds BF1, BF2, CP1 and CaCS02 were 8, 0.22, 47 and 29  $\mu\text{M}$ , respectively.

More recently, virtual screening using a three-dimensional structure of CS from *Paracoccidioides brasiliensis* (*PbCS*, homology model based on the crystallographic structure of the *Saccharomyces cerevisiae* enzyme) in combination with molecular dynamics calculations led to the discovery of two inhibitors, a quinoline derivative (1*S*,2*S*,3*aS*,4*S*,9*bR*)-1-chloro-6-nitro-2-(2-nitrophenyl)sulfanyl-2,3,3*a*,4,5,9*b*-hexahydro-1*H*-cyclopenta[*c*]quinoline-4-carboxylic acid (CP1) and the azo dye nitrazine yellow 1-(2,4-dinitrophenylazo)-1-hydroxynaphthalene-3,6-disulfonic acid (CaCS02) (Figure 2). The dimorphic fungus *P. brasiliensis* was chosen as a target because it causes paracoccidioidomycosis (PCM), an invasive fungal disease,

which is the most prevalent systemic mycosis in Latin America. In Brazil, for example, PCM is responsible for approximately half of the deaths caused by systemic mycoses<sup>14,15</sup>.

Compounds CP1 and CaCS02 bind to *PbCS* with dissociation constants of  $64 \pm 1$  and 20  $\mu\text{M}$  and inhibit the enzyme with  $\text{IC}_{50}$  values of  $47 \pm 5$  and  $29 \pm 3 \mu\text{M}$ , respectively<sup>13</sup>. Since several azo-dyes featuring different substitution patterns at the phenyl and naphthyl ring are commercially available, we decided to extend our study by testing a series of azo-dyes for their binding affinity to *PbCS*, leading to the discovery of an azo-dye (PH011669) with a ca. 20-fold higher affinity than the initial hit and an  $\text{IC}_{50}$  value of

10  $\mu\text{M}$ . To gain further insight into the binding of PH011669 to *PbCS*, we conducted molecular docking and molecular dynamics (MD) simulations.

## Materials and methods

### Azo-dyes tested in the current study

Supplementary Table 1 provides the names and structures of the azo-dyes investigated in the current study. The compounds were purchased from Sigma-Aldrich and used without further purification.

### Production and purification of CSs

CSs from 16 organisms were produced. The genes coding for the respective proteins were cloned and protein production was performed in different pET *E. coli* T7 expression vectors and *E. coli* strains (Supplementary Table 2). The pET-21a(+) plasmid containing the gene coding for *PbCS* was received from Dr. Flavio AV Seixas (State University of Maringá, Brazil). Recombinant protein production and purification from all CSs, except for CS from *Mycobacterium tuberculosis*, which was produced and purified as reported by Dias and colleagues<sup>16</sup>, was adapted from Rodrigues-Vendramini and coworkers<sup>13</sup>. Briefly, *E. coli* was grown in an LB medium containing the appropriate antibiotic and inoculated with a fresh overnight culture to an optical density at 600 nm ( $\text{OD}_{600}$ ) of 0.1. The main cultures were grown to an  $\text{OD}_{600}$  of 0.7 at 37 °C and 140 rpm. To start protein production, isopropyl- $\alpha$ -D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM. After induction, the main cultures were incubated at 20 °C and 140 rpm overnight (~16 h). The cells were then harvested by centrifugation ( $5,500 \times g$  for 15 min), and the resulting cell pellet was stored at -20 °C until further use. To obtain pure protein, cell pellets were resuspended in lysis buffer (50 mM  $\text{NaH}_2\text{PO}_4$ , 150 mM NaCl, 10 mM imidazole, pH 8) and lysed by sonication ( $2 \times 10$  min) using a Labsonic P sonication probe (B. Braun Biotech, Berlin, Germany). After sonication, the lysate was centrifuged ( $38,500 \times g$  for 45 min) and loaded onto a 5 mL nickel-nitrilotriacetic acid (Ni-NTA) column. The column was equilibrated with lysis buffer, and the target protein was eluted with elution buffer (50 mM  $\text{NaH}_2\text{PO}_4$ , 150 mM NaCl, 300 mM imidazole, pH 8). Fractions containing the target protein were pooled and dialysed against 50 mM  $\text{NaH}_2\text{PO}_4$  and 150 mM NaCl (pH 8) overnight. The target protein was then concentrated using an ultrafiltration device (30-kDa cut-off, Merck-Millipore, Darmstadt, Germany). The purified protein was flash-frozen using liquid nitrogen and stored at -80 °C until further use.

### Spectrophotometric titrations/binding assay

For the binding studies, PH011669 was dissolved in 50 mM  $\text{NaH}_2\text{PO}_4$  and 150 mM NaCl at pH 8, diluted to a final concentration of approximately 10  $\mu\text{M}$  ( $A_{580\text{nm}} \sim 0.25$ ; 800  $\mu\text{L}$ ), and transferred to a quartz cuvette (with 800  $\mu\text{L}$  of buffer in the reference cuvette). After recording an initial UV-visible absorption spectrum between 300 and 800 nm, 10  $\mu\text{L}$  aliquots of *PbCS* (~55  $\mu\text{M}$ ) were added, and spectral changes were monitored after each addition. To quantify the binding affinity of PH011669 to *PbCS*, we plotted the relative absorption changes at 510 nm as a function of the protein concentration in the cuvette<sup>17</sup>.

### Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) was performed on a MicroCal PEAQ-ITC. All ITC experiments were conducted in storage buffer (50 mM  $\text{NaH}_2\text{PO}_4$ , 150 mM NaCl, pH 8) at 25 °C and a stirring rate of 750 rpm. Titration experiments consisted of 13 injections (first injection: 0.4  $\mu\text{L}$ ; duration time: 0.8 s; spacing time: 300 s; 2nd–13th injection: 3  $\mu\text{L}$ ; duration time: 6 s; spacing time: 150 s) and were performed in triplicates. The azo-dye PH011669 (500  $\mu\text{M}$ ) was loaded into the injection syringe and titrated into the sample cell containing *PbCS* (25  $\mu\text{M}$ ); the reference cell was filled with degassed  $\text{ddH}_2\text{O}$ . Dissociation constants were obtained with the MicroCal PEAQ-ITC Analysis Software (Malvern) using the “one set of sites fitting” model.

### Inhibition assay

The influence of PH011669 on the activity of *PbCS* was tested by forward coupling the CS reaction with the anthranilate synthase reaction. The final reaction solution contained 80  $\mu\text{M}$  EPSP, 4  $\mu\text{M}$  *PbCS*, 7  $\mu\text{M}$  YcnD, 20  $\mu\text{M}$  anthranilate synthase, 4 mM  $\text{MgSO}_4$ , 10 mM L-glutamine (L-Gln), 30 mM  $(\text{NH}_4)_2\text{SO}_4$ , 1 mM dithiothreitol (DTT) and different concentrations of the inhibitor (0.1–250  $\mu\text{M}$ ). The assay was prepared in 10 mM potassium phosphate at pH 7.6 and pipetted into a 96-well plate. The reaction was started by the addition of 500  $\mu\text{M}$  NADPH. *PbCS* activity could be studied spectrofluorometrically at 37 °C by following the increase in the fluorescence intensity at 390 nm ( $\lambda_{\text{ex}} = 340$  nm) for 5 min, which correlates with the amount of anthranilate formed in the course of the reaction. The slopes of the initial velocities were determined and plotted as a function of the log of the corresponding PH011669 concentration (triplicates were determined).

### Structure prediction of *PbCS*

The structure of *PbCS* was modelled using AlphaFold2 via the Collaboratory service from Google Research ([https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2\\_advanced.ipynb](https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb))<sup>18</sup>. Briefly, the sequence of *PbCS* was submitted to AlphaFold2, and, as a first step, a multiple sequence alignment was performed employing the mmseq2 method<sup>19</sup>. Five models per enzyme were predicted and ranked by pLDDT (predicted local distance difference test). The five predicted *PbCS* models were superimposed on the crystal structure of *SpCS* complexed with oxidised FMN and EPSP (PDB ID: 1QXO<sup>6</sup>). The model with the lowest Ca-RMSD to PDB entry 1QXO and the highest pLDDT score was refined with the Amber-Relax protocol of the AlphaFold2 pipeline and validated by the MOLPROBITY and RAMPAGE servers<sup>20</sup>. *PbCS* was modelled as a homotetramer (biological unit), but only one protomer was used for the subsequent docking experiment and MD simulations.

### Molecular docking

Molecular docking was performed with AutoDock Vina (version 1.2.1)<sup>21</sup> to study the binding mode of PH011669 in the active site of *PbCS* and provide initial coordinates and topology parameters for the MD simulations. The predicted AlphaFold model was used as the receptor, and an FMN molecule was added to the model before docking by superimposing the catalytic core of the crystal

structure 1QXO onto our model. The structure of PH011669 was generated with the NCI Online SMILES Translator and Structure File Generator (<https://cactus.nci.nih.gov/translate/>). The model structures of PH011669 and *PbCS*-FMN were then formatted into *pdbqt* files using AutoDockTools (version 1.5.7)<sup>22</sup>, specifying and sampling all rotatable bonds and computing partial charges for both structures. Binding poses of PH011669 were searched in a grid box of  $14 \times 16 \times 10 \text{ \AA}^3$  (spacing =  $1.0 \text{ \AA}$ ) centred on the substrate EPSP after superimposing the AlphaFold model of *PbCS* on the crystal structure 1QXO. Docking exhaustiveness was set to 50. AutoDock Vina was run 10 times to verify the consistency and reliability of the docking results. The final docking settings were validated by re-docking EPSP into the active site of 1QXO, which lacked the substrate EPSP. The resulting docking poses closely resembled the EPSP pose found in the crystal structure, verifying our docking protocol. Upon docking, the binding poses of PH011669 were compared to that of EPSP in 1QXO and evaluated by docking scores and for reasonability. Poses that deviated significantly from the pose of EPSP in the reference structure 1QXO were excluded. The pose with the most reasonable orientation and highest docking score was selected as input for the following MD simulations.

### Molecular dynamics

The MD experiment was performed using GROMACS 2022<sup>23</sup>, as previously described<sup>24</sup>. The receptor (i.e. *PbCS*-FMN) and ligand (i.e. PH011669) GRO coordinates and topology files were generated and merged to obtain the complex (i.e. *PbCS*-FMN-PH011669) GRO and topology files. The flavin cofactor was parameterised in its reduced form (i.e. FMNH<sub>2</sub>), as the enzyme requires reduced FMN for activity.

Three systems were investigated by MD: apo*PbCS*, *PbCS*-PH011669 (binary system) and *PbCS*-FMN-PH011669 (ternary system). The MD unit cell was defined as a dodecahedron with a minimum distance of 1.0 nm between the complex and the box edges. The system was solvated with the SPC/E water model and neutralised by adding Na<sup>+</sup> ions. The system was then subjected to standard energy minimization and equilibration. Energy minimization was performed using the steepest descent algorithm with a maximum force of 1000 kJ/mol/nm and a maximum of 5000 steps. The equilibration step consisted of two phases: NVT (constant number of particles, volume and temperature) and NPT (constant number of particles, pressure and temperature). The NVT phase was run for 100 ps with a temperature coupling of 300 K using the v-rescale thermostat. The NPT phase was run for 200 ps with a pressure coupling of 1 bar using the Parrinello-Rahman barostat.

Initially, three independent MD production runs were performed for each system starting from the equilibration output, each for 300 ns with initial velocity generation (i.e. 900 ns for each system). For the binary system, we needed to increase the sampling and thus performed six more independent runs of 300 ns. The parameters for the production runs were the same for all systems: temperature and pressure were set to 300 K and 1 bar by the v-rescale thermostat and Parrinello-Rahman barostat, respectively; bonds were constrained using the LINCS algorithm; the Verlet cutoff scheme was used to process intra-atomic interactions; the PME method was implemented to account for Coulombic and Lennard-Jones interactions; and a van der Waals cutoff radius of 1.0 was applied.

Note that only one protomer was used for all calculations to reduce computational cost.

### Molecular dynamics analysis

The analysis density module of the MDAnalysis python toolkit<sup>25,26</sup> was applied to calculate residence densities for PH011669 from the trajectories of the binary and ternary systems using default parameters. The resulting density maps were visualised as volumes (carve =  $1.6 \text{ \AA}$ ) using PyMOL<sup>27</sup>.

A principal component analysis (PCA) of the PH011669 across the binary and ternary complex trajectories was performed. First, all the trajectories were aligned on the receptor backbone; then, the PH011669 coordinates were extracted into separate sub-trajectories. The full ensemble of these sub-trajectories was used to train the PCA algorithm provided by the MDAnalysis.analysis package, which was then used to transform the individual ones. Finally, the first two principal components were used to generate the data for the kernel density estimate (KDE) plots.

### Binding free energy calculations

Molecular mechanics/Poisson-Boltzmann surface area (MM/PBSA) calculations were performed with the program *gmx\_MMPBSA*<sup>28</sup> to approximate binding energies for the enzyme-inhibitor complexes. The last 30 ns of each production run (i.e.  $3 \times 30$  ns for the ternary and  $9 \times 30$  ns for the binary system) were used for energy calculations. The internal dielectric constant of the solute, the internal dielectric constant of the gas phase, temperature, salt concentration and Generalised Born method value were set to 1, 80, 298.15 K, 0.15 M and 5, respectively. Per-residue energy decomposition analysis was performed using default parameters to determine the binding energy contribution of all residues within  $6 \text{ \AA}$  of PH011669. The obtained data were averaged over three (ternary system) and nine (binary system) independent trajectories.

### ADME prediction

The absorption, distribution, metabolism and excretion (ADME) properties for PH011669 and an optimised version of PH011669 were assessed using the SwissADME web tool<sup>29</sup>. For this purpose, the SMILES string for PH011669 and its optimised version were submitted to the server.

## Results and discussion

### Azo-dyes as inhibitors of chorismate synthases

As the azo-dye CaCS02 (Figure 2) showed a promising dissociation constant ( $20 \mu\text{M}$ ) and  $\text{IC}_{50}$  value ( $29 \mu\text{M}$ ), we screened a series of structurally similar and commercially available azo-dyes for their binding affinity using *PbCS*. The binding affinities of azo-dyes listed in Table 1 were determined by spectrophotometric titrations employing *PbCS* as the target enzyme. As shown in Figure 3, binding affinities varied over a wide range, with PH011669 exhibiting by far the lowest dissociation constant of  $1.1 \pm 0.1 \mu\text{M}$ .

In addition to the spectrophotometric titrations, we also determined the binding affinity of PH011669 using ITC. As shown in Figure 4, the binding of PH011669 to *PbCS* was an exothermic process. The experimental data were fitted with a single site binding model and yielded a dissociation constant of  $3.8 \pm 1.8 \mu\text{M}$ , which is comparable to that obtained by spectrophotometric titration ( $1.1 \pm 0.1 \mu\text{M}$ ).

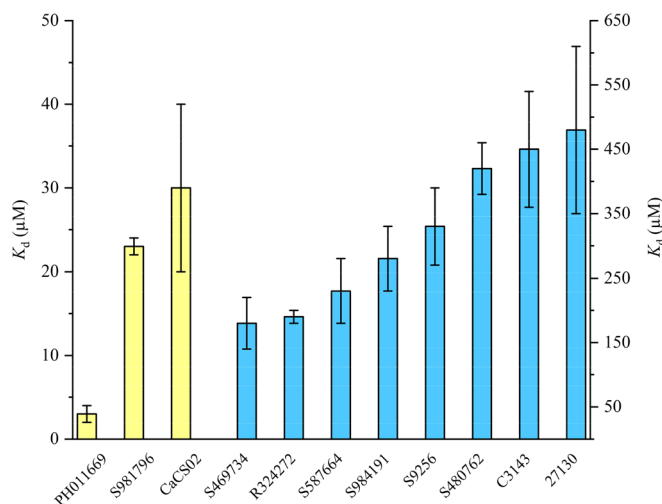


**Table 1.** Affinity of PH011669 towards CSs from different organisms.

Species	Classification	Yield <sup>a</sup>	$K_d$ ( $\mu\text{M}$ )
<i>Anabaena variabilis</i>	cyanobacterium	+++	1.8 $\pm$ 0.1
<i>Aquifex aeolicus</i>	Gram <sup>-</sup> bacterium	+++	3.2 $\pm$ 0.1
<i>Botrytis fuckeliana</i>	fungus	+++	4.0 $\pm$ 0.6
<i>Candida albicans</i>	fungus	++	4.6 $\pm$ 0.7
<i>Corydalis sempervirens</i>	plant	+	1.0 $\pm$ 0.1
<i>Escherichia coli</i>	Gram <sup>-</sup> bacterium	++	3.7 $\pm$ 1.0
<i>Methanobacterium thermoautotrophicum</i>	Gram <sup>+</sup> bacterium	+	2.7 $\pm$ 0.0
<i>Mycobacterium tuberculosis</i>	Gram <sup>+/-</sup> bacterium	n. s.	-
<i>Neurospora crassa</i>	fungus	+	3.5 $\pm$ 0.1
<i>Paracoccidioides brasiliensis</i>	fungus	+++	1.1 $\pm$ 0.0
<i>Saccharomyces cerevisiae</i>	yeast	+++	2.4 $\pm$ 0.1
<i>Solanum lycopersicum</i>	plant	++	1.8 $\pm$ 0.1
<i>Staphylococcus aureus</i>	Gram <sup>+</sup> bacterium	+	3.7 $\pm$ 0.1
<i>Tetrahymena thermophila</i>	unicell. eukaryote	+	4.4 $\pm$ 0.1
<i>Thermotoga maritima</i>	Gram <sup>-</sup> bacterium	+	1.5 $\pm$ 0.0
<i>Toxoplasmodium gondii</i>	protozoon	+	1.7 $\pm$ 0.0

<sup>a</sup>Yields of purified protein are (+) = 1–4.9 mg/g cell wet weight, (++) = 5–9.9 mg/g cell wet weight and (+++) > 10 mg/g cell wet weight.

n. s.: not soluble;  $K_d$ : dissociation constant of PH011669.



**Figure 3.** Azo-compounds PH011669, S981796, CaCS02, S469734, R324272, S587664, S984191, S9256, S480762, C3143 and 27130 ranked according to their  $K_d$ -values against *PbCS*. To quantify the binding affinities of the azo-compounds to *PbCS*, the relative absorption changes at a certain wavelength were plotted as a function of the protein concentration in the cuvette, with the standard deviations displayed as error bars (from three titrations). The scales for the yellow and blue bars are indicated on the left and right sides, respectively.

To gain insight into the affinity of PH011669 to CSs from other sources, we successfully produced fifteen CSs from bacterial, fungal and plant sources. As shown in Table 1, the highest and lowest affinity was found with the CS from the plant *Corydalis sempervirens* (1.0 $\pm$ 0.1  $\mu\text{M}$ ) and the fungus *Candida albicans* (4.6 $\pm$ 0.7  $\mu\text{M}$ ), respectively. Thus, PH011669 possesses low micromolar dissociation constants across the CSs used in our study. This indicates that the interactions governing the binding of PH011669 are fairly conserved in the active site of CSs.

### Inhibition assay

To test the inhibitory effect of PH011669 on the activity of CS, we used a coupled assay involving EPSP, CS, YcnD and anthranilate synthase. Our inhibition assay has been used to determine the ability of a drug to reduce the activity of the enzyme. The

resulting  $\text{IC}_{50}$  value gives the concentration of a drug required for 50% inhibition of the enzymatic activity. Using a nonlinear curve fit based on the Hill function, an  $\text{IC}_{50}$  value of 10 $\pm$ 1  $\mu\text{M}$  was determined (Figure 5). A comparable compound tested with *PbCS* provided an  $\text{IC}_{50}$  value of 29 $\pm$ 3  $\mu\text{M}$ , indicating an improved inhibitory activity<sup>17</sup>. Compared to the existing value and combined with the data obtained, PH011669 is an excellent *PbCS* binder and a strong inhibitor.

### Evaluation of the predicted *PbCS* model

As no crystal structure for *PbCS* is available, we generated an AlphaFold model. The final model exhibited an excellent average pLDDT value of 91.9% and good stereochemical quality, with 94.0% of the residues being in the favoured region of the Ramachandran plot (Figure 6). The final *PbCS* model represents the active conformation of the enzymes, as the active site loops were modelled in their closed conformation. The predicted structure comprises 11  $\alpha$ -helices and 12  $\beta$ -sheets, with its core adopting the  $\beta$ - $\alpha$ - $\beta$  sandwich fold, which is unique to CSs<sup>6</sup>. An FMN molecule was incorporated by superimposing the binding site of our model on that of the crystal structure of the ternary complex *SpCS*-FMN-EPSP (PDB ID: 1QXO). The calculated Ca-RMSD between our model and the experimentally determined crystal structure 1QXO is 1.25  $\text{\AA}$ , further corroborating the accuracy of our model. As AlphaFold models fail to account for conformational flexibility upon ligand binding, it was important that our model resembles the ligand-bound structure 1QXO as much as possible.

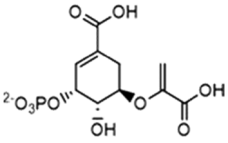
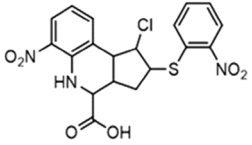
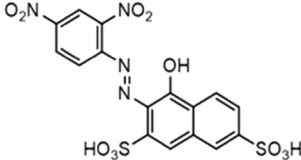
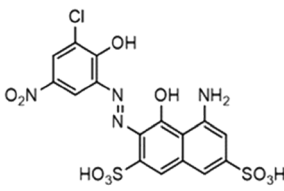
### Molecular docking

To gain more information on the interactions between PH011669 and *PbCS*, we performed molecular docking using the AlphaFold2-generated *PbCS* model, including the cofactor FMN. In addition to PH011669, the substrate EPSP and the previously identified CS inhibitors CP1 and CaCS02 were docked into the active site of *PbCS*-FMN for comparison<sup>13,17</sup>. PH011669 exhibited the best docking score and predicted inhibition constant ( $K_i$ ) among all tested ligands, outperforming the natural substrate EPSP, although both compounds are involved in the same number of polar interactions (Table 2). This is corroborated by a direct comparison of the binding affinity of EPSP and PH011669 to the CS from *N. crassa*, exhibiting  $K_d$  values of 17  $\mu\text{M}$ <sup>30</sup> and 3.5  $\mu\text{M}$ , respectively.

As the catalytic pocket of *PbCS* is positively charged, PH011669 binding to *PbCS* is governed by ion-ion interactions between the inhibitor's negatively charged sulfonate groups and the enzymes' positively charged arginine residues (Figure 7A). The sulfonate groups of PH011669 interact with R46, R362 and R103, exhibiting bond distances between 2.9 and 3.7  $\text{\AA}$  (Figure 7B). Besides these electrostatic interactions, PH011669 is involved in a strong H-bond with the backbone amine of A128 (3.0  $\text{\AA}$ ) and weaker ones with T54 (3.6  $\text{\AA}$ ), T55 (3.1  $\text{\AA}$ ), S127 (3.4  $\text{\AA}$ ) and R354 (3.1 and 3.6  $\text{\AA}$ ), which contribute only little to the binding of the inhibitor (Figure 7B).

Comparison of the docking results of PH011669 and EPSP show that the inhibitor and substrate exhibit comparable binding modes with *PbCS* (Figure 7C). Both molecules interact with the same set of arginine residues (i.e. R46, R103 and R362); however, in contrast to PH011669, EPSP is engaged in further salt bridges with R129

**Table 2.** Docking scores, predicted  $K_i$  values and number of polar interactions.

Ligand	Docking score [kcal/mol]	$K_i$ [nM]	No. of polar interactions <sup>a</sup>
	$-8.4 \pm 0.1$	$651 \pm 188$	9
EPSP 	$-7.9 \pm 0.6$	$587 \pm 123$	5
CP1 	$-9.8 \pm 0.2$	$61 \pm 19$	8
CaCS02 	$-10.2 \pm 0.2$	$37 \pm 9$	9
PH011669			

Docking scores and  $K_i$ -values were calculated with AutoDock Vina and represent the average of 10 independent runs.

<sup>a</sup>Number of polar interactions between the ligand and the enzyme in the best docking pose (i.e. best docking score).

and R354 via its 3-phosphate group. Superimposition of EPSP and PH011669 show that their carboxy and sulfonate groups, respectively, are located at approximately the same positions within the binding site; the distance between the carboxy groups in EPSP is almost identical to that between the sulfonate groups in PH011669 (Figure 7C,D). Thus, the naphthalene-2,7-disulfonate moiety of PH011669 seems to mimic the 5-enolpyruvylshikimate core of EPSP within the enzyme's binding pocket. While the whole EPSP molecule is buried in the binding pocket of *PbCS*, only the 5-amino-4-hydroxynaphthalene-2,7-disulfonate core of PH011669 is buried, with its 3-chloro-2-hydroxy-5-nitrophenyl group protruding out of the pocket into the bulk solvent.

### Molecular dynamics simulations

We used MD simulations of 300 ns to study the interactions between PH011669 and *PbCS* in more detail. Simulations were run in the presence and absence of FMN because, in a previous study, we observed that the inhibitory compound CP1 bound slightly more tightly to the flavin-free than to the flavin-containing enzyme<sup>13</sup>. Simulations were also performed for *apoPbCS* as a reference.

Root mean square deviation (RMSD) analysis showed that all systems, *apoPbCS*, *PbCS*-PH011669 and *PbCS*-FMN-PH011669,

reached equilibrium after approximately 50 ns and remained stable throughout the simulation (Figure 8A). In the inhibitor-bound systems, PH011669 remained stable in the binding pocket throughout the simulation; only the 3-chloro-2-hydroxy-5-nitrophenyl group protruded from the binding pocket at times. The RMSD values of the inhibitor's atoms are higher in the flavin-lacking *PbCS*-PH011669 system than those in the *PbCS*-FMN-PH011669 system, indicating less specific binding of PH011669 in the binary system, which will be addressed later (Figure 8B).

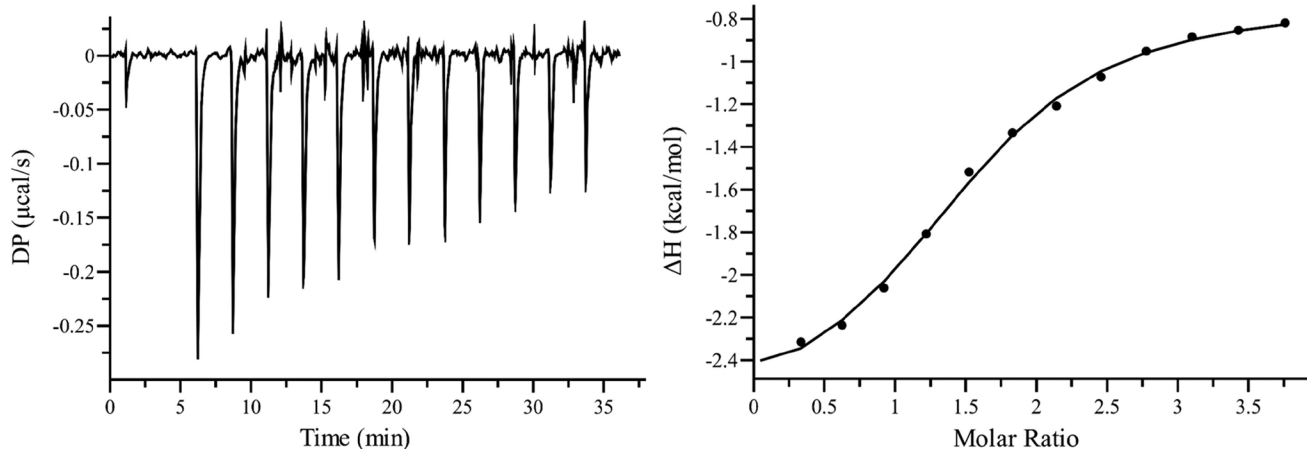
The root mean square fluctuation (RMSF) profiles show that the amino acid residues in the PH011669-bound systems fluctuated more than those in the apo system (Figure 8C). Besides the termini, nine regions exhibited pronounced residue flexibility (RMSF > 3 Å) in the inhibitor-bound systems, most of which were loop regions. Five of these nine flexible regions constitute the substrate binding site, which comprises the loops L1, L3, L4, L6–L8 and L20–L22, the  $\alpha$ -helices  $\alpha 2$ – $\alpha 4$  and the  $\beta$ -sheets  $\beta 11$  and  $\beta 12$ . The inhibitor-bound complexes exhibited lower RMSD but higher RMSF values than *apoPbCS*, suggesting that PH011669 induces local structural changes within the binding region during the binding event (Figure 8A,C).

### MD Analysis of the *PbCS*-FMN-PH011669 system

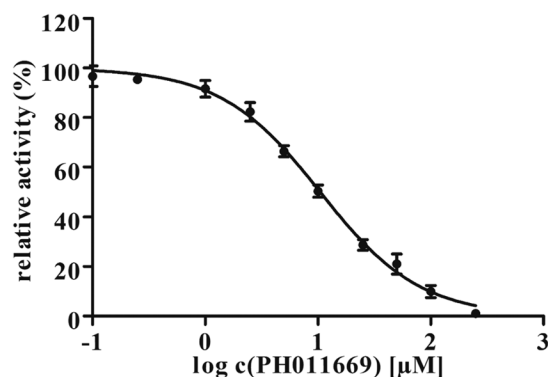
PH011669 exhibited only one binding mode in the ternary complex throughout the simulation. The inhibitor was stably bound within the binding site and showed no significant fluctuations, except for its 3-chloro-2-hydroxy-5-nitrophenyl moiety, which underwent considerable rotations and movements during the simulation. The 3-chloro-2-hydroxy-5-nitrophenyl group was mainly located at the entrance of the active site in direct contact with the bulk solvent and failed to engage in significant interactions with the enzyme, explaining its high flexibility for most of the simulation time.

As predicted by molecular docking, PH011669 binding is dominated by ionic interactions between the inhibitor's negatively charged sulfonate groups and the enzyme's positively charged amino acid residues: R46, R103, R129 and R362. These main binding contributors were within interaction distance with PH011669 (<4 Å) throughout the simulation, except for R362, which at the end of the simulation slightly lost contact with PH011669, as it shifted toward D356 to form a salt bridge (Supplementary Figure 1A–D). R46 and R362 interacted with the inhibitor's 7-sulfonate group, whereas R103 bound to its 2-sulfonate group (Supplementary Figure 1A,B,D). Besides interacting with the 2-sulfonate group, R103 was also engaged in a moderate electrostatic interaction with the nitro group of the 3-chloro-2-hydroxy-5-nitrophenyl moiety, which, however, failed to stabilise this group notably (Supplementary Figure 1B). R129 was mainly located above the naphthalene moiety of PH011669, exhibiting cation- $\pi$  interactions with the ring system for a significant fraction of the simulation time. Furthermore, it interacted moderately with the 7-sulfonate group and very weakly with the 2-hydroxy group of the 3-chloro-2-hydroxy-5-nitrophenyl moiety (Supplementary Figure 1C).

Even though the active site mainly comprises flexible regions, R46, R103, R129 and R362 did not significantly fluctuate in their positions, as evidenced by their location in the RMSF plot (Supplementary Figure 1E). Supplementary Figure 1F,G show the interactions between PH011669 and the enzyme using the energy-minimised structure of *PbCS*-FMN-PH011669 as a representative structure, as most of these interactions persisted during the simulation.



**Figure 4.** Determination of the dissociation constant of PH011669 to *PbCS* by isothermal titration calorimetry. The experiment was performed in triplicates, and data were fitted with a single site binding model.



**Figure 5.** The inhibitory effect of PH011669 on the catalytic activity of chorismate synthase from *P. brasiliensis*. The assay was performed in triplicate, with the standard deviation shown as error bars; where no error bars are shown, the statistical error is smaller than the dot.

We also determined the contact frequency (i.e. the time in percent that atoms of a given residue were within 4 Å of PH011669) and type of interaction for each residue that was interacting with PH011669 during the simulation to provide a comprehensive picture of PH011669 binding to *PbCS*-FMN (Supplementary Table 3).

#### MD Analysis of the *PbCS*-PH011669 system

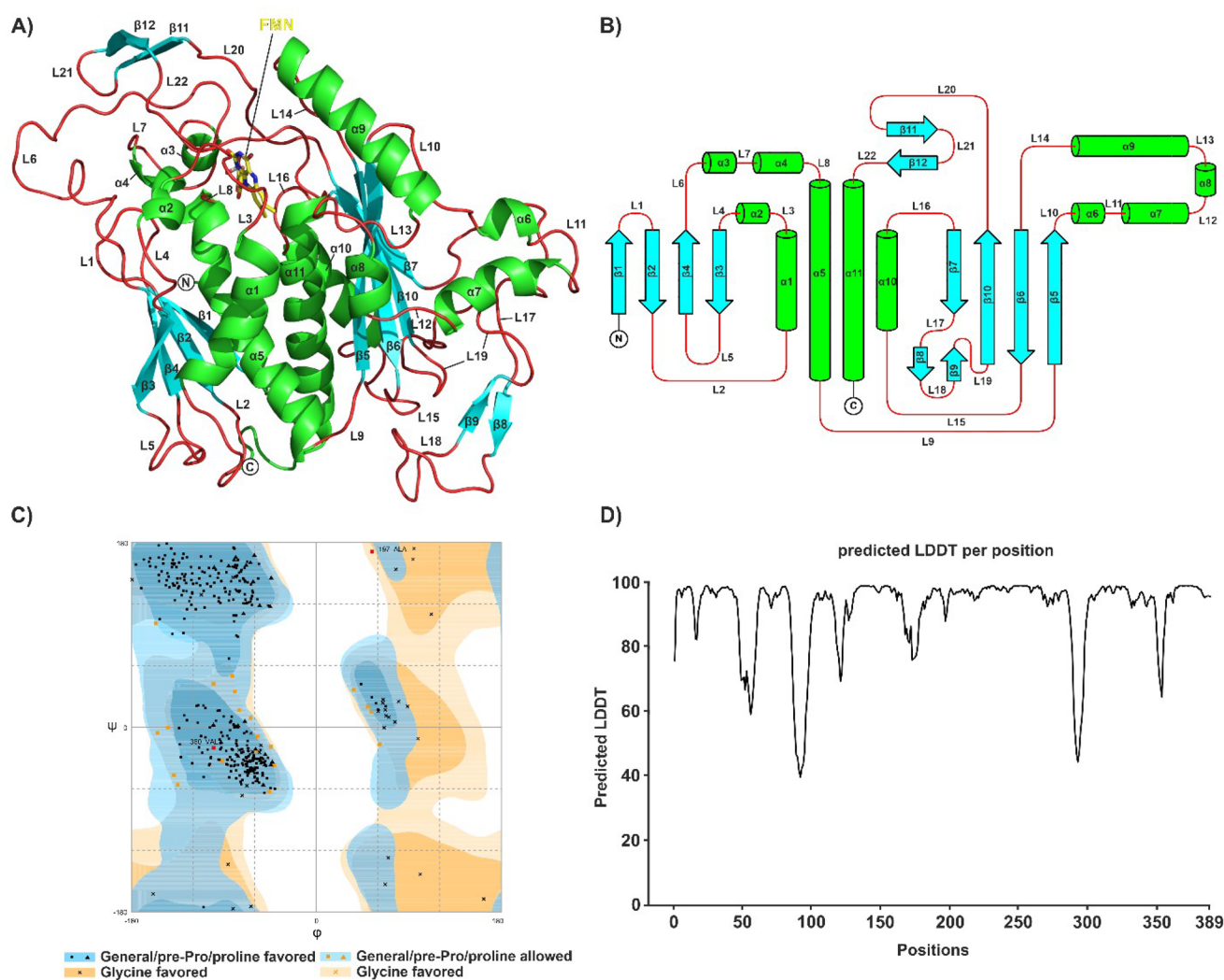
In contrast to the ternary *PbCS*-FMN-PH011669 system, the *PbCS*-PH011669 system did not exhibit a single dominant binding mode. Initially, we ran three independent MD runs for each system to ensure the robustness of the simulation, but in the binary system, the inhibitor exhibited a different binding mode in each trajectory. Therefore, we performed six more runs (in total, nine 300-ns runs) to improve the conformational sampling (and statistics) and identify a dominant binding mode for PH011669 in the flavin-free system.

After increasing the sampling, we investigated the range of motion of PH011669 during the MD simulations by calculating its residence density with the density tool of the MDAAnalysis toolkit<sup>25,26</sup>. This tool generates a three-dimensional grid and counts the number of times an atom of the inhibitor has spent at each grid point during the simulation, providing a time-averaged distribution of the inhibitor's atoms within a specific volume (i.e. the

enzyme's binding pocket). The analysis revealed that despite increasing the sampling, PH011669 still exhibited a markedly larger and more diffused residence density in the flavin-free than in the flavin-containing system, confirming that the inhibitor sampled a wide range of conformations and poses during the simulation of the binary system (Figure 9A,B). On the other hand, the inhibitor produced a concentrated and well-defined residence density in the ternary system, verifying the existence of only one dominant binding mode.

To better understand the dynamic behaviour of PH011669, we carried out PCA on the Cartesian coordinates of the inhibitor in both systems. The PCA plots corroborate the presence of one major binding mode for the inhibitor in the ternary system and multiple binding poses and conformations in the binary system; however, the plot of the binary system also shows a cluster of similar inhibitor conformations (Figure 9C,D). Looking at the PCA plots of the averaged trajectories (Figure 9C,D) and those of each trajectory separately (Supplementary Figure 2), it is evident that the inhibitor is relatively stable in each trajectory but not across the trajectories. Therefore, we concluded that there are several stable binding poses for PH011669 within the flavin-free system, and once the inhibitor adopts one of these poses, it remains stably bound. This could explain the previously observed slightly higher affinity of the inhibitory compound CP1 toward FMN-free *PbCS*, as the inhibitor may have exhibited multiple stable binding conformations in the absence of FMN, outperforming the binding of the inhibitor's primary binding mode in FMN-loaded *PbCS*<sup>13</sup>.

The MD trajectories of the *PbCS*-PH011669 system showed binding poses for the inhibitor that were similar to that in the *PbCS*-FMN-PH011669 system (highest populated poses in the binary system) but also poses in which the inhibitor intruded deeper into the binding pocket, but only partially overlapping with the FMN binding site. PH011669 interacted mainly with the same set of arginine residues as in the ternary complex (i.e. R46, R103, R129 and R362) and was therefore electrostatically trapped above the FMN binding site for most of the simulation time. Figure 9E,F show the enzyme-inhibitor interactions of a representative structure from a trajectory that is located in the center of the major cluster in the PCA plot, thus representing the most populated binding pose for the binary system (Figure 9D). In this binding pose, PH011669 is located above the FMN binding site and interacts via ionic interactions with R46, R103, R129 and



**Figure 6.** AlphaFold prediction and evaluation. (A) AlphaFold2 model of *PbCS* colour-coded by secondary structure ( $\alpha$ -helices, green;  $\beta$ -sheets, cyan; loops, red). FMN is shown as yellow sticks; secondary structure elements are annotated. (B) Topology plot of *PbCS* colour-coded according to (A). (C) Ramachandran plot of *PbCS*. Of all residues, 94.0%, 5.4% and 0.6% are in the favoured, allowed and outlier regions, respectively. Red squares indicate residues (A197 and V360) in the outlier region of the plot. (D) The predicted-LDDT-per-residue plot for the AlphaFold2 model of *PbCS*.

R362 and via H-bonding with Y101, H106 and the backbone of A128.

Collectively, the lack of FMN seems to lead to a less specific—but still tight—binding of PH011669 to *PbCS* compared with that in the presence of FMN. The main contributors to PH011669 binding (i.e. R46, R103, R129 and R362) are spatially orchestrated so that ligand binding above the FMN binding site is favoured in CSs. However, our MD simulations also suggest that in the absence of FMN, PH011669 exploits the increased space and freedom of movement within the active site to a certain degree, exhibiting multiple stable binding conformations. We did not observe any significant FMN-induced structural changes that affected inhibitor binding. This is also corroborated by the fact that the active sites of experimentally determined apo structures of CSs are almost identical to that of the FMN-bound CS structure 1QXO, exhibiting  $\alpha$ -RMSD values between 0.7 and 1.5 Å.

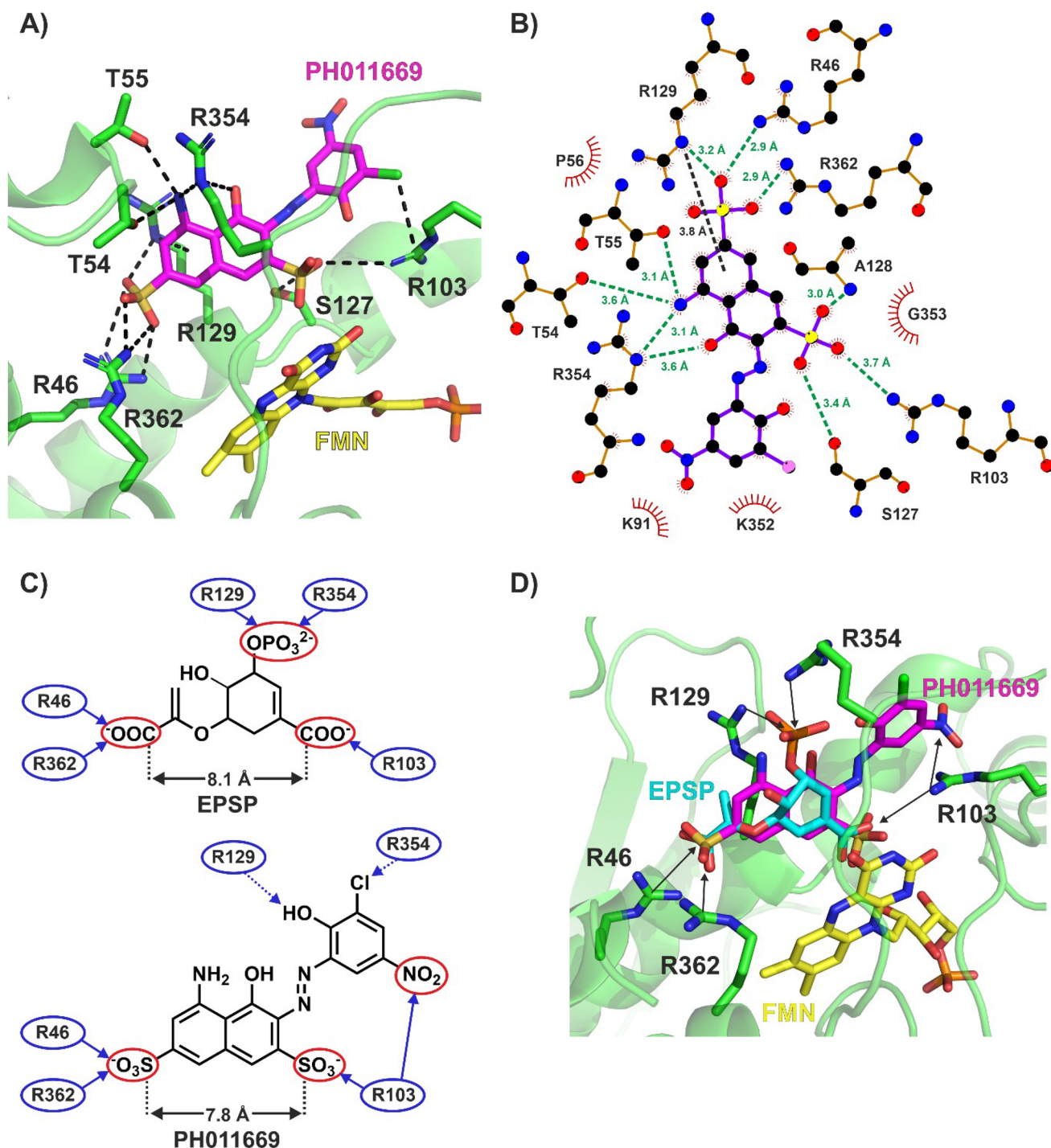
#### Binding free energy estimation and decomposition

To evaluate the affinity of PH011669 to both systems and verify our findings, we estimated the binding free energies using the MM/PBSA

method. The average Gibbs binding free energies ( $\Delta G_{\text{binding}}$ ) for the *PbCS*-FMN-PH011669 and *PbCS*-PH011669 systems were  $-50.2$  kcal/mol and  $-39.6$  kcal/mol, respectively (Table 3). MM/PBSA calculations confirmed that electrostatic interactions ( $\Delta E_{\text{EL}}$ ) govern PH011669 binding in both systems. However, the positive solvation-free energy change ( $\Delta G_{\text{SOL}}$ ) and the negative gas-phase contribution to the free energy change ( $\Delta G_{\text{GAS}}$ ) indicate strong interactions between PH011669 and the solvent. This is corroborated by trajectory structures in which the 3-chloro-2-hydroxy-5-nitrophenyl group of the inhibitor was highly solvent-exposed in both systems. Thus, our Gibbs binding free energy analysis suggests that reducing the polarity and size of the inhibitor should be considered to improve its binding to *PbCS*.

To confirm and identify the key residues for PH011669 binding, we performed a per-residue decomposition analysis and determined the contribution of all inhibitor-binding residues to  $\Delta G_{\text{binding}}$ . In both systems, R46, R103, R129 and R362 contributed the most to  $\Delta G_{\text{binding}}$  (Figure 10). However, the total binding energy of PH011669 was distributed more evenly over more residues in the binary than in the ternary system, corroborating the less specific binding of PH011669 in the binary complex.



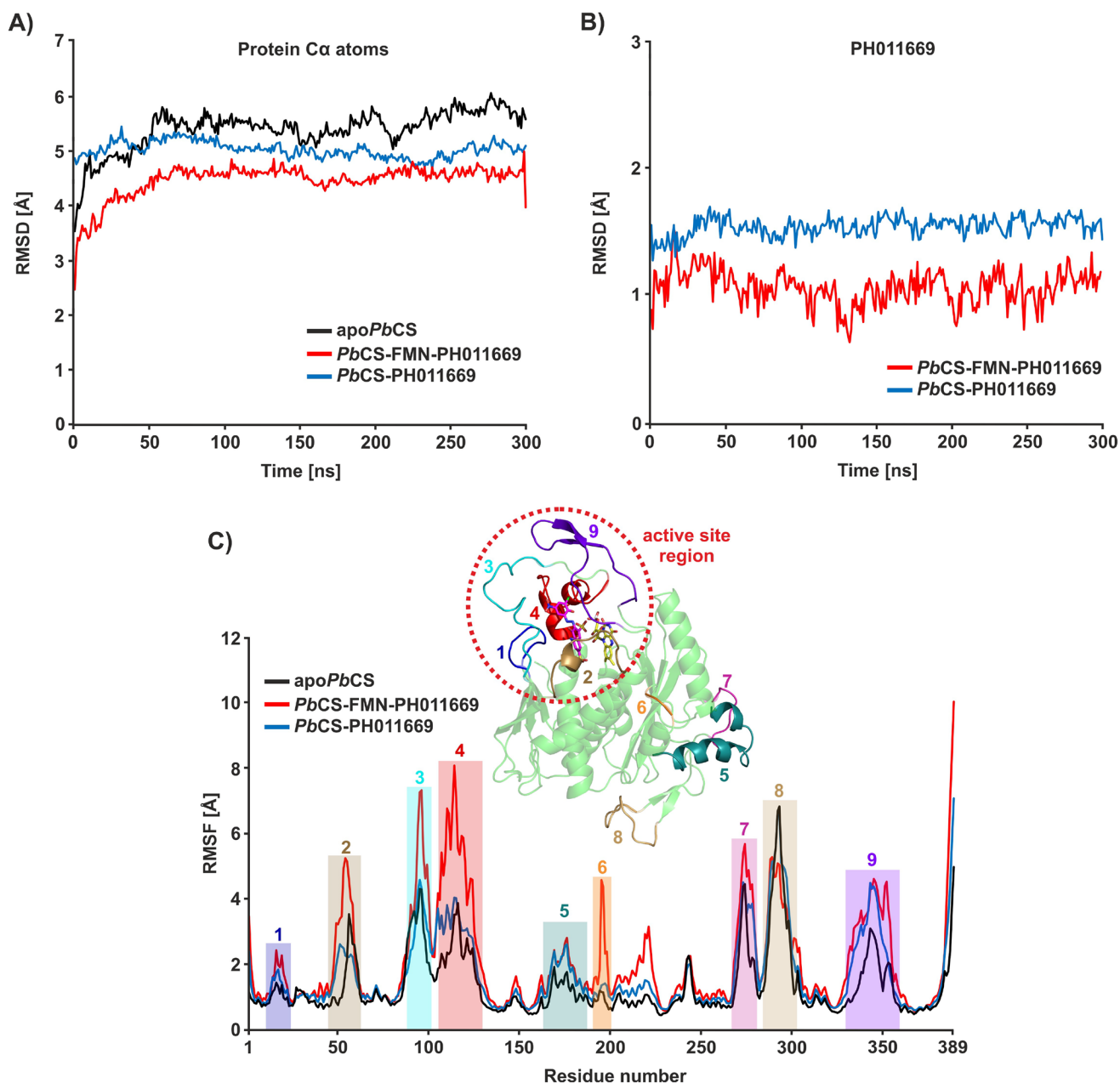


**Figure 7.** Interactions between PH011669 and *PbCS*-FMN, as predicted by molecular docking. (A) Best binding pose of PH011669 within the active site of *PbCS*-FMN. The enzyme is illustrated as a transparent green cartoon, with interacting residues shown as sticks. Black dashed lines indicate electrostatic and polar interactions. The interaction between R129 and the inhibitor is a cation- $\pi$  interaction. (B) Protein-ligand interaction plot (LigPlot) of the calculated *PbCS*-FMN-PH011669 complex. Green dashed lines indicate electrostatic and H-bond interactions while the black dashed line indicates cation- $\pi$  interactions; bond distances are also shown. (C) Schematic comparison of the binding mode of EPSP and PH011669 in *PbCS*. Only the main contributors to the binding of EPSP and PH011669 are depicted for clarity. Solid and dashed blue arrows indicate strong (interaction distance:  $<3.2$  Å) and weak (interaction distance:  $3.3$ – $4.0$  Å) interactions, respectively. (D) Superimposition of the best docking pose of PH011669 on the EPSP pose found in the crystal structure 1QXO. The enzyme is shown as a green cartoon, with interacting amino acids depicted as sticks. PH011669 and EPSP are illustrated as magenta and cyan sticks, respectively. Black arrows indicate interactions between the ligands and *PbCS*.

#### ADME properties of PH011669 and optimisation strategies

To evaluate the propensity of PH011669 as a drug, we analysed the ADME properties of the inhibitor (Table 4). As a large

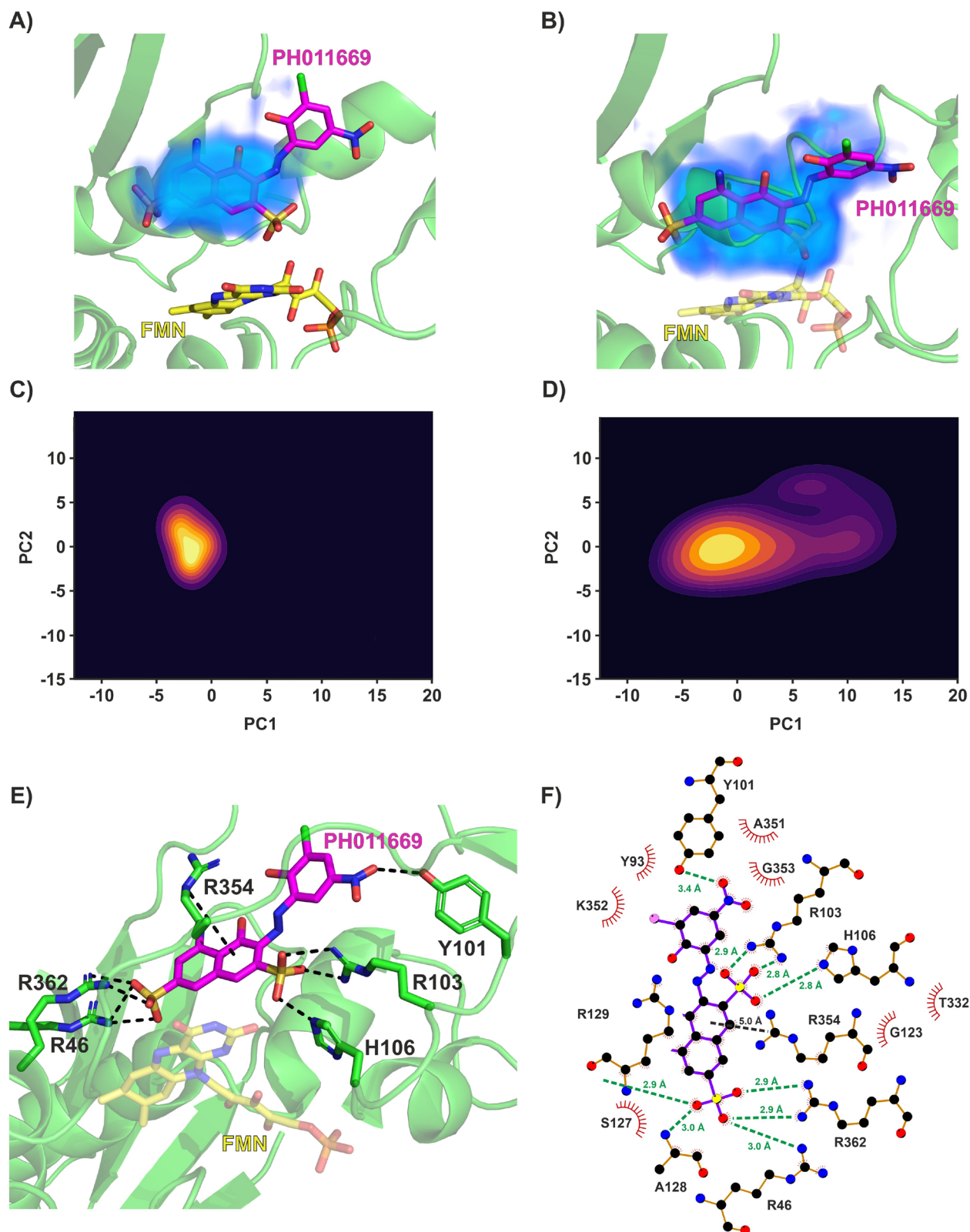
molecule decorated with many functional groups, PH011669 violates two of the four Lipinski's rule of 5, as it has a molecular mass of  $> 500$  Da (518.9 Da) and  $>10$  H-bond acceptors<sup>31</sup>. Furthermore, PH011669 exhibits a large topological polar surface



**Figure 8.** RMSD and RMSF plots for all systems. (A) RMSD plots for the Ca atoms of *PbCS* in all systems. (B) RMSD plots for the atoms of PH011669 in the binary and ternary complex. (C) RMSF plots for all systems. Residue regions exhibiting pronounced flexibility are annotated and marked by coloured boxes. The *PbCS* model is depicted as a green cartoon with the flexible regions annotated and coloured according to the RMSF profile to visualise where those regions are located in the structure. The dashed red circle highlights the active site region. RMSD and RMSF values were averaged over three trajectories for the apo and ternary systems and over nine trajectories for the binary system.

area of  $262.5 \text{ \AA}^2$ , which most likely will reduce its bioavailability significantly. The presence of an azo-group is also undesirable because of the instability towards reduction, releasing potentially harmful aromatic amines<sup>32</sup>. Therefore, PH011669 must be optimised to achieve an acceptable pharmacokinetic profile. Our *in-silico* results indicate that PH011669 binding mainly relies on ionic interactions of its 5-amino-4-hydroxynaphthalene-2,7-disulfonate core, while the 3-chloro-2-hydroxy-5-nitrophenyl moiety contributes only marginally to the binding. In addition, the binding contributions of the 5-amino and 4-hydroxy groups seem also negligible. Thus, removing these groups is expected

to hardly affect the inhibitor's binding but improve its pharmacokinetic profile significantly. To also enhance the binding affinity of PH011669, we propose adding a negatively charged group, such as a carboxy group, instead of the 4-hydroxy group to enable further salt bridge formation with R129 and R354, as observed for EPSP binding in *SpCS* (Figure 7C,D). Applying these modifications would lead to a pharmacologically more suitable inhibitor, as evidenced by a lower molecular mass of 332.3 Da, the absence of the azo group and a lower topological polar surface area of  $162.8 \text{ \AA}^2$ , which is still considered high but very close to the upper limit of the Muegge filter<sup>33</sup> (Table 4).



**Figure 9.** Post-molecular dynamics analysis of the *PbCS*-FMN-PH011669 and *PbCS*-PH011669 systems. Sampling of inhibitor conformations from the MD trajectories of the (A) ternary and (B) binary systems. Sampled conformations are shown as residence densities for PH011669 (blue volumes). For better comparison, FMN (transparent sticks) is also depicted in the FMN-lacking binary system (B). Note the differences in the size and shape of the density between the binary (B) and ternary (A) complex. Due to its high flexibility, the 3-chloro-2-hydroxy-5-nitrophenyl ring did not produce sufficient data points to be visualised as a density, so the shown densities mainly originate from the naphthalene disulfonate core. (C) Principal component analysis (PCA) of PH011669 in the ternary complex. (D) PCA of PH011669 in the binary complex. PCAs are depicted as heatmaps showing the motion of the inhibitor. The heatmaps are coloured by PH011669 density (yellow=highly populated inhibitor poses; purple=less populated inhibitor poses). (E) Representative PH011669 pose within the active site of *PbCS*. The representative structure was taken from a frame of a trajectory located in the centre of the most populated cluster in the PCA plot (D). The enzyme is illustrated as a transparent green cartoon, with interacting residues shown as sticks. Black dashed lines indicate electrostatic and polar interactions. The interaction between R354 and the inhibitor is a cation- $\pi$  interaction. (F) Protein-ligand interaction plot (LigPlot) of the representative PH011669 pose. Green dashes indicate electrostatic and polar interactions while black dashes indicate cation- $\pi$  interactions; bond distances are also shown.

## Conclusions

The shikimate pathway is essential for prokaryotes, protozoa, fungi and plants but is absent in humans, rendering the enzymes of this pathway, such as CS, attractive targets for developing antifungal or antimicrobial agents. In this study, we screened commercially available azo-dyes for CS inhibitors and identified the novel lead compound PH011669 for the development of highly potent therapeutic agents targeting the shikimate pathway. Molecular docking and dynamics simulations predicted that the azo-dye binds to the active site of *PbCS* primarily via strong ionic interactions between its negatively charged sulphonate

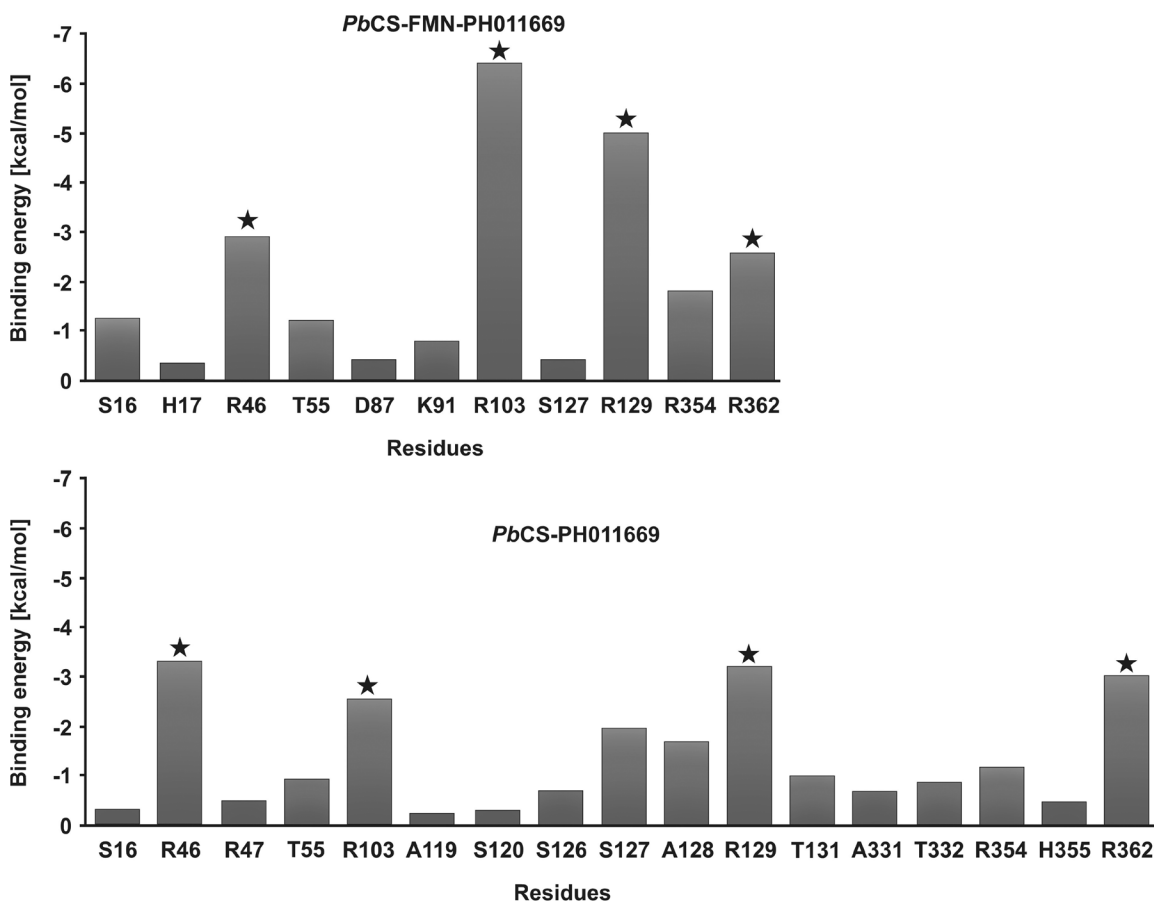
groups and three arginine residues (R46, 103 and R362). These interactions essentially mimic those observed in the crystal structure of *SpCS* in complex with its natural substrate EPSP. MD simulations in the presence and absence of the FMN cofactor demonstrated that the interacting arginine residues R46, R103 and R362 are arranged in such a way within the binding site so that they can orchestrate the binding of negatively charged ligands directly above the cofactor. The calculated binding poses showed that the naphthalene-2,7-disulfonate core of PH011669 is mainly responsible for its binding to CS, suggesting that the remaining functional groups can be removed without considerably affecting its binding affinity. Such modifications will be necessary to obtain a pharmacologically acceptable inhibitor that can be considered for medicinal purposes, as the current structure of PH011669 has a poor pharmacokinetic profile. Furthermore, our computational results provide valuable guidance for future inhibitor designs, as they show which hot spots in the active site of CS need to be targeted (e.g. R46, R103 and R362). For example, our results suggest that adding a further negative charge at the 4-position of the naphthalene ring of PH011669 could produce a compound with higher affinity to CS, as additional electrostatic interactions will be established.

Additional studies need to be conducted to optimise the present lead compound, especially in terms of improving its drug-likeness, although our data can support future research toward antifungal or antimicrobial therapeutics that target the shikimate pathway via CS inhibition.

**Table 3.** Gibbs free energies calculated by MM/PBSA.

	Energies [kcal/mol]	
	<i>PbCS</i> -FMN-PH011669	<i>PbCS</i> -PH011669
$\Delta E_{VDW}$	$-55.35 \pm 0.07$	$-47.11 \pm 0.06$
$\Delta E_{EL}$	$-123.52 \pm 0.41$	$-154.79 \pm 0.37$
$\Delta E_{POL}$	$135.78 \pm 0.34$	$206.78 \pm 0.33$
$\Delta E_{NONPOL}$	$-7.05 \pm 0.01$	$-6.48 \pm 0.01$
$\Delta G_{GAS}$	$-178.86 \pm 0.40$	$-239.90 \pm 0.37$
$\Delta G_{SOLV}$	$128.73 \pm 0.34$	$200.3 \pm 0.32$
$\Delta G_{binding}$	$-50.20 \pm 0.11$	$-39.59 \pm 0.10$

$\Delta E_{VDW}$ : van der Waals energy;  $\Delta E_{EL}$ : electrostatic contribution to the solvation energy;  $\Delta E_{POL}$ : polar component of the solvation energy;  $\Delta E_{NONPOL}$ : non-polar component of the solvation energy;  $\Delta G_{GAS}$ : total gas phase free energy;  $\Delta G_{SOLV}$ : total solvation free energy;  $\Delta G_{binding}$ : binding free energy. The values are averages over all trajectories (three for the ternary system and nine for the binary system).



**Figure 10.** Per-residue Gibbs binding free energy decomposition analysis of the *PbCS*-FMN-PH011669 (top panel) and *PbCS*-PH011669 system (bottom panel). Stars highlight the main contributors to PH011669 binding. Values are averaged over all trajectories (three for the ternary system and nine for the binary system).



**Table 4.** Lipinski's rule of 5 violations and ADME properties of PH011669 and its optimised version.

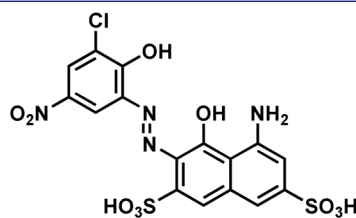
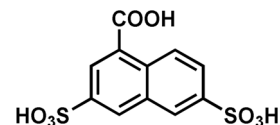
	PH011669	PH011669 optimized
Lipinski's rule of 5		
150 Da ≤ MW ≤ 500 Da	518.9 Da	332.3 Da
-0.7 ≤ Log P ≤ 5	2.06	0.19
≤ 10 H-bond acceptors	12	8
≤ 5 H-bond donors	5	3
Other pharmacokinetic properties		
20 Å <sup>2</sup> ≤ TPSA ≤ 150 Å <sup>2</sup>	262.5 Å <sup>2</sup>	162.8 Å <sup>2</sup>
-6 ≤ Log S ≤ 0	-4.38	-2.17
No. of rotatable bonds ≤ 7 <sup>a</sup>	5	3
Presence of azo group	Yes	No
Gastrointestinal absorption <sup>b</sup>	Low	Low
Interactions with CYP isoenzymes <sup>c</sup>	No	No

MW: molecular weight; TPSA: topological polar surface area; CYP: cytochromes P450; BBB: blood-brain barrier.

<sup>a</sup>Drugs with more than seven rotatable bonds are flexible and can lead to unfavourable ADME properties, as the increased conformational entropy can distort target binding.

<sup>b</sup>For a drug to be effective, its gastrointestinal absorption should be high to enter the bloodstream.

<sup>c</sup>Assessing the inhibitory interactions of a drug with the CYP isoenzymes CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4 is essential because inhibition of these isoenzymes can lead to drug-drug interactions, increasing the risk of adverse effects. In this category, "no" means that the drug is not an inhibitor of any CYP isoenzyme.

**PH011669****PH011669 optimized**

## Acknowledgements

We would like to thank Rosemarie Trenker-El-Toukhy for her expert technical support.

## Author contributions

**Katharina Fuchs:** Conceptualisation, Formal Analysis, Investigation, Visualisation, Writing – original draft, Writing – review & editing. **Massimo G. Totaro:** Conceptualisation, Formal Analysis, Investigation, Data curation, Resources, Methodology, Software, Validation, Visualisation, Writing – review & editing. **Marina Toplak:** Formal Analysis, Investigation. **Aleksandar Bijelic:** Conceptualisation, Data curation, Formal Analysis, Investigation, Methodology, Supervision, Validation, Visualisation, Writing – original draft, Writing – review & editing. **Peter Macheroux:** Conceptualisation, Funding acquisition, Methodology, Project administration, Resources, Visualisation, Writing – original draft, Writing – review & editing.

## Disclosure statement

The authors report no conflicts of interest.

## Funding

This research was funded in whole or in part by the Austrian Science Fund (FWF) (Doctoral program "Molecular Enzymology" W908 to P.M.). For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## Data availability statement

The datasets (MD trajectories) presented in this study have been deposited in the repository at TU Graz (<https://repository.tugraz.at/records/p37ee-ph630>).

## References

- Macheroux P, Schmid J, Amrhein N, Schaller A. A unique reaction in a common pathway: mechanism and function of chorismate synthase in the shikimate pathway. *Planta*. 1999;207(3):325–334.
- Lauhon CT, Bartlett PA. Substrate analogs as mechanistic probes for the bifunctional chorismate synthase from *Neurospora crassa*. *Biochemistry*. 1994;33(47):14100–14108.
- Macheroux P, Petersen J, Bornemann S, Lowe DJ, Thorneley RN. Binding of the oxidized, reduced, and radical flavin species to chorismate synthase. An investigation by spectrophotometry, fluorimetry, and electron paramagnetic resonance and electron nuclear double resonance spectroscopy. *Biochemistry*. 1996;35(5):1643–1652.
- Osborne A, Thorneley RN, Abell C, Bornemann S. Studies with substrate and cofactor analogues provide evidence for a radical mechanism in the chorismate synthase reaction. *J Biol Chem*. 2000;275(46):35825–35830.
- Dmitrenko O, Wood HB, Bach RD, Ganem B. A theoretical study of the chorismate synthase reaction. *Org Lett*. 2001;3(26):4137–4140.
- Maclean J, Ali S. The structure of chorismate synthase reveals a novel flavin binding site fundamental to a unique chemical reaction. *Structure*. 2003;11(12):1499–1511.
- Dias MVB, Ely F, Palma MS, Azevedo WF, de Basso LA, Santos DS. Chorismate synthase: an attractive target for drug development against orphan diseases. *Curr Drug Targets*. 2007;8(3):437–444.

8. Derrer B, Macheroux P, Kappes B. The shikimate pathway in apicomplexan parasites: implications for drug development. *Front Biosci.* 2013;18(3):944–969.
9. Balasubramanian S, Davies GM, Coggins JR, Abell C. Inhibition of chorismate synthase by (6R)- and (6S)-6-fluoro-5-enolpyruvylshikimate 3-phosphate. *J Am Chem Soc.* 1991;113(23):8945–8946.
10. Bornemann S, Ramjee MK, Balasubramanian S, Abell C, Coggins JR, Lowe DJ, Thorneley RN. *Escherichia coli* chorismate synthase catalyzes the conversion of (6S)-6-fluoro-5-enolpyruvylshikimate-3-phosphate to 6-fluorochorismate. Implications for the enzyme mechanism and the antimicrobial action of (6S)-6-fluoroshikimate. *J Biol Chem.* 1995;270(39):22811–22815.
11. Davies GM, Barrett-Bee KJ, Jude DA, Lehan M, Nichols WW, Pinder PE, Thain JL, Watkins WJ, Wilson RG. (6S)-6-fluoroshikimic acid, an antibacterial agent acting on the aromatic biosynthetic pathway. *Antimicrob Agents Chemother.* 1994;38(2):403–406.
12. Thomas MG, Lawson C, Allanson NM, Leslie BW, Bottomley JR, McBride A, Olusanya OA. A series of 2(Z)-2-benzylidene-6,7-dihydroxybenzofuran-32H-ones as inhibitors of chorismate synthase. *Bioorg Med Chem Lett.* 2003;13(3):423–426.
13. Rodrigues-Vendramini FAV, Marschalk C, Toplak M, Macheroux P, Bonfim-Mendonça PdS, Svidzinski TIE, Seixas FAV, Kioshima ES. Promising new antifungal treatment targeting chorismate synthase from *Paracoccidioides brasiliensis*. *Antimicrob Agents Chemother.* 2019;63(1):e01097-18.
14. Martinez R. Epidemiology of paracoccidioidomycosis. *Rev Inst Med Trop Sao Paulo.* 2015;57(Suppl 19):11–20.
15. Bocca AL, Amaral AC, Teixeira MM, Sato PK, Shikanai-Yasuda MA, Soares Felipe MS. Paracoccidioidomycosis: eco-epidemiology, taxonomy and clinical and therapeutic issues. *Future Microbiol.* 2013;8(9):1177–1191.
16. Dias MVB, Borges JC, Ely F, Pereira JH, Canduri F, Ramos CHI, Frazzon J, Palma MS, Basso LA, Santos DS, et al. Structure of chorismate synthase from *Mycobacterium tuberculosis*. *J Struct Biol.* 2006;154(2):130–143.
17. Bueno PS, Rodrigues-Vendramini FA, Toplak M, Macheroux P, Kioshima ES, Seixas FA. New inhibitors of chorismate synthase present antifungal activity against *Paracoccidioides brasiliensis*. *Future Microbiol.* 2019;14(11):969–980.
18. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods.* 2022;19(6):679–682.
19. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol.* 2017;35(11):1026–1028.
20. Lovell SC, Davis IW, Arendall WB, Bakker PIW, de Word JM, Pisant MG, Richardson JS, Richardson DC. Structure validation by C $\alpha$  geometry: phi, psi and C $\beta$  deviation. *Proteins.* 2003;50(3):437–450.
21. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010;31(2):455–461.
22. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem.* 2009;30(16):2785–2791.
23. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015;1–2:19–25.
24. Kerschbaumer B, Totaro MG, Friess M, Breinbauer R, Bijelic A, Macheroux P. Loop 6 and the  $\beta$ -hairpin flap are structural hotspots that determine cofactor specificity in the FMN-dependent family of ene-reductases. *Febs J.* 2024;291(7):1560–1574.
25. Gowers R, Linke M, Barnoud J, Reddy T, Melo M, Seyler S, Domański J, Dotson D, Buchoux S, Kenney I, et al. MDAnalysis: A Python package for the rapid analysis of molecular dynamics simulations. In Proceedings of the 15th Python in Science Conference. July 11–July 17 2016; Austin, Texas. *SciPy*; 2016. p. 98–105
26. Michaud-Agrawal N, Denning EJ, Woolf TB, Beckstein O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J Comput Chem.* 2011;32(10):2319–2327.
27. DeLano WL. The PyMOL molecular graphics system. New York: Schrödinger LLC;2002.
28. Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA, Moreno E. gmx\_MMPPBSA: A new tool to perform end-state free energy calculations with GROMACS. *J Chem Theory Comput.* 2021;17(10):6281–6291.
29. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep.* 2017;7(1):42717.
30. Kitzing K, Auweter S, Amrhein N, Macheroux P. Mechanism of chorismate synthase. Role of the two invariant histidine residues in the active site. *J Biol Chem.* 2004;279(10):9451–9461.
31. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods.* 2000;44(1):235–249.
32. Gičević A, Hindija L, Karačić A. Toxicity of azo dyes in pharmaceutical industry. In: Badnjević A, Škrbić R, Gurbeta Pokvić L, editors. *CMBEBIH 2019*. Vol. 73. Cham: Springer International Publishing; 2020. p. 581–587. *IFMBE Proceedings*.
33. Muegge I, Heald SL, Brittelli D. Simple selection criteria for drug-like chemical matter. *J Med Chem.* 2001;44(12):1841–1846.