Scientific Research Publishing

International Journal of Organic Chemistry, 2026, 16(3), 79-107
https://www.scirp.org/journal/ijoc
ISSN Online: 2161-4695
ISSN Print: 2161-4687

Rational Design and Computational Analysis of Covalent Benzofuran Fluorophosphonate Inhibitors in a Heme-Depleted Model of CYP1A1

Filip Latkovic[1*], Fabio Di Ricco[2]

[1]GCAS College Dublin, Dublin, Ireland
[2]Department of Pharmacy, University of Pisa, Pisa, Italy
Email: *filip.latkovic@gcas.ie, fabio.diricco@phd.unipi.it

How to cite this paper: Latkovic, F. and Di Ricco, F. (2026) Rational Design and Computational Analysis of Covalent Benzofuran Fluorophosphonate Inhibitors in a Heme-Depleted Model of CYP1A1. *International Journal of Organic Chemistry*, **16**, 79-107.
https://doi.org/10.4236/ijoc.2026.163005

Received: June 8, 2026
Accepted: July 26, 2026
Published: July 29, 2026

Abstract

The cytochrome P450 isoform, CYP1A1, occupies a paradoxical role in cancer biology, functioning both as a metabolic gatekeeper of xenobiotics and as a catalyst in the bioactivation of procarcinogens. While conventional inhibitor design has largely targeted the catalytically-active, heme-bound (*holo*-) form of CYP1A1, comparatively little attention has been paid to its dynamic equilibrium with its heme-free states (heme-depleted *holo*- and *apo*-) under conditions of oxidative stress and altered heme homeostasis: hallmarks of the neoplastic microenvironment. In this study, we explore the hypothesis that transient heme depletion (ligand exchange) may expose a therapeutically exploitable window for covalent inhibition. To this end, a series of six benzofuran fluorophosphonate derivatives (BMFP-C1 - C6) were rationally designed as hypothetical electrophilic inhibitors capable of engaging nucleophilic residues within the active site of heme-depleted CYP1A1. A multi-tiered computational workflow was employed, comprising molecular docking (AutoDock Vina), pseudo-covalent docking (rDock), molecular dynamics (GROMACS), binding affinity prediction (K_{DEEP}), geometric analysis of nucleophilic attack (distance and angles), and normal mode analysis (NMA). Across these analyses, ligand BMFP-C1 consistently exhibited a favorable convergence of properties, including strong predicted binding affinity, rapid attainment of ligand-protein equilibrium, and a geometrically-permissive orientation for nucleophilic substitution at phosphorus. Notably, structural inspection of the BMFP-C1-CYP1A1 complex revealed the proximal arrangement of Asp320, Lys499, and Thr497, suggestive of a pseudo-"catalytic triad" capable of facilitating a two-step nucleophilic substitution (S_N2_P) mechanism. While this mechanistic pro-

posal remains speculative, the alignment of geometric and energetic indicators supports the plausibility of covalent engagement mediated by transient, heme-depleted *holo*-CYP1A1 accessibility. Taken together, these findings suggest that targeting the heme-depleted, pre-*apo* conformational landscape of CYP1A1 may represent a viable, albeit underexplored, strategy for hypothetical covalent inhibitor design. The present work thus serves as a computationally-grounded, hypothesis-generating framework for the future synthesis and biochemical evaluation of fluorophosphonate-based CYP1A1 inhibitors.

Keywords

Antineoplastics, Covalent Inhibitors, Cytochrome P450 1A1 (CYP1A1), Fluorophosphonates, Normal Mode Analysis (NMA)

1. Introduction

The Cytochrome P450 (CYP) superfamily of metalloproteins are a related class of heme-containing enzymes responsible for up to 80% - 90% of the metabolism of drugs and xenobiotics, and span three distinct families: CYP1, CYP2, and CYP3 [1] [2]. Functioning mostly as hepatic, endoplasmic reticulum (ER) membrane-bound monooxygenases, the CYP isoforms which play the most dominant role in drug metabolism—with relative drug turnover rate measured across 248 different compounds—include: CYP3A4/5 (30%), CYP2D6 (20%), CYP2C9 (13%), CYP1A2 (9%), and CYP2B6 and 2C19 (7%, respectively) [3]. Notwithstanding the hepatic abundance of most CYPs, notable families of the enzymes are expressed extrahepatically, including: CYP1A1 (lungs, esophagus, small intestine, skin), CYP1A2 (lungs, colon), CYP2A6 (tracheal and bronchiolar epithelial cells), CYP2A13 (nasal mucosa, trachea, bronchial epithelial cells), and CYP1B1 (lungs, small intestine, colon) [1] [4].

While their endogenous functions entail the processing of complex hydrocarbons such as 17β-estradiol (2-/4-hydroxylation), arachidonic acid (ω-16 hydroxylation), and retinoids (retinol/al oxidation), the roles of the CYP1A1 and CYP1B1 isoforms from the CYP1 family have demonstrated an etiological culpability in cancers such as lung, colon, and breast cancer [5] [6]. Cardinally, their propensity to activate procarcinogens such as polycyclic aromatic hydrocarbons (PAHs), nitro-aromatic hydrocarbons, and amine/azo-amine compounds has been implicated in the origin and progression of malignant tumors [7]. In this respect, a study conducted by Androutsopoulos *et al.* [5] revealed that—compared to non-neoplastic bladder and colon cells—65% of bladder cancers and 80% of colon cancers overexpressed CYP1A1 mRNA, while 65% of bladder cancers and 60% of colon cancers overexpressed CYP1B1 mRNA, respectively. The same study also found that up to 35% of all bladder and colon cancer cell CYP1A1/1B1 activity was increased compared to non-neoplastic tissue [5]. These findings are, moreover, supported by research conducted by Carrera *et al.* [8] and Kumarakulasing-

ham *et al.* [9], respectively, which found a significant risk of lymph node metastasis in CYP1B1-overexpressing tumor cells. Similarly, while necessary to activate certain chemotherapy prodrugs, the same CYPs were found to be responsible for the inactivation of cancer-retarding chemotherapy agents [10]-[12].

In particular, the CYP1A1 isoform, known more prominently than CYP1B1 in its role in xenobiotic-based procarcinogen activation (*i.e.* benzo[a]pyrene conversion to the carcinogenic metabolite, benzo[a]pyrene-diol-epoxide, as seen in **Figure 1**), remains a compelling object of study primarily due to its largely undetectable expression in healthy, non-carcinogenic tissue [13]-[15].

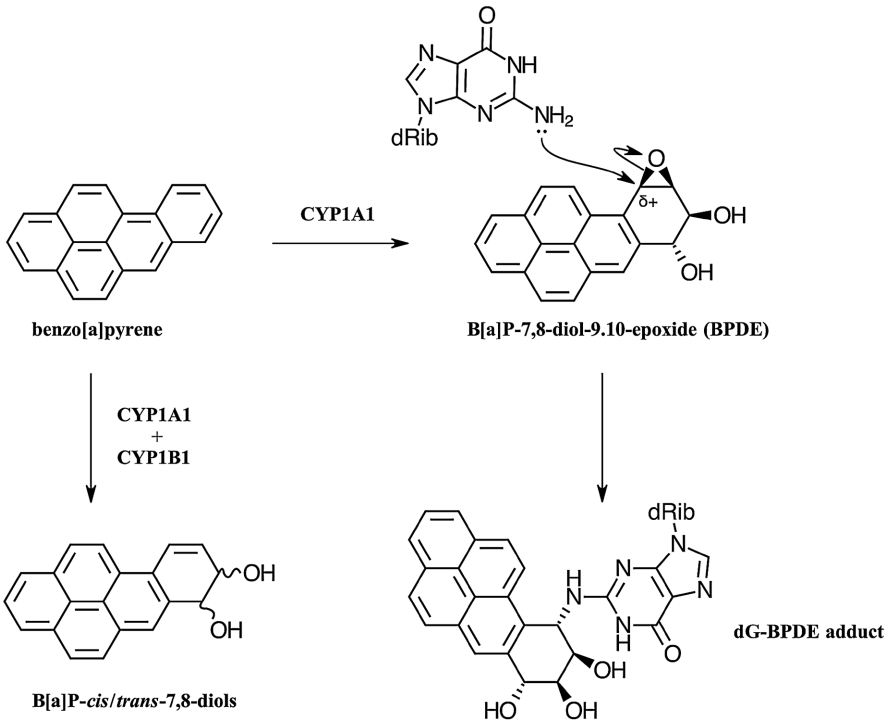

Figure 1. The respective metabolic pathways of benzo[a]pyrene by CYP1A1 and CYP1B1, adapted from Uppstad *et al.* [15]. Note the mutagenic mechanism of action (MoA) of B[a]P-7,8-diol-9,10-epoxide (BPDE) in an alkylation reaction with DNA base deoxyguanosine (dG) through the formation of an *N*2-adduct, as noted by Costa [16].

CYP1A1 has hence also been noted for the positive feedback loop by which it is overexpressed, primarily via the activation of the aryl hydrocarbon receptor (AhR) stimulated by polycyclic aromatic hydrocarbons (PAHs), activating the aforementioned procarcinogens into carcinogens in lung and liver tissue [17]. In contrast to CYP1A1 expression induced primarily via the AhR, and thus levels remaining largely undetectable in non-AhR-stimulated tissue, CYP1B1 is expressed at detectable levels in non-carcinogenic cells [13]. As such, it may be argued that CYP1A1, as opposed to CYP1B1, may be a more defining marker for tumorigenesis—particularly those of suspected xenotoxic, polycyclic aromatic hydrocarbon (PAH) etiology—in tissues such as breast, lung, and colon, due to its lower absolute abundance in normal tissues compared to CYP1B1 [14] [15]. Nev-

ertheless, Liu *et al.* [13] have demonstrated that CYP1A1-mediated resveratrol metabolism can induce apoptosis in medulloblastoma cells, thus playing an anti-tumorigenic role, likewise; a trait absent in CYP1B1.

Considering the relationality between CYP1A1 and tumorigenesis, tumor progression has been noted to be stimulated by CYP1A1-dependent suppression of AMP-activated protein kinase (AMPK) signaling via repression of AMPK phosphorylation, which consequently decreases the expression of gene-silencing CYP1A1 siRNA [14]. Such a suppression of AMPK signaling, Rodriguez & Potter [14] deduce, contributes to CYP1A1's other pro-neoplastic activities, such as: 1) the overexpression of cyclin D_1, 2) the overexpression of survivin, and 3) the dis-inhibition of the ERK1/2 & PI3K/AKT pathways.

One of the, arguably, most intriguing findings about the CYP superfamily in general is its propensity to equilibrate its heme prosthetic group with the surrounding heme exchange pool (HEP), with a notable reduction of heme prosthesis when cellular HEP levels are low [18]. The CYPs readily equilibrate between an inactive heme-free, *apo*-form, and a heme-containing, active, *holo*-form, by establishing a dynamic equilibrium with cellular free heme reserves [18]-[20]. As Correia *et al.* [18] explain, excess free heme is prone to creating cytotoxicity via reactive oxygen species (ROS) generation, and—while bound to cellular proteins to mitigate cytosolic toxicity—requires an upregulation of heme oxygenase-1 (HO-1) to break down said excess heme, which further promotes an equilibrium shift to the *apo*-form of CYP. Strongly acidic conditions (pH 2.5 - 3.0), leading to a protonation of the cysteine (Cys) thiolate ($CysS^- \rightarrow CysSH$) coordinating the central iron cation of the heme in CYPs, and strongly-oxidative conditions (induced by H_2O_2 addition *in vitro*) are considered responsible for favoring the dissociation of heme from *holo*-CYP, forming *apo*-CYP [20]-[23]. This is theorized to be due to the reduced nucleophilicity, and thus coordinating capability, of the thiol (RSH) form of Cys, the oxidation of the Cys sulfhydryl moiety to create cross-linked protein aggregates which prevents correct heme binding, and even ferric (Fe^{III})-ferrous (Fe^{II}) oxidation state fluctuations of the central iron cation [22] [23]. The aforementioned conditions have, moreover, been found to trigger an often-irreversible conformational change in CYP into an inactive variant, Cytochrome P420, particularly if a histidine near the heme replaces the protonated Cys thiol in a process known as "ligand switching", leading to a stable P420 protein conformation [24]. Tetreau *et al.* [25] have, moreover, discovered that a transient state exists between the heme-depleted *holo*- conformation of CYP1A1 and the *apo*- conformation, with a lifetime of less than 100 ns, offering an exploitable window for druggability as a transient "heme-depleted *holo*-" enzyme.

Interestingly, this pH/oxidation-dependent switching from *holo*- to *apo*-CYP relates to neoplastic tissues such that, while cancer cells tend to maintain a more alkaline intracellular pH (pH_i, 7.3 - 7.6, compared to 7.0 - 7.4 in normal cells) and a more acidic extracellular pH (pH_e, 6.8 - 7.0, compared to 7.2 - 7.4 in normal cells), oxidative stress has consistently been shown to be higher in neoplastic cells

than normal cells [26]-[29]. This occurs due to the increased levels of mitochondrial-mediated metabolic processes, such as the tricarboxylic acid cycle (TCA) and oxidative phosphorylation, which—even at basal levels—generate superoxide ($O_2^{\cdot-}$) that is consequently transformed into H_2O_2 by superoxide dismutases (SODs) [26]. While oxidative stress, caused by increased reactive oxygen species (ROS), typically causes apoptosis, the persistent increase in ROS levels in cancer cells, such as those of H_2O_2 and superoxide—coupled with their roles as a cellular signaling molecules in phosphatase/kinase pathways—leads to a dependence of cancer survival and proliferation on chronically-increased ROS levels in many neoplastic cells (*i.e.* breast and skin cancer) [26]. Moreover, recent research by Brandl *et al.* [30] has demonstrated the correlative effects between ROS generation and pH homeostasis in cancer cells, chiefly through hypoxia-inducible factor 1*α* (HIP-1*α*) signalling.

While research into the trivariate correlation between how the oxidative stress environment of cancer cells and their pH_i-pH_e gradient influences CYP heme exchange is lacking, McFadyen *et al.* [31] noted the influence of the hypoxic, ROS-rich environment on CYP metabolism in selective prodrug activation of AQ4N by CYP1A1, CYP3A4, and CYP1B1 into the anticancer drug AQN. Moreover, HO-1 levels have consistently been found to be elevated in cancer cells compared to non-neoplastic cells due primarily to the increase in ROS, with the concomitant increase in HO-1 serving a cytoprotective role against elevated ROS levels, ultimately contributing to cancer cell survival and proliferation [32]. On the other hand, it was discovered that an increase in heme degradation brought on by elevated HO-1 levels in certain cancers (*i.e.* prostate, colorectal, and neuroblastoma) can induce iron-mediated ferroptosis [32].

As such, we hypothesized that devising competitive inhibitors that would occupy the active site of CYP1A1 during heme exchange could serve as a plausible direction for future anticancer research, considering the interplay between ROS, HO-1 levels, and the influence on *holo-apo* conformational shifts in CYP1A1. Based on this hypothesis, we devised a series of six primary benzofuran fluorophosphonates that would opportunistically engage with the heme-free CYP1A1 during heightened rates of heme exchange in HO-1-overexpressing cancer cells, covalently inhibiting the re-binding of heme to Cys457 on the enzyme's active site through steric hindrance (see **Figure 2**).

These six ligands: BMFP-C1, BMFP-C2, BMFP-C3, BMFP-C4, BMFP-C5, and BMFP-C6 were developed through a combinatorial strategy of ligand-based rational design initiated by Joshi *et al.* [17] in their efforts to create a viable CYP1A1 inhibitor for cancer chemoprevention, and refined by our team previously. Considering the narrow time window, moreover, for the heme-dissociated *holo*-form of CYPs to structurally "relax" into the *apo*-form (>10 - 20 ns at 293 K), we had to keep in mind the devising of ligand structures which would non-covalently bind and stabilize the active site of CYP1A1 within a drastically shorter timeframe, preventing structural hindrance to the ligand binding site [25]. Nevertheless, several

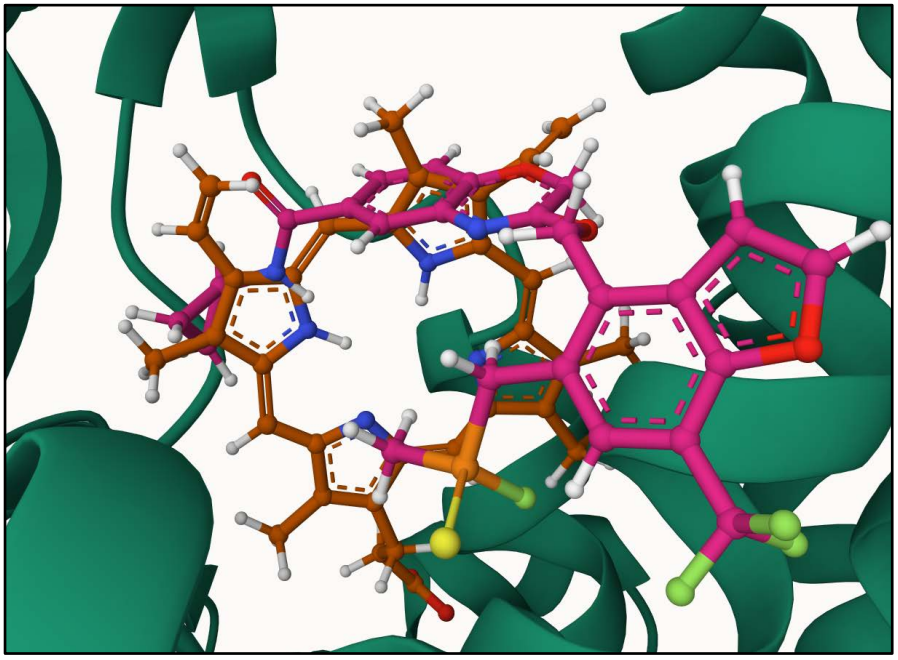

Figure 2. A superimposition depiction (PDB) of the binding pocket of the active site of *holo*-CYP1A1 (green), with the porphyrin ring of the heme (brown), and ligand BMFP-C1 (magenta), depicting steric clash/overlap following AutoDock Vina docking. *MolStar** was used to visualize and superimpose PDB files.

studies have confirmed the possibility of reinstitution of ligands even in the "relaxed", *apo*-form of CYP [19] [25].

CYP1A1 was chosen as a primary target for our paper because of its ready ability to serve as a biomarker for tumorigenesis, as well as its complex pro-neoplastic mechanism of action described previously. Furthermore, the fluorophosphonate group was chosen as the electrophilic group susceptible to nucleophilic substitution (S_N)—particularly bimolecular nucleophilic substitution at phosphorus (S_N2_P)—due to its relative hydrolytic stability at physiological pH compared to other halophosphonates, whilst still retaining high, serine/threonine-selective reactivity, thus potentially reducing dangerous off-site effects before a sufficiently-long, non-covalent interaction with the targeted CYP1A1 active site [33] [34]. Similarly, while fluorophosphonates and organophosphates are known to undergo addition-elimination substitution reactions, these tend to be much less prominent mechanistically than the S_N2_P mechanistic pathway, particularly in media with low water concentration, such as the catalytic site of CYP once bound to a substrate/ligand [35] [36].

2. Materials and Methods

2.1. Molecular Docking: AutoDock Vina and AceDock

The first phase of this study consisted of using *Mcule*'s AutoDock Vina feature to prepare the heme-depleted *holo*- and *apo*- protein crystal structures for comparative heme-free molecular docking based on a RCSB PDB template of heme-

containing *holo*-CYP1A1 in complex with *α*-naphthoquinone (PDB ID: 4I8V) [37] [38]. Six ligands (BMFP-C1, BMFP-C2, BMFP-C3, BMFP-C4, BMFP-C5, and BMFP-C6) were designed as structural derivatives of the parent compound, BML-1, previously developed by our team prior as a *holo*-CYP1A1 inhibitor (Figure 3).

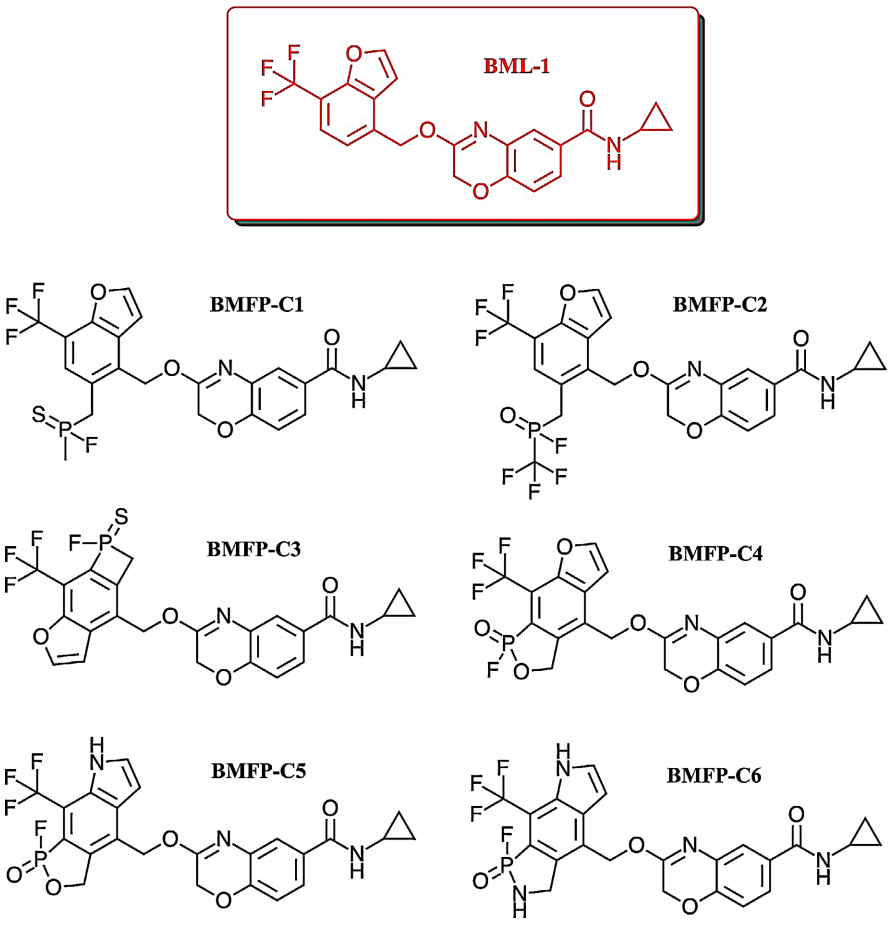

Figure 3. Structural formulae of the six ligands designed in the study (BMFP-C1, BMFP-C2, BMFP-C3, BMFP-C4, BMFP-C5, BMFP-C6), along with the parent compounds, BML-1.

The central binding region surrounding residue Phe224 (X: −17.915, Y: 34.362, Z: −25.594, exhaustiveness: 8, pH: 7.4, ligand protonation state: N/A, number of top-scoring poses retained: 4) within the active-site cavity of both heme-depleted *holo*- and *apo*-CYP1A1 was selected as the docking target because of the well-documented π–π interactions formed between polycyclic aromatic ligands and Phe224 [17] [39]. Compared to the parent compounds of the BM class (BM6, BM13-N, BM13-O, BM14-N, and BM14-O) originally designed and synthesized by our team, the (7-trifluoromethyl)benzofuran-functionalized derivatives (BML-1, BML-2, BML-3, and BML-4) of the best-docking parent, BM14-O, were found to have slightly more favorable (more negative) docking scores (Figure 4).

For preliminary docking, iron-free protoporphyrin IX (POR) was likewise

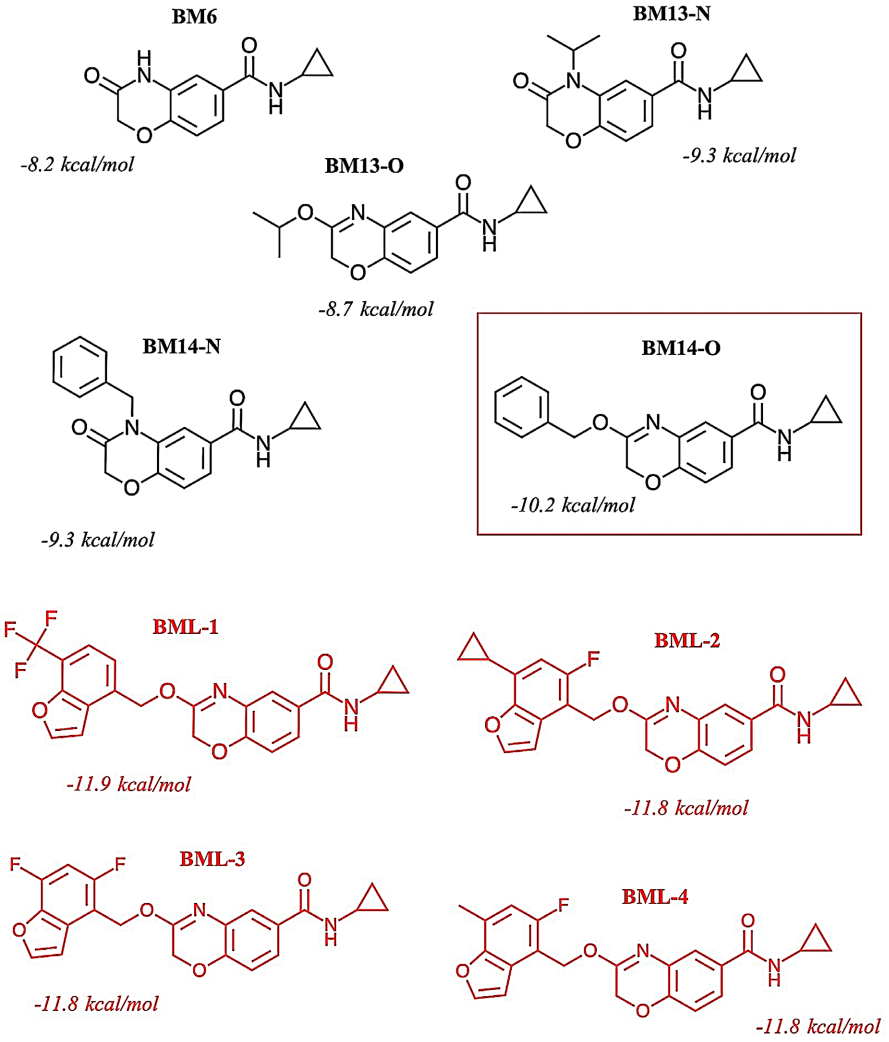

Figure 4. Compound bank and best docking scores (kcal/mol) of the original five parent compounds (BM6, BM13-N, BM13-O, BM14-N, BM14-O) and the (7-trifluoromethyl)benzofuran derivatives (BML-1, BML-2, BML-3, BML-4) of best-docking BM14-O, designed by our team previously. Of these, BML-1, having the most negative docking score of the four (−11.9 kcal/mol) was chosen for the synthesis of the BMFP fluorophosphonate class studied in this paper.

incorporated in its respective cysteine 457-centered (Cys457) binding site for comparison to the binding energies (kcal/mol) of the other six ligands due to the steric hindrance rationale provided (see Introduction). Iron-free protoporphyrin IX was chosen due to computational issues surrounding proper parameterization of chelated metal heteroatoms (*i.e.* Fe) during docking. The most favorable (most negative) absolute docking scores obtained for the *holo* and *apo* forms were subsequently compared, and the protein-ligand complexes exhibiting the more negative predicted binding energies (kcal/mol) of the two were selected for subsequent pseudo-covalent docking using the AceDock (rDock) module in *Acellera*'s "PlayMolecule AI" software [40]. This pseudo-covalent docking was carried out from a conformationally-relative docking score (identical to absolute docking

score in all but one of the ligands) PDB file that was selected to account for the most favorable (best docking score) ligand-nucleophile non-covalent alignment. As such, when speaking of "relative docking scores" in this study, this should be taken to indicate interchangeability with the term "conformationally-relative" docking scores.

2.2. Molecular Dynamics Simulation and Binding Affinity Prediction

The second phase of this study consisted of molecular dynamics (MD) simulation via *SiBioLead*'s GROMACS platform for each of the six ligands. The constants supplied for GROMACS were 310 K (corresponding to physiological temperature), 300 ns NVT/NPT equilibration time, 10 mM NaCl, 1 bar pressure, and 10 ns simulation time. 10 ns was chosen due to the need to detect ligand-protein complex stabilization within the specified timeframe due to the swift (<100 ns) relaxation of heme-depleted *holo*-CYP1A1 into its *apo*- conformation following heme depletion (see Introduction). The force field utilized during MD simulation was AMBER99SB, with a triple-point simple point charge (SPC) water model, and a dodecahedron box type used for optimal globular complex parameterization. Energy minimization (EM) during MD simulation was conducted using the "Steepest Descent" EM integrator, with the default 5000 steps setting. The simulation integrator utilized was "Leap Frog". Two result criteria were selected for further analysis and interpretation post-MD: root mean square deviation (RMSD), measuring the average structural deviation of the ligand-protein complex over time (10 ns), and root mean square fluctuation (RMSF), recording the average fluctuations (nm) of individual amino acid residues over the duration of the simulation. A single MD run was performed to obtain preliminary rankings.

"PlayMolecule AI" was subsequently used to calculate K_{DEEP} values (pK_d) via docked complexes generated through rDock, which were then manually converted into predicted dissociation constant (K_d) and free-energy (ΔG) values via the following equation:

$$\Delta G = -RT\ln(K_d)$$

where R = 8.314 $J \cdot mol^{-1} \cdot K^{-1}$, and T = 310 K. Predicted K_d values reported in Table 2 were converted from molar (M) to nanomolar (nM) units, while calculated predicted ΔG values were converted from $J \cdot mol^{-1}$ into kcal/mol, for added clarity.

2.3. Geometric Analysis: Distance and Angle of Nucleophilic Attack (AONA)

Following MD simulation, the PDB files of the six ligands with the end trajectories were subsequently analyzed using Schrödinger's PyMOL (Version 3.0). PyMOL was used to calculate the proximal phosphorus-nucleophile (P-Nu) distance (Å) for each of the six ligands, with the nucleophilic amino acid side chains considered being: serine (Ser), threonine (Thr), aspartate (Asp), glutamate (Glu), asparagine

(Asn), glutamine (Gln), and lysine (Lys). Nucleophilic precedence, due to steric factors and the pH constant selected being 7.4 to account for cancer cell pHi, was given in the following order (highest to lowest): Asp/Glu, Ser, Thr, Arg/Glu, and Lys. Potential "catalytic triad" (His/Lys/Arg-Glu/Asp-Ser/Thr) pockets that would increase the rate of S_N2_P were likewise recorded.

After identifying the proximal nucleophilic residue to the fluorophosphonate warhead, the angle of nucleophilic attack (AONA) was determined based on optimal S_N2_P geometry (140˚ - 180˚), specifically by measuring the angle (˚) between the nucleophile (Nu), phosphorus (P), and fluoride (F) leaving group [41] [42]. Then, manual covalent bonding of Nu to P was performed, with a displacement of F, via PyMOL's editing software. If the nucleophile forming the predicted P-Nu bond was of the carboxylate (*i.e.* Glu & Asp) or amide (*i.e.* Asn & Gln) class, an unstable phosphonocarbonate or *N*-phosphonamide intermediate would be formed; thus, a further geometric analysis and covalent binding would be attempted with the "second-closest" nucleophilic residue of a stable, phosphonate/phosphonamide-forming class (*i.e.* Ser, Thr, or Lys), instead.

2.4. Energy Minimization of Covalent Complexes

Following the forced covalent bonding for each of the six ligand-CYP1A1 complexes, PyMOL's Merck Molecular Force Field (MMFF94) feature was used for preliminary energy minimization, alongside transient use of the "sculpting" feature to relieve bond molecular geometry stress given the hybridization of atoms. Following said preliminary energy minimization, the exported PDB file was input into YASARA force field energy minimization software via *Yasara.org* in order to obtain a more reliable and academically-established energy-minimized conformation of the covalently-bound ligand-CYP1A1 complexes.

2.5. Normal Mode Analysis (NMA)

The final phase of this study consisted of conducting normal mode analysis (NMA) via BioSig Lab's *DynaMut* [43]. The deformation energies (0-120 units) and atomic fluctuations (0 - 3 Å) were plotted in respective line graphs via *Google Sheets*, first and foremost, for the heme-depleted, post-MD simulation *holo*-CYP1A1 to serve as a comparison for the ligand-CYP1A1 complexes. Deformation energies (0 - 120 units) for both the six non-covalent ligand-CYP1A1 complexes (post-MD simulation PDB files) and six covalent ligand-CYP1A1 complexes (residues 36 - 512) were then plotted in comparative line graphs via *Google Sheets*, with the same procedure being undertaken for the comparative plotting of atomic fluctuations (0 - 3 Å) of the covalent and non-covalent ligand-CYP1A1 complexes, respectively.

2.6. Limitations

While this study has concerned itself with the rational design of the six aforementioned hypothetical covalent inhibitors of CYP1A1, the study itself remains ex-

perimentally yet to be validated. In such a way, definitively confirming the inhibitory mechanism of action (MOA) of our ligands is currently putative at best, and based on promising inhibition bio-assays and analyses of parent compound classes synthesized by Joshi *et al.* [17] previously. Furthermore, while much of the software used was utilized in consideration of both: 1) its availability on web-based, cost-effective GUI services for our current, non-externally-funded research capacities, and 2) its reliability in academic research already conducted in the field of medicinal chemistry and chemical biology, limitations to not utilizing more expensive software, such as GLIDE, remain to be considered in the practicable utility for producing accurate computational results.

It should be noted, moreover, that certain software used—such as YASARA energy minimization—may be considered old and outdated (with consideration that *Yasara.org* is a site that has been functional, albeit with the most recent update being in 2026, since 1993). On the other hand, certain software used may be considered too novel, and thus too unexplored, for ascertaining research optimality and reliability (*i.e. PlayMolecule*). This has been exemplified there where different, inconsistent periods of K_{DEEP} calculation have yielded somewhat different results (*i.e.* pK_d ranging from 9.00 - 11.00, corresponding to 0.95 nanomolar to 1.04 nanomolar uncertainty for K_d values), likely due to calibration issues with parameterization. Thus, the "benefit of the doubt" has been given to *PlayMolecule*, and the most optimal K_{DEEP} scores (pK_d) computed for all six ligands (as well as protoporphyrin IX and BML-1) were reported in a large, continuous computation, such that relative pK_d values remained reliable for comparison and contrast. Quantitatively absolute pK_d values, on the other hand, could not be reliably accounted for in this study, and as such pK_d computation had to remain as a largely qualitative examination of covalent binding possibility utilizing the provided relative binding affinities, and could not be treated trans-comparatively to other docking conducted (*i.e.* non-covalent docking). The important ramification of this was that, during the lead optimization phase of our research, calibrations with prior ligands of the BMFP series were used prior to any K_{DEEP} scoring.

Another paramount limitation affecting this study was the inability to calculate covalent docking scores using CovDock, *CHARMM-GUI*, or *DOCKovalent* due to cost-ineffectiveness/high cost-barrier (CovDock) for this preliminary study, and the inability of the software to accurately parameterize fluorophosphonate electrophiles used (*CHARMM-GUI* and *DOCKovalent*). Hence, we had to rely on *PlayMolecule*'s rDock feature (AceDock) for obtaining "pseudo-covalent", distance-minimized poses between the nucleophilic CYP1A1 residues and electrophilic phosphorus of our ligands, further precluding any absolute docking scores from being interpretable aside from relative comparability. Similarly, covalently-bound ligand-CYP1A1 complexes could not be parameterized by AMBER99SB force field via *SiBioLead*'s molecular docking software to enable MD simulation; hence, only NMA could be used for covalent complex analysis.

3. Results and Discussion

3.1. Docking and Molecular Dynamics Simulation

Preliminary docking scores between heme-free *holo*-CYP1A1 and *apo*-CYP1A1 demonstrated a progressive increase (greater negative values) in best docking score binding energy (kcal/mol) from BMFP-C1 (−10.20) to BMFP-C6 (−12.30) with heme-depleted *holo*-CYP1A1, whereas the same trend could not be observed for the ligand-*apo*-CYP1A1 docking complexes (Table 1). Protoporphyrin IX (POR) docking scores were markedly more negative with the heme-depleted *holo*-CYP1A1 model (−10.30) than with *apo*-CYP1A1 (−8.00); a trend likewise observed with the six BMFP ligands (BMFP-C1, BMFP-C2, BMFP-C3, BMFP-C4, BMFP-C5, and BMFP-C6) and the parent compound BML-1 (Table 1).

Table 1. The best absolute non-covalent docking scores (kcal/mol) of the six ligands devised (BMFP-C1, BMFP-C2, BMFP-C3, BMFP-C4, BMFP-C5, and BMFP-C6) and the parent compounds, BML-1—inclusive of protoporphyrin IX (POR)—with *apo*-CYP1A1 and heme-depleted *holo*-CYP1A1, respectively. Docking was conducted via *Mcule*'s AutoDock Vina software [37].

Ligand	Best *holo*-CYP1A1 Docking Score (kcal/mol)	Best *apo*-CYP1A1 Docking Score (kcal/mol)
POR	−10.30	−8.00
BML-1	−11.90	−6.70
BMFP-C1	−10.20	−0.80
BMFP-C2	−10.50	−2.90
BMFP-C3	−11.20	−2.20
BMFP-C4	−12.00	−5.70
BMFP-C5	−12.10	−4.80
BMFP-C6	−12.30	−5.70

Nevertheless, BMFP-C1, BMFP-C2, and BMFP-C3 all reported less optimal (more positive) docking scores compared to BML-1, whilst BMFP-C4, BMFP-C5, and BMFP-C6 retained better (more negative) scores than the same, in the heme-depleted *holo*-CYP1A1 model, whilst none of the BMFP class ligands outcompeted BML-1 in docking preference in the *apo*- model. This suggests that BML-1 potentially has a preferential binding for the *apo*- form of CYP1A1 in contrast to the heme-depleted *holo*- form, compared to the other ligands. In this light, BML-1 is only preferentially outcompeted in docking by POR, which maintains the strongest—and thus suggestively most stable—binding computed in the *apo*-CYP1A1 model (−8.00 kcal/mol). Nonetheless, BMFP-C6 has been computed to be the strongest-binding ligand in the heme-depleted *holo*-CYP1A1 model, and the strongest-binding ligand in the computed data overall (−12.30 kcal/mol).

The above comparative findings justified further investigation into docking, particularly poses and ligand-protein complex conformations, as well as pseudo-

covalent docking (rDock) and physiologically-adapted binding energy calculations (ΔG) (Table 2).

Table 2. A comparative table of the six ligands (BMFP-C1, BMFP-C2, BMFP-C3, BMFP-C4, BMFP-C5, and BMFP-C6) across: best absolute non-covalent docking scores and best relative non-covalent docking scores (kcal/mol, computed via AutoDock Vina), best pseudo-covalent docking scores (rDock score, computed via AceDock), predicted ΔG (kcal/mol) at physiological temperature (310 K), K_{DEEP} (pK_d, computed with *PlayMolecule*), and predicted K_d (nM).

Ligand	Best Absolute Docking Score—Non-Cov. (kcal/mol)	Best Relative Docking Score—Non-Cov. (kcal/mol)	Best Docking Score—Pseudo-Cov. (rDock score*)	Predicted ΔG (kcal/mol)	K_{DEEP} (pK_d)	Predicted Dissociation Constant (K_d, nM)
BMFP-C1	−10.20	−10.20	3.03	−15.32	10.80	0.02
BMFP-C2	−10.50	−10.40	1.83	−14.11	9.95	0.11
BMPF-C3	−11.20	−11.20	14.70	−10.71	7.55	28.18
BMPF-C4	−12.00	−12.00	17.51	−13.80	9.73	0.19
BMFP-C5	−12.10	−12.10	−15.58	−14.79	10.43	0.04
BMFP-C6	−12.30	−12.30	−10.95	−14.23	10.03	0.09

*"rDock scores" are arbitrary units on *PlayMolecule* that allow for comparative analysis of pseudo-covalent docking poses. More negative values indicate favorable binding, whilst more positive values indicate unfavorable binding [40].

For reference, the calculated K_{DEEP} values computed for protoporphyrin IX (POR) and BML-1 were 7.71 and 9.41, respectively, corresponding to ΔG values of −10.94 kcal/mol (POR) and −13.35 kcal/mol (BML-1), as well as K_d values of 19.50 nM (POR) and 0.39 nM (BML-1). Adjusted for physiological conditions (310 K, pH 7.4), no longer does there remain a progressive increase in binding energy strength (as in AutoDock Vina preliminary docking) from BMFP-C1 to BMFP-C6, but instead the largest predicted binding energy (ΔG) was reported for the following three ligands: BMFP-C1, BMFP-C5, and BMFP-C6 (Table 2). BMFP-C1, peculiarly, had a calculated predicted ΔG value of −15.32 kcal/mol at 310 K, the largest value of the sample, with BMFP-C5 (−14.79 kcal/mol), and BMFP-C6 (−14.23 kcal/mol) following in second and third place, respectively. This suggests that BMFP-C1 has the largest non-covalent binding affinity of the six ligands designed for CYP1A1, corresponding to a predicted dissociation constant (K_d) of 0.02 nM (20 picomolar). This, theoretically, places the ligand in the "high affinity" (<100 nM) range compared to most compounds [44]. BMFP-C3, on the other hand, had the least binding affinity (predicted K_d = 28.18 nM), albeit still placing it in the "high affinity" range theoretically comparable to some pyrazolopyridazinone PDEδ inhibitors and cannabinoid receptor-1 (CB1) agonists verified experimentally [45] [46]. BMFP-C5 and BMFP-C6, furthermore, scored rDock pseudo-covalent docking scores of −15.58 and −10.95 units, respectively, suggesting much more favorable covalent binding compared to the other four ligands (all of which had positive values).

The above analysis and interpretation considered, MD simulation conducted on

the six ligands yielded variable RMSD and RMSF results (Figure 5 and Figure 6).

Interpreting the above, the RMSD for all ligand-CYP1A1 complexes post-MD simulation ranged from 0.075 to 0.25 nanometers (nm), indicating normal average deviation in protein structure when bound to each respective ligand. For

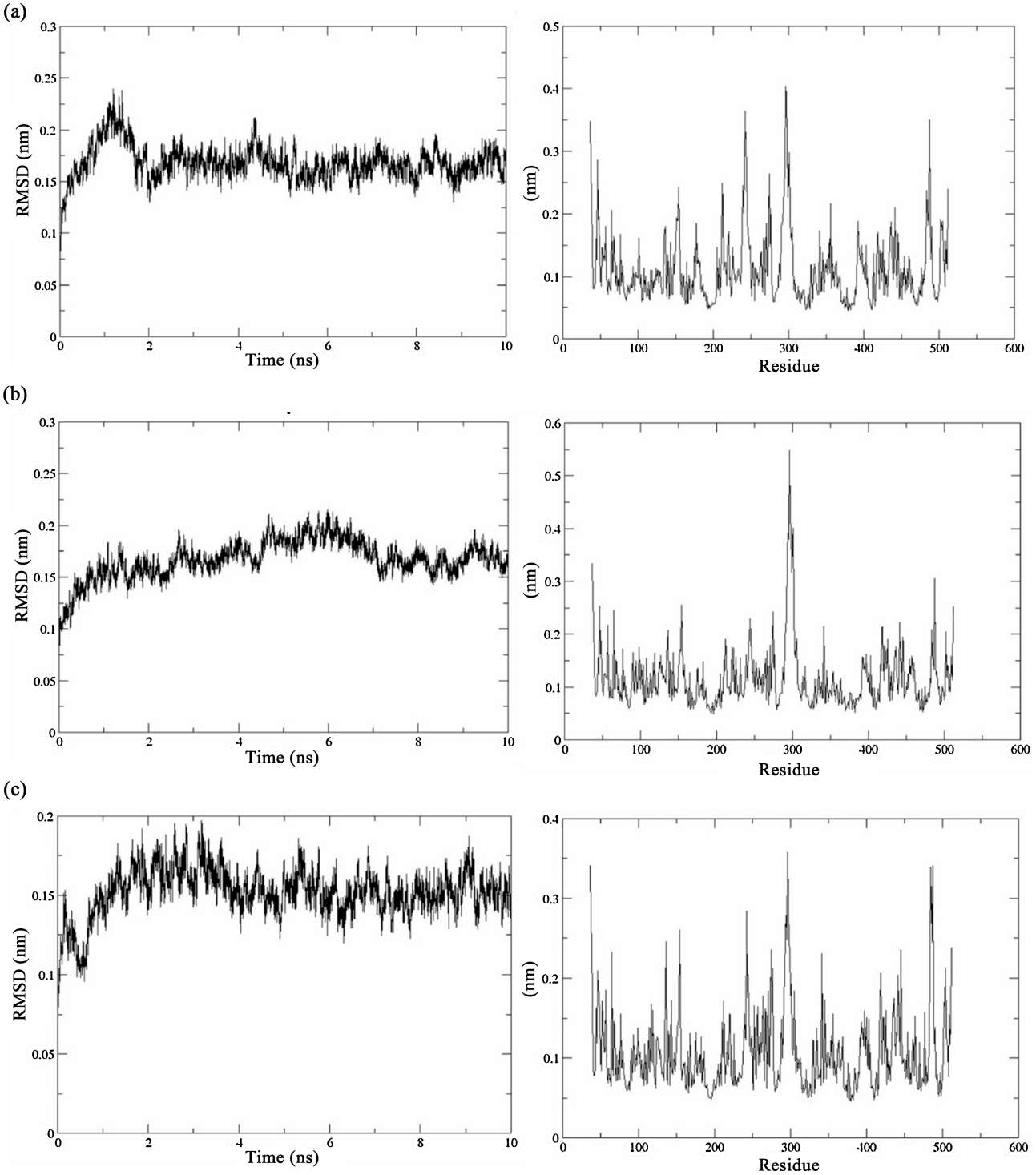

Figure 5. RMSD (left) and RMSF (right) graphs of (a) BMFP-C1, (b) BMFP-C2, and (c) BMFP-C3, computed via AMBERSS9B force field in GROMACS MD simulation (conditions: 10 mM NaCl, 310 K).

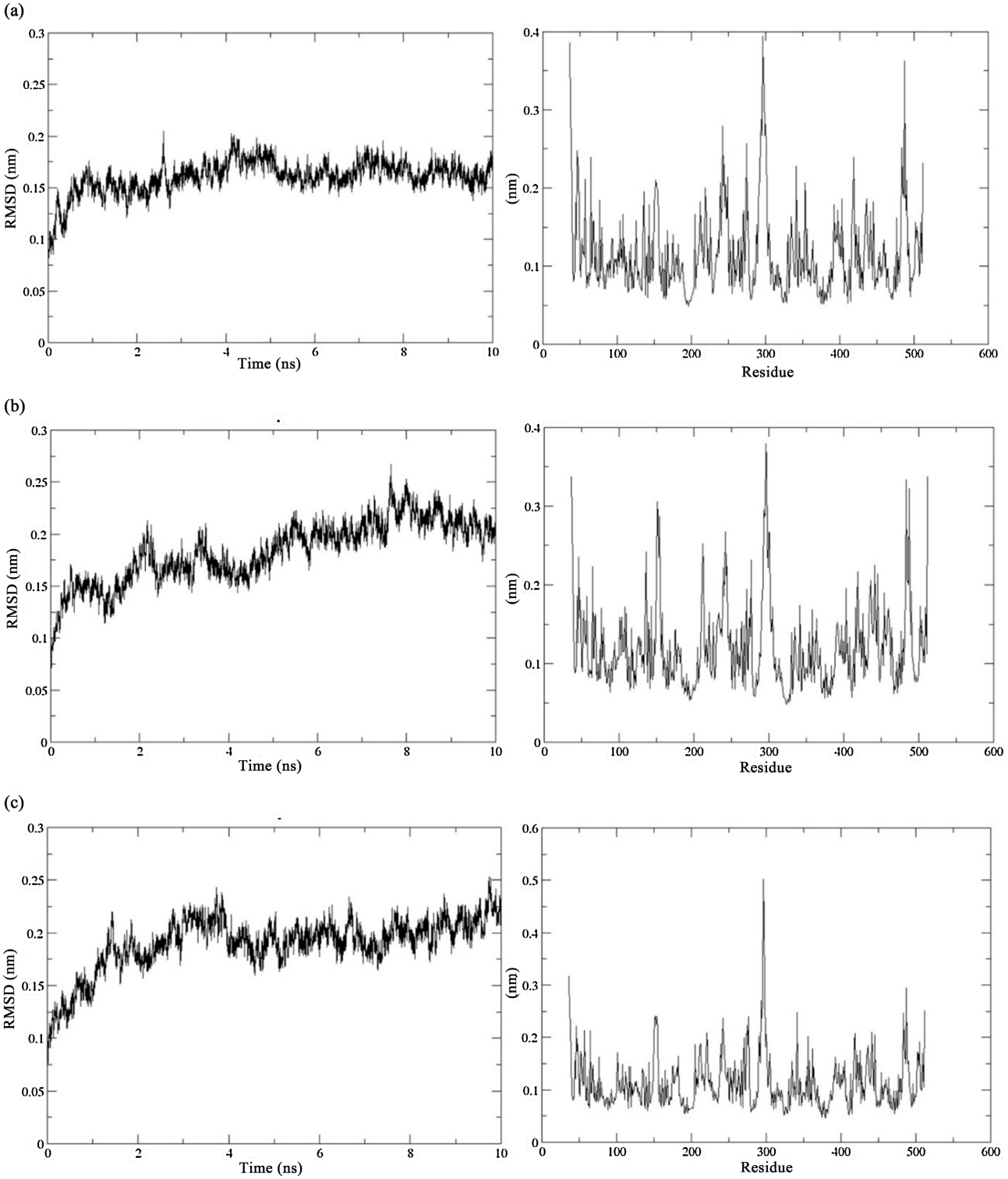

Figure 6. RMSD (left) and RMSF (right) graphs of (a) BMFP-C4, (b) BMFP-C5, and (c) BMFP-C6, computed via AMBERSS9B force field in GROMACS MD simulation (conditions: 10 mM NaCl, 310 K).

reference, values with the upper bound at 0.2 - 0.3 nm (2 - 3 Å) are considered within the "normal" range of RMSD for ligand-protein complexes [47]. The time to plateau (ns) ranges accordingly: 2.5 ns (BMFP-C1), 7 ns (BMFP-C2), 4 ns

(BMFP-C3), 5 ns (BMFP-C4), and >10 ns (BMFP-C5 & BMFP-C6). This suggests that BMFP-C1 achieves the most rapid ligand-protein equilibrium, placing it well within the lower boundary of the heme-exchange equilibration time of 10 - 20 ns [25], whilst BMFP-C5 and BMFP-C6 both take upwards of 10 ns to achieve stable ligand-protein equilibrium. Similarly, the plateaus of the ligand-CYP1A1 complexes averaged at: 0.16 nm (BMFP-C1, BMFP-C2, and BMFP-C4) and 0.15 nm (BMFP-C3), with no plateaus observed for BMFP-C5 or BMFP-C6.

Thus, it may be inferred that BMFP-C1 could pose the best computed, competitive inhibition strategy out of the six ligands in terms of ligand-protein thermodynamics across time, whereas BMFP-C3 could pose the most optimal competitive inhibition strategy out of the six ligands in terms of overall average protein deviation. Nevertheless, the difference between average protein deviation between BMFP-C1 and BMFP-C3 remains ~0.01 nm (0.1 Å), suggesting a relatively insignificant difference in terms of non-covalent/intermolecular attractions predominating in secondary and tertiary protein structure (range > 1.0 Å [48]). On the other hand, ligands BMFP-C5 and BMFP-C6 could be predicted to pose the least optimal competitive inhibition strategy of the same kind in terms of their longer time to plateau, and large average protein deviation (>0.2 nm).

With consideration to the RMSF data, on the other hand, the range of data spans from 0.05 to 0.55 nm (0.5 - 5.5 Å) across the six ligands investigated. The largest average residue fluctuations—considered uniformly across all six ligands—were centered around residue 300, corresponding to a highly flexible loop structure (Glu293 to Ser303) (**Figure 7**).

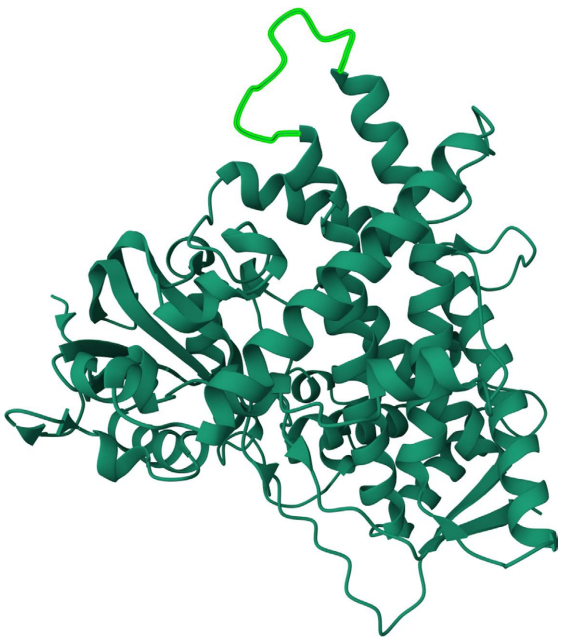

Figure 7. CYP1A1 3D structure, depicting the highly flexible loop structure (highlighted lime-green) seen across ligand-protein complexes of the six ligands, spanning eleven residues: Glu293, Leu294, Asp295, Glu296, Asn297, Ala298, Asn299, Val300, Gln301, Leu302, and Ser303.

Of the six ligand-CYP1A1 complexes, the BMFP-C2-CYP1A1 complex had the largest RMSF (residue 300), at 0.55 nm (5.5 Å), while the lowest maximum was found with BMFP-C3 (0.36 nm, 3.6 Å) at the same residue range. Overall RMSF data, on the other hand, suggested the least variation for ligands BMFP-C2 and BMFP-C6, with most RMSF values remaining below 0.2 nm (2 Å). Moreover, further high residue fluctuations were uniformly observed around residues 150 (0.23 - 0.3 nm), 240 (0.23 - 0.28 nm), and 495 (0.25 - 0.37 nm), corresponding to various flexible structural motifs depicted in Figure 8.

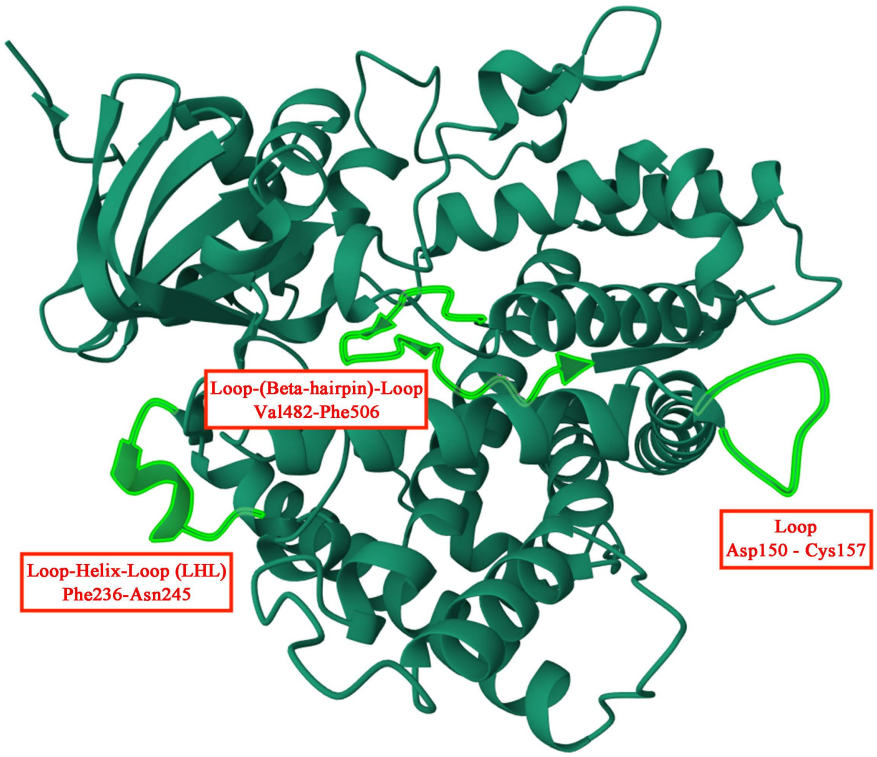

Figure 8. Labeled 3D structure of CYP1A1 and structural motifs analyzed by RMSF to be highly flexible in the ligand-protein complex post-MD simulation.

Importantly, the aforementioned flexible motifs are likely to influence the propensity for nucleophilic attack—and subsequent energy minimization—of the ligand-protein complex, which will be discussed in more detail in Section 3.2.

In terms of molecular dynamics (MD) results, particularly considering RMSD and RMSF data, it may thus be deduced that BMFP-C3 could remain the overall most favorable ligand for non-covalent binding to the active site of heme-depleted *holo*-CYP1A1 in this computational study due to its computed ability to stabilize the ligand-protein complex swiftly (4 ns), effectively (0.15 nm RMSD), and with the lowest residue fluctuation maxima (0.36 nm) over a duration of 10 ns of MD simulation. Nevertheless, the overall least-flexible—and thus more conformationally-stable—ligand-CYP1A1 complex is suggested to be that of which BMFP-C6 is the ligand. Thus, a further investigation of steric factors and the feasibility of nucleophilic attack for an efficient S_N2_P reaction between the ligand and CYP1A1

was conducted in the following section.

3.2. Geometric Analysis of Covalent Binding Feasibility, Covalent Binding, and Energy Minimization

Geometric analyses conducted on each of the six ligands in terms of proximal P-Nu distance (Å), identity of the nucleophilic residue, and angle of nucleophilic attack (AONA), yielded variable results (Table 3).

Table 3. Table depicting the proximal phosphorus-nucleophile (P-Nu) distance (Å), nucleophilic residue, and angle of nucleophilic attack (AONA) for each of the six ligands in respective ligand-CYP1A1 complexes. Data was computed via PyMOL.

Ligand	Proximal P-Nu Distance (Å)	Nucleophilic Residue	Angle of Nucleophilic Attack (AONA)
BMFP-C1	4.7	Asp320	157.6˚
BMFP-C2	3.8	Ser116	56.1˚
BMFP-C3	4.0	Ser116	90.1˚
BMFP-C4	5.4	Ser116	104.3˚
BMFP-C5	4.7	Ser116	92.8˚
BMFP-C6	8.5	Asp320	119.3˚

As per the above data, the proximal P-Nu distance ranged from 3.8 Å (BMFP-C2) to 8.5 Å (BMFP-C6), with Ser116 being the predominant closest nucleophile for all ligands except BMFP-C1 and BMFP-C6. Likewise, the range of angles of nucleophilic attack (AONA) was from 56.1˚ (BMFP-C2) to 157.6˚ (BMFP-C1), with solely BMFP-C1 being within the acceptable AONA range for S_N2_P reaction (see Section 2.3).

While the most optimal AONA in terms of S_N2_P geometry was to be found within the ligand-CYP1A1 complex of BMFP-C1 (157.6˚), the P-Nu distance (4.7 Å) is considered—viz. literature—too large for effective S_N2_P nucleophilic attack, which ranges from 1.5-4 Å (average being 3.4 Å) [42] [49] [50]. Nevertheless, Wang *et al.* [51] have predicted that—in the case of the catalytic triad-based serine residue in acetylcholinesterase (AChE)—the distance between the terminal oxygen (Oγ) and sarin phosphorus in pre-reactive S_N2_P states is >4 Å. This puts BMFP-C1 within the upper boundaries of the literature-accepted range of S_N2_P reactions in terms of its susceptibility to nucleophilic attack at its phosphorus atom.

It nonetheless remains that two ligands approach closer to nucleophilic hydroxyl residues, specifically on Ser116: BMFP-C2 (3.8 Å) and BMFP-C3 (4.0 Å), possibly favoring a more thermodynamically-stable product than the Asp320 nucleophile at the proximal distance to BMFP-C1 phosphorus: a phosphate ester, as opposed to a phosphonocarbonate intermediate. Nevertheless, for a standard, backside S_N2_P reaction, the AONA for both BMFP-C2 (56.1˚) and BMFP-C3 (90.1˚) remain unfavorable under standard physiological conditions. While

frontside S_N2_P reactions do occur, they are usually limited to polar-protic solvent systems not found within the relatively anhydrous binding cavities and active sites of enzymes upon ligand binding [35] [52]. It should be noted, nevertheless, that incoming nucleophiles capable of forming a hydrogen-bonded transition state with the phosphorus species tends to favor frontside over backside S_N2_P, thus remaining a mechanistic plausibility for the phosphoryl (P=O) and trifluoromethyl (CF_3)-containing BMFP-C2 [52]. BMFP-C3, on the other hand, lacks the hydrogen-acceptor groups aforementioned, thus making a potential frontside attack unlikely.

Interestingly, a pseudo-"catalytic triad" group of amino acids in close proximity has been computationally observed in the vicinity of the fluorophosphonate moiety of BMFP-C1 in the ligand-protein complex, containing Asp320, Lys499, and Thr497 (**Figure 9**).

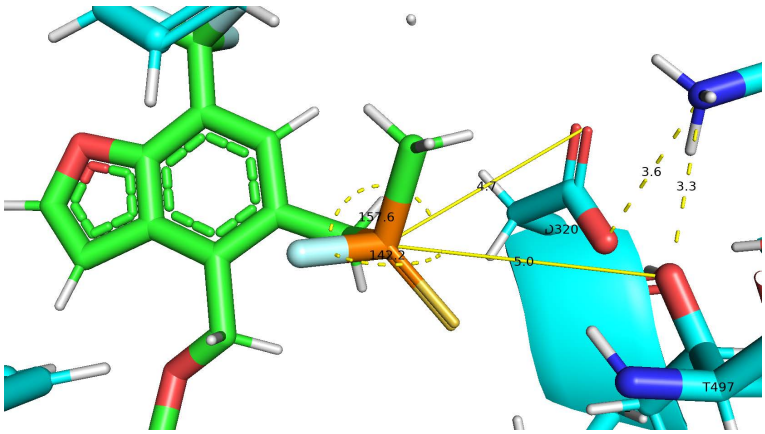

Figure 9. 3D structure of BMFP-C1 bound to the active site of CYP1A1, with pseudo-"catalytic triad" residues (Asp320, Lys499, and Thr497) in close, hydrogen-bonding (H-bonding) proximity (3.3 - 3.6 Å).

The above pseudo-"catalytic triad" moiety has already been experimentally observed in amidase and class-D β-lactamase enzymes, indicating its functional utility in enzyme-mediated hydrolysis [53] [54]. Shin *et al.* [54] support the function of lysine (Lys) in said triad as an "oxyanion hole", stabilizing the intermediate oxide of amide hydrolysis, whilst Leonard *et al.* [53] have found that carbamylated Lys functions as an effective nucleophile-base bifunctional moiety in β-lactam hydrolysis. It should be noted, likewise, that the P-Nu distance between the phosphorus of the ligand and either the resonance-stabilized oxides of Asp320 or hydroxyl of Thr497 differs only by 0.3 Å, suggesting nucleophile exchange between Asp320 and Thr497 as a kinetic possibility. Both oxygen species are, moreover, aligned at favorable angles for S_N2_P with a backside attack [41] [42].

Moreover, the predicted quantitative thermodynamic payout of exchange between an unstable phosphonocarbonate intermediate (Asp320-P) and a stable phosphonate ester makes this reaction especially interesting, considering phosphonate ester standard Gibbs free energy of formation (ΔG°_f) amounts up to -165.1 kcal/mol (diethyl phosphonate) [55]. For comparison, while the standard

Gibbs free energy of formation (ΔG°_{f}) of an analogous phosphonocarbonate species (*i.e. O*-ethylphosphonyl acetate) has not been reported in literature, Milev *et al.* [56] have found that—in an acetic acid and ethylene glycol solvent with a calcium diethoxide catalyst at 25˚C—only 36.1% yield was obtained from the diethyl hydrogen phosphonate reactant after 384 hours, indicating very slow reaction rate even with cation-mediated catalysis to stabilize oxyanion intermediates. As further comparison, the standard Gibbs free energy of formation (ΔG°_{f}) of unimolar ratios of acetate and ethyl chloride, ethyl bromide, and ethyl iodide have been experimentally observed to be 22.3, 20.0, and 16.6 kcal/mol, respectively, in DMSO, indicating significantly greater spontaneity in phosphonate ester formation under standard conditions [57]. The mechanistic hypothesis thus developed has been illustrated below (**Figure 10**).

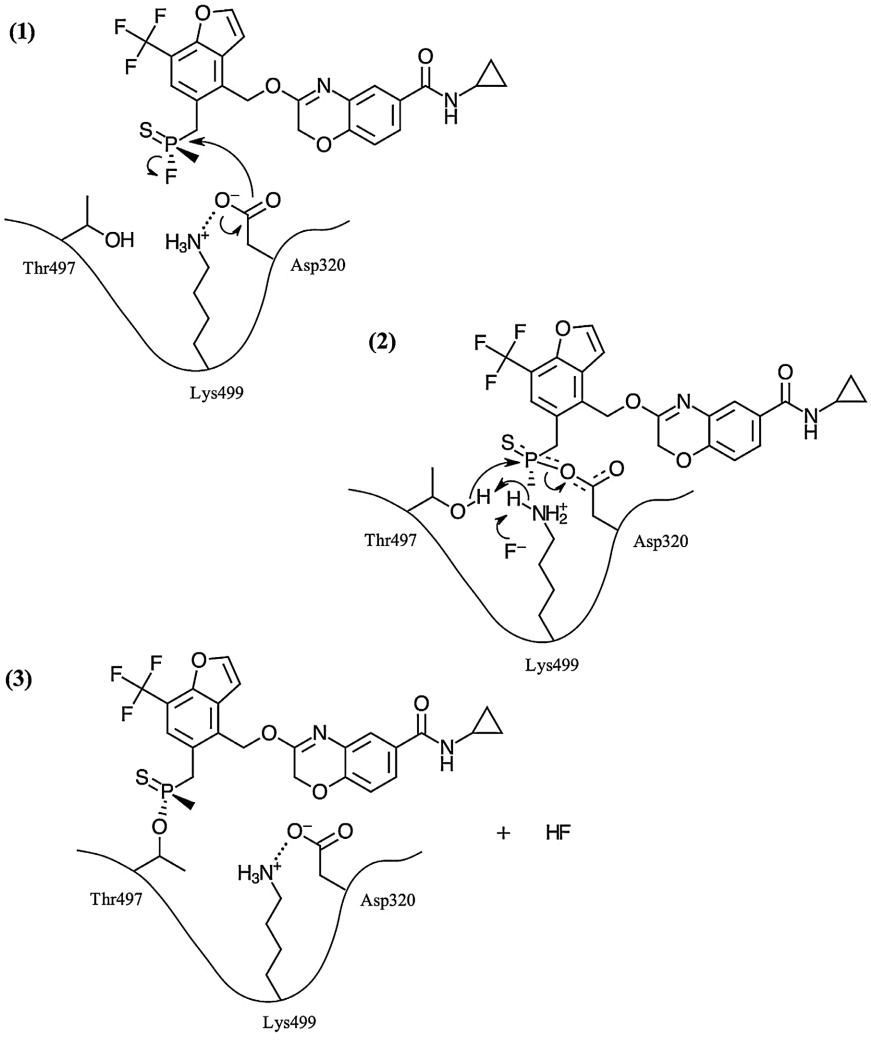

Figure 10. A mechanistic hypothesis considering the role of the pseudo-"catalytic triad" (Asp320, Lys499, Thr497) in BMFP-C1 covalent inhibition of CYP1A1, adapted from mechanistic research by Leonard *et al.* [53] and Shin *et al.* [54]. Note the protracted resonance stabilization of the phosphonocarbonate intermediate, and speculated retention of phosphorus stereochemistry in steps 1 and 3, as per Kolodiazhnyi & Kolodiazhna [35].

Nevertheless, it should be noted that—whilst proximity, AONA, and certain literature support the mechanistic hypothesis outlined above—this hypothesis remains purely speculative, and not catalytically proven *in vitro* or *ex vivo*.

Considering the above data holistically, computational covalent binding was conducted for each of the six ligands, with subsequent energy minimization of the covalent complexes in order to account for greater stabilization, so that further normal mode analysis (NMA) could be conducted to support a more encompassing conclusion on which of the six ligands could be deemed most mechanistically promising for heme-depleted CYP1A1 inhibition. In the cases of BMFP-C1 and BMFP-C6, and due to the instability of the phosphonocarbonate complex theorized in the Asp320-P intermediates, the second-closest, phosphonate ester-forming residue, Thr497, was covalently bound to each for the purpose of NMA as it would favor the thermodynamic product in accordance with the mechanism discussed previously.

3.3. Protein Motion Analysis: NMA

Normal mode analysis (NMA) was foremost conducted for ligand-free, heme-depleted *holo*-CYP1A1 as a reference standard for the six ligand-protein complexes (Figure 11).

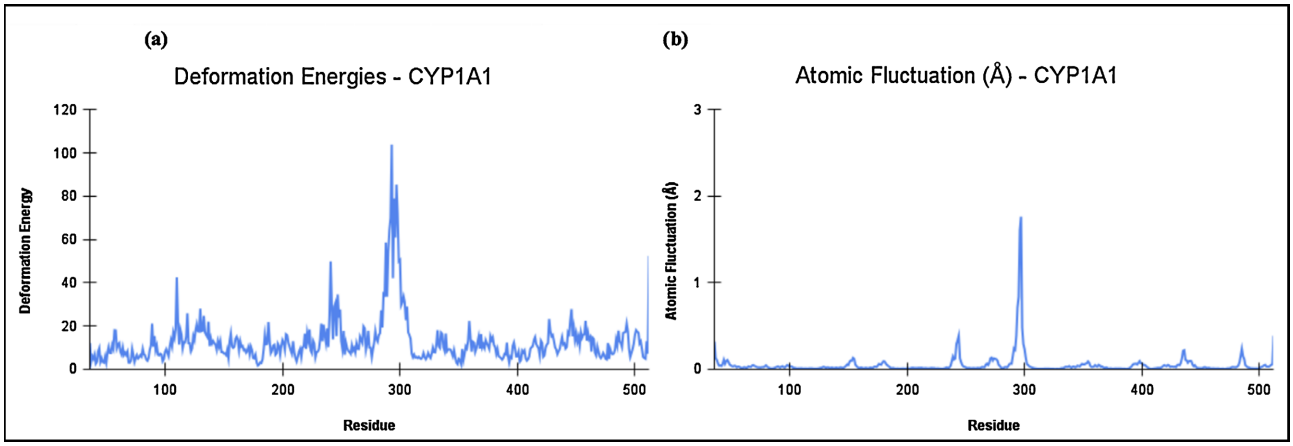

Figure 11. Deformation energies and atomic fluctuation (Å) of heme-depleted *holo*-CYP1A1. Note the deformation energies being normalized, thus having no units. Note the outlier peak corresponding to terminal Ser512.

The above data depicts the highest deformation energies around residues 290 - 300 (60 - 102 units), with the largest atomic fluctuation (~1.8 Å) being reported in the same region. This corresponds to the flexible loop structure discussed in Section 3.1, spanning Glu293 to Ser303 residues. While the modal deformation energies of the heme-depleted *holo*-CYP1A1 remained under 20 units, notable peaks reaching >40 units were seen around residues 105 and 240, corresponding to a loop-backbone basic pocket of Arg106 and Arg455, and an *α*-helix of the loop-helix-loop (LHL) motif discussed previously (see Figure 8) spanning residues Pro238 to Tyr242 (Figure 11). These motifs have been illustrated for further clar-

ity in Figure 12 below.

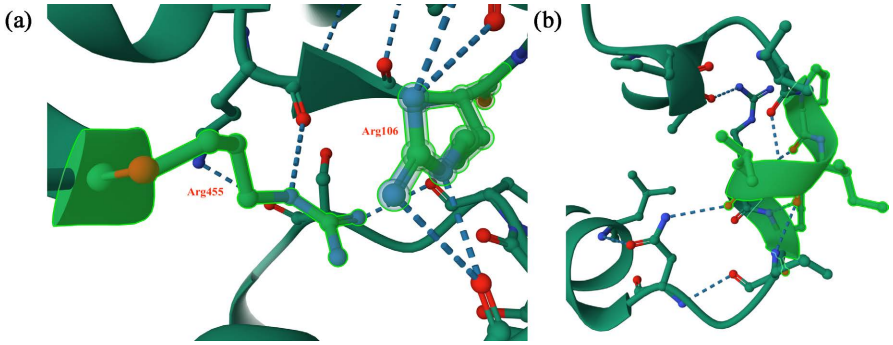

Figure 12. 3D protein structure depictions of (a) basic pocket of Arg106 and 455, and (b) short *α*-helix residue of a loop-helix-loop (LHL) motif spanning residues Phe236 to Asn245.

The calculated discrepancy between high deformation energies and simultaneous large atomic fluctuations at the above residue ranges indicate a relative instability in secondary protein structure (*i.e.* flexible loops), whilst maintaining a notable level of tertiary structure stability there where inter-side chain and side chain-backbone interactions (*i.e.* H-bonding between Arg455 and Arg106 with backbone amides) dominate.

Considering the NMA of the six ligand-CYP1A1 complexes in both the non-covalent/pre-covalent and covalent binding modes, furthermore, notable differences arise between both the complexes and the NMA of heme-depleted *holo*-CYP1A1, as well as inter-complex values (Figure 13).

As seen below, the deformation energies for both non-covalent and covalent binding modes of the six ligand-protein complexes do not exceed the deformation energies of heme-depleted *holo*-CYP1A1—rather, while remaining largely similar in the non-covalent mode, the covalent mode had a notable decrease thereof (102 units to a maximum of 90, residue Asp295). Moreover, the maximum value of deformation energy for the non-covalent mode (102 units) was reported with BMFP-C2, whilst BMFP-C5 retained the highest score for the covalent mode (90 units). Interestingly, both modes—in spite of the modal similarity (<20 units) of deformation energies with heme-depleted *holo*-CYP1A1—had a marked increase thereof around residues 110 - 120 (~20 to ~70 units) and 150 - 160 (~15 to ~65 units), specifically with ligands BMFP-C2, BMFP-C5, and BMFP-C6. Deformation energies at residue 400 (Tyr) likewise showed dramatic increases between heme-depleted *holo*-CYP1A1 and BMFP-C3 & BMFP-C6 (non-covalent), and BMFP-C1 & BMFP-C3 (covalent), rising from ~10 units (heme-depleted *holo*-CYP1A1) to maximum values of 70 units for both BMFP-C3 (non-covalent) and BMFP-C1 (covalent), respectively.

The atomic fluctuations (Å) for the non-covalent binding mode of the six ligand-protein complexes, on the other hand, had reported a higher maximum distance (2.6 Å) than both the covalent and ligand-free *holo*-CYP1A1 (1.8 Å,

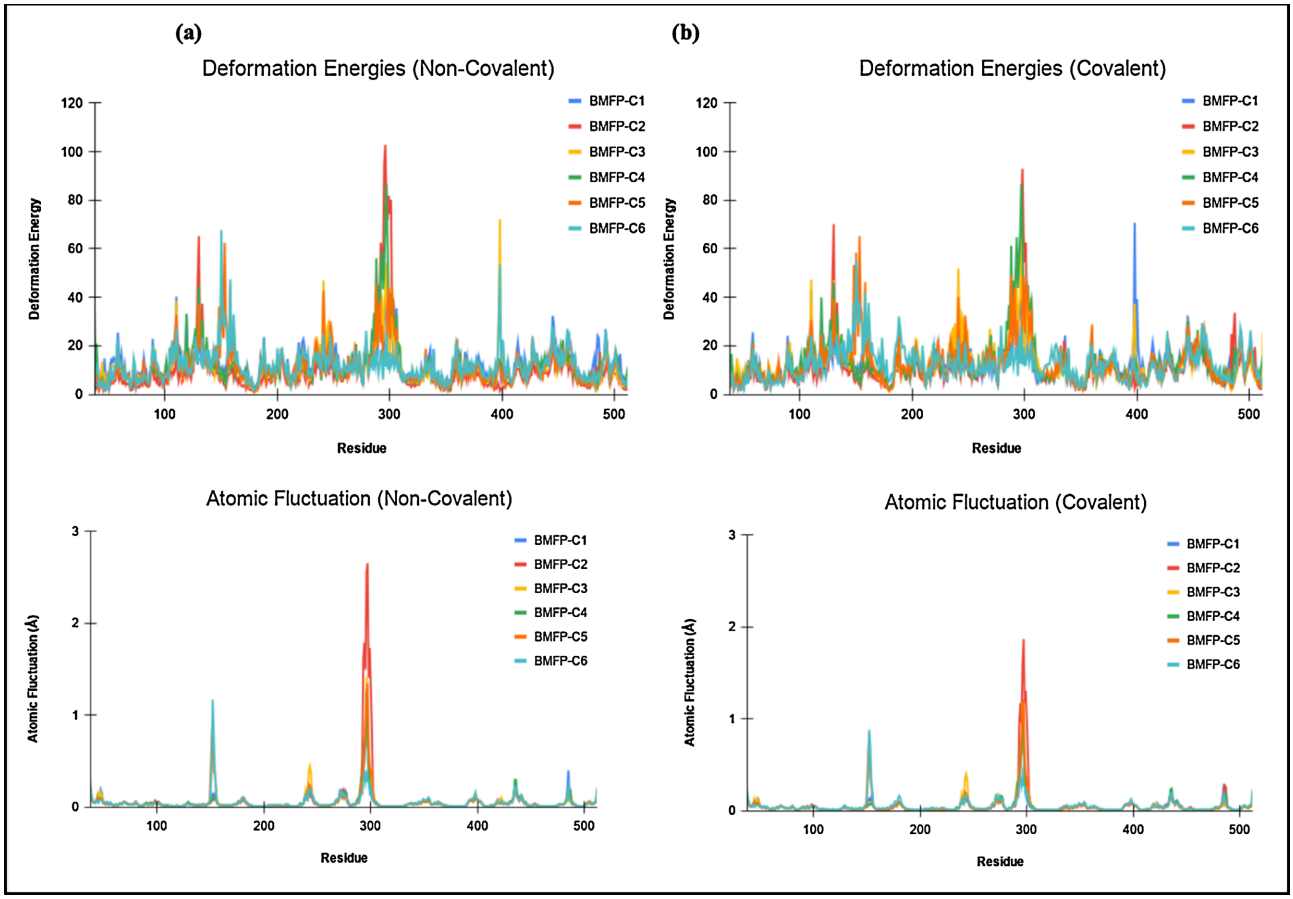

Figure 13. Deformation energies and atomic fluctuation (Å) of the respective six ligand-CYP1A1 complexes (BMFP-C1, BMFP-C2, BMFP-C3, BMFP-C4, BMFP-C5, and BMFP-C6), compared for both non-covalent (a) and covalent (b) binding modes.

respectively). This maximum fluctuation distance for both the non-covalent and covalent binding modes (2.6 Å and 1.8Å) was attributed to BMFP-C2. Moreover, a significant increase of atomic fluctuation at residue 150 has been observed in the non-covalent (~1.2 Å), and—to a lesser extent—covalent (~0.9 Å) modes of BMFP-C6, as compared to heme-depleted *holo*-CYP1A1 (~0.1 Å). This correlates to a helix-loop-helix (HLH) motif loop, ranging from Ile147 to Cys157. Nevertheless, no changes in atomic fluctuation are seen around the nucleophilic residues responsible for binding the phosphorus species in any of the ligands: Ser116 or Thr497. Furthermore, the other ligands have had no noticeable change on atomic fluctuations in comparison to heme-depleted *holo*-CYP1A1 aside from a slight decrease in atomic fluctuation around residue 240—in both non-covalent and covalent modes—by all except BMFP-C3.

The most stabilizing ligand in terms of computed deformation energies required to deform the protein thus appears to be BMFP-C2, considering its deformation energy in both non-covalent and covalent binding modes at residues 110 (Tyr), 295 (Asp), and 490 (Met) compared to the other five ligands (BMFP-C1, BMFP-C3, BMFP-C4, BMFP-C5, and BMFP-C6) being largest. Nevertheless, its remarkably poor performance in reducing atomic fluctuation of the ligand-pro-

tein complex renders it an unfavorable candidate for optimal thermodynamic stability in our opinion, alongside its general inability to raise deformation energy compared to heme-depleted *holo*-CYP1A1 at most residues (with the exception of Tyr110). Moreover, the related ligand, BMFP-C6, has a marked predicted superiority in increasing deformation energy requirements at residue 150 (Asp) in the non-covalent mode. Similarly, in spite of its deformation energies in the covalent mode generally not exceeding those of BMFP-C2, BMFP-C6 remains a lower overall contributor to atomic fluctuation distance than the latter, and BMFP-C3 and BMFP-C5.

BMFP-C1 and BMFP-C4, on the other hand, had no peaks indicating drastically different atomic fluctuations at specific residues, in comparison, aside from a brief 0.2 - 0.3 Å peak difference at residues 490 - 500 in the covalent mode for BMFP-C1. On the other hand, BMFP-C4 reported consistently high deformation energies comparable to BMFP-C2/BMFP-C5 (residues 293 - 303) in both non-covalent and covalent binding modes, superseded by BMFP-C1 only at residue 400 (Tyr) in the covalent mode.

With regards to the above workflow results throughout the prior three sections, it was thus interpreted that—in spite of certain shortcomings compared to other ligands (*i.e.* BMFP-C2-C6 in absolute non-covalent—and BMFP-C2, C5, and C6 in rDock—docking scores, BMFP-C2 and C3 in RMSD/RMSF and proximal P-Nu distance, and BMFP-C2 and C6 in NMA)—its relative consistency in ranking positively across the computational testing conducted rendered it the most optimal candidate for hypothetical heme-depleted *holo*-CYP1A1 inhibition. It was, thereby, observed that no other ligand of the BMFP class of compounds presented such a "Goldilocks zone" series of results that did not drastically fluctuate between different methods of analysis of the same simulation (ie. RMSD versus RMSF in MD, or Atomic Fluctuations and Deformation Energies in NMA), thus rendering BMFP-C1 the most promising *in silico* candidate for this paper's objective, with numerous favorable parameters (*i.e.* absolute docking score of −10.2 kcal/mol, 4.7 Å P-Nu distance and AONA of 157.6˚ within suitable S_N2_P range, RMSD plateau time of 2.5 ns and <0.25 nm well within the lower end of the time window of heme-depleted *holo*-CYP1A1 relaxation to *apo*-CYP1A1, and relatively consistent NMA results).

In further consideration to the pharmacodynamics and pharmacokinetics of fluorophosphonates, as the BMFP class of compounds herein investigated, on the other hand, it should be noted that the risk of significant off-site effects remains, nevertheless. While mitigations via rational design and targeting of CYP1A1 have been made in this study for a corresponding decrease in potential off-site effects, the risk of the fluorophosphonate moiety cross-reacting with other—potentially dangerous—nucleophilic residues (*i.e.* Ser203 Oγ of the Ser203-His447-Glu334 catalytic triad of acetylcholinesterase [AChE], akin to lethal nerve agents such as sarin and soman) cannot be excluded without adequate toxicological assays [58]. As further support to the selectivity of BMFP-C1-C6 for heme-depleted *holo*-CYP1A1 over other related isoforms, such as CYP1A2 and CYP1B1, preliminary

comparative computational screening (*Autodock Vina*) as target validation yielded a more selective average docking score for the active site of heme-depleted *holo*-CYP1A1 than the latter, though the scope of this preliminary validation remains beyond the objective of this paper. Nonetheless, any future synthetic—and subsequently pharmacological—directions with the BMFP compound class should, for safety purposes, note this hazard—and its corresponding high risk—posed, in spite of the targeted design of the ligands herein computed.

4. Conclusions

With consideration of all of the above analyses overall, we conclude that ligand BMFP-C1 was computationally suggested to be most optimal for the purpose of theoretical covalent inhibition of heme-depleted *holo*-CYP1A1, via a hypothesized "catalytic triad"-mediated S_N2_P reaction, initially with Asp320 (kinetic product), followed by Thr497 (thermodynamic product). This is because—in spite of advantages posed by certain other ligands in each of the analytic methods scored in the previous section (docking, K_{DEEP}, RMSD/RMSF, covalent attack geometry, and NMA)—BMFP-C1's unique mechanistic susceptibility to covalent attack in the "catalytic triad" (Thr497-Lys499-Asp320) due to its geometry (4.7 Å P-Nu distance, 157.6˚ AONA), its large docking score (−10.2 kcal/mol) indicative of strong non-covalent binding, its fast-plateauing and stable RMSD (2.5 ns, <0.25 nm), and its relatively-consistent and non-fluctuating NMA results (deformation energies and atomic fluctuations), all serve indicative purposes to it being chosen as the best plausible inhibitor molecule for CYP1A1.

Any future direction should, we are of the opinion, thus point towards further *in vitro* synthetic attempts at synthesizing and testing (*i.e.* IC_{50} bioassays) the computed properties of BMFP-C1 in order to verify or annul our predictions made within this study. Moreover, within this future research, we recommend the specific screening of selectivity for CYP1A1 over related isozymes (*i.e.* CYP1A2 and CYP1B1) of this ligand and any/all of its derivatives.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions

Filip Latkovic: Writing—Original Draft, Conceptualization, Methodology, Investigation, **Fabio Di Ricco:** Conceptualization, Validation.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Ding, X. and Kaminsky, L.S. (2003) Human Extrahepatic Cytochromes P450: Func-

tion in Xenobiotic Metabolism and Tissue-Selective Chemical Toxicity in the Respiratory and Gastrointestinal Tracts. *Annual Review of Pharmacology and Toxicology*, **43**, 149-173. https://doi.org/10.1146/annurev.pharmtox.43.100901.140251

[2] Patel, T., Rahimi, N. and Cassagnol, M. (2026) Biochemistry, Cytochrome P450. StatPearls. https://www.ncbi.nlm.nih.gov/books/NBK557698/

[3] Zhao, M., Ma, J., Li, M., Zhang, Y., Jiang, B., Zhao, X., *et al.* (2021) Cytochrome P450 Enzymes and Drug Metabolism in Humans. *International Journal of Molecular Sciences*, **22**, Article 12808. https://doi.org/10.3390/ijms222312808

[4] Chiang, H., Wang, C. and Tsou, T. (2012) Differential Distribution of CYP2A6 and CYP2A13 in the Human Respiratory Tract. *Respiration*, **84**, 319-326. https://doi.org/10.1159/000339591

[5] Androutsopoulos, V.P., Spyrou, I., Ploumidis, A., Papalampros, A.E., Kyriakakis, M., Delakas, D., *et al.* (2013) Expression Profile of CYP1A1 and CYP1B1 Enzymes in Colon and Bladder Tumors. *PLOS ONE*, **8**, e82487. https://doi.org/10.1371/journal.pone.0082487

[6] Li, F., Zhu, W. and Gonzalez, F.J. (2017) Potential Role of CYP1B1 in the Development and Treatment of Metabolic Diseases. *Pharmacology & Therapeutics*, **178**, 18-30. https://doi.org/10.1016/j.pharmthera.2017.03.007

[7] Stipp, M.C., Nardin, J.M., Casali-da-Rocha, J.C., Ioshii, S. and Acco, A. (2025) Implications of CYP1B1 in the Treatment and Prognosis of Breast Cancer. *Beni-Suef University Journal of Basic and Applied Sciences*, **14**, Article No. 22. https://doi.org/10.1186/s43088-025-00609-x

[8] Carrera, A.N., Grant, M.K.O. and Zordoky, B.N. (2020) CYP1B1 as a Therapeutic Target in Cardio-Oncology. *Clinical Science*, **134**, 2897-2927. https://doi.org/10.1042/cs20200310

[9] Kumarakulasingham, M., Rooney, P.H., Dundas, S.R., Telfer, C., Melvin, W.T., Curran, S., *et al.* (2005) Cytochrome P450 Profile of Colorectal Cancer: Identification of Markers of Prognosis. *Clinical Cancer Research*, **11**, 3758-3765. https://doi.org/10.1158/1078-0432.ccr-04-1848

[10] Fasinu, P.S. and Rapp, G.K. (2019) Herbal Interaction with Chemotherapeutic Drugs—A Focus on Clinically Significant Findings. *Frontiers in Oncology*, **9**, Article 1356. https://doi.org/10.3389/fonc.2019.01356

[11] Hrycay, E.G. and Bandiera, S.M. (2015) Involvement of Cytochrome P450 in Reactive Oxygen Species Formation and Cancer. *Advances in Pharmacology*, **74**, 35-84. https://doi.org/10.1016/bs.apha.2015.03.003

[12] Ingelman-Sundberg, M. and Lauschke, V.M. (2020) Can CYP Inhibition Overcome Chemotherapy Resistance? *Trends in Pharmacological Sciences*, **41**, 503-506. https://doi.org/10.1016/j.tips.2020.05.007

[13] Liu, J., Wang, Q., Wu, D., Wang, X., Sun, Y., Chen, X., *et al.* (2004) Differential Regulation of CYP1A1 and CYP1B1 Expression in Resveratrol-Treated Human Medulloblastoma Cells. *Neuroscience Letters*, **363**, 257-261. https://doi.org/10.1016/j.neulet.2004.03.075

[14] Rodriguez, M. and Potter, D.A. (2013) CYP1A1 Regulates Breast Cancer Proliferation and Survival. *Molecular Cancer Research*, **11**, 780-792. https://doi.org/10.1158/1541-7786.mcr-12-0675

[15] Uppstad, H., Øvrebø, S., Haugen, A. and Mollerup, S. (2010) Importance of CYP1A1 and CYP1B1 in Bioactivation of Benzo[a]Pyrene in Human Lung Cell Lines. *Toxicology Letters*, **192**, 221-228. https://doi.org/10.1016/j.toxlet.2009.10.025

[16] Costa, P.M. (2021) Current Aspects of DNA Damage and Repair in Ecotoxicology: A Mini-Review. *Ecotoxicology*, **31**, 1-11. https://doi.org/10.1007/s10646-021-02487-2

[17] Joshi, P., McCann, G.J.P., Sonawane, V.R., Vishwakarma, R.A., Chaudhuri, B. and Bharate, S.B. (2017) Identification of Potent and Selective CYP1A1 Inhibitors via Combined Ligand and Structure-Based Virtual Screening and Their *in Vitro* Validation in Sacchrosomes and Live Human Cells. *Journal of Chemical Information and Modeling*, **57**, 1309-1320. https://doi.org/10.1021/acs.jcim.7b00095

[18] Almira Correia, M., Sinclair, P.R. and De Matteis, F. (2010) Cytochrome P450 Regulation: The Interplay between Its Heme and Apoprotein Moieties in Synthesis, Assembly, Repair, and Disposal. *Drug Metabolism Reviews*, **43**, 1-26. https://doi.org/10.3109/03602532.2010.515222

[19] Chien, S., Shoji, O., Morimoto, Y. and Watanabe, Y. (2017) Use of Apomyoglobin to Gently Remove Heme from a H_2O_2-Dependent Cytochrome P450 and Allow Its Reconstitution. *New Journal of Chemistry*, **41**, 302-307. https://doi.org/10.1039/c6nj02882a

[20] Wagner, G.C., Perez, M., Toscano, W.A. and Gunsalus, I.C. (1981) Apoprotein Formation and Heme Reconstitution of Cytochrome P-450$_{cam}$. *Journal of Biological Chemistry*, **256**, 6262-6265. https://doi.org/10.1016/s0021-9258(19)69157-x

[21] Correia, M.A. and Meyer, U.A. (1975) Apocytochrome P-450: Reconstitution of Functional Cytochrome with Hemin *in Vitro*. *Proceedings of the National Academy of Sciences of the United States of America*, **72**, 400-404. https://doi.org/10.1073/pnas.72.1.400

[22] Karuzina, I.I., Zgoda, V.G., Kuznetsova, G.P., Samenkova, N.F. and Archakov, A.I. (1999) Heme and Apoprotein Modification of Cytochrome P450 2B4 during Its Oxidative Inactivation in Monooxygenase Reconstituted System. *Free Radical Biology and Medicine*, **26**, 620-632. https://doi.org/10.1016/s0891-5849(98)00252-4

[23] Sun, L., Wang, Z., Jiang, H., Tan, X. and Huang, Z. (2010) Novel Conformational Transitions of Human Cytochrome P450 2C8 during Thermal and Acid-Induced Unfolding. *Chinese Journal of Chemistry*, **28**, 1491-1502. https://doi.org/10.1002/cjoc.201090255

[24] Sun, Y., Zeng, W., Benabbas, A., Ye, X., Denisov, I., Sligar, S.G., *et al.* (2013) Investigations of Heme Ligation and Ligand Switching in Cytochromes P450 and P420. *Biochemistry*, **52**, 5941-5951. https://doi.org/10.1021/bi400541v

[25] Tetreau, C., Tourbez, M. and Lavalette, D. (2000) Conformational Relaxation in Hemoproteins: The Cytochrome P-450$_{cam}$ Case. *Biochemistry*, **39**, 14219-14231. https://doi.org/10.1021/bi000920u

[26] Liou, G. and Storz, P. (2010) Reactive Oxygen Species in Cancer. *Free Radical Research*, **44**, 479-496. https://doi.org/10.3109/10715761003667554

[27] Madshus, I.H. (1988) Regulation of Intracellular Ph in Eukaryotic Cells. *Biochemical Journal*, **250**, 1-8. https://doi.org/10.1042/bj2500001

[28] Nakamura, H. and Takada, K. (2021) Reactive Oxygen Species in Cancer: Current Findings and Future Directions. *Cancer Science*, **112**, 3945-3952. https://doi.org/10.1111/cas.15068

[29] White, K.A., Grillo-Hill, B.K. and Barber, D.L. (2017) Cancer Cell Behaviors Mediated by Dysregulated Ph Dynamics at a Glance. *Journal of Cell Science*, **130**, 663-669. https://doi.org/10.1242/jcs.195297

[30] Brandl, N., Seitz, R., Sendtner, N., Müller, M. and Gülow, K. (2025) Living on the Edge: ROS Homeostasis in Cancer Cells and Its Potential as a Therapeutic Target.

Antioxidants, **14**, Article 1002. https://doi.org/10.3390/antiox14081002

[31] McFadyen, M.C.E., Melvin, W.T. and Murray, G.I. (2004) Cytochrome *p*450 Enzymes: Novel Options for Cancer Therapeutics. *Molecular Cancer Therapeutics*, **3**, 363-371. https://doi.org/10.1158/1535-7163.363.3.3

[32] Nitti, M., Ivaldo, C., Traverso, N. and Furfaro, A.L. (2021) Clinical Significance of Heme Oxygenase 1 in Tumor Progression. *Antioxidants*, **10**, Article 789. https://doi.org/10.3390/antiox10050789

[33] Dutra, J.K., Foley, T.L., Huang, Z., Fisher, E.L., Lachapelle, E.A., Mahapatra, S., *et al.* (2021) Fluorophosphonates On-Demand: A General and Simplified Approach toward Fluorophosphonate Synthesis. *ChemBioChem*, **22**, 1769-1774. https://doi.org/10.1002/cbic.202000852

[34] Liu, Y., Patricelli, M.P. and Cravatt, B.F. (1999) Activity-Based Protein Profiling: The Serine Hydrolases. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 14694-14699. https://doi.org/10.1073/pnas.96.26.14694

[35] Kolodiazhnyi, O.I. and Kolodiazhna, A. (2017) Nucleophilic Substitution at Phosphorus: Stereochemistry and Mechanisms. *Tetrahedron*: *Asymmetry*, **28**, 1651-1674. https://doi.org/10.1016/j.tetasy.2017.10.022

[36] Miao, Y. and Baudry, J. (2011) Active-Site Hydration and Water Diffusion in Cytochrome P450cam: A Highly Dynamic Process. *Biophysical Journal*, **101**, 1493-1503. https://doi.org/10.1016/j.bpj.2011.08.020

[37] Kiss, R., Sandor, M. and Szalai, F.A. (2012) http://Mcule.com: A Public Web Service for Drug Discovery. *Journal of Cheminformatics*, **4**, Article No. P17. https://doi.org/10.1186/1758-2946-4-s1-p17

[38] Walsh, S.E. (2013) RCSB PDB-4I8V: Human Cytochrome P450 1A1 in Complex with *α*-Naphthoflavone. https://www.rcsb.org/structure/4I8V

[39] Walsh, A.A., Szklarz, G.D. and Scott, E.E. (2013) Human Cytochrome P450 1A1 Structure and Utility in Understanding Drug and Xenobiotic Metabolism. *Journal of Biological Chemistry*, **288**, 12932-12943. https://doi.org/10.1074/jbc.m113.452953

[40] Torrens-Fontanals, M., Tourlas, P., Doerr, S. and De Fabritiis, G. (2024) Playmolecule Viewer: A Toolkit for the Visualization of Molecules and Other Data. *Journal of Chemical Information and Modeling*, **64**, 584-589. https://doi.org/10.1021/acs.jcim.3c01776

[41] Kolodiazhnyi, O.I. (2018) Stereochemistry of Electrophilic and Nucleophilic Substitutions at Phosphorus. *Pure and Applied Chemistry*, **91**, 43-57. https://doi.org/10.1515/pac-2018-0807

[42] Torres, R.A. and Bruice, T.C. (1998) Molecular Dynamics Study Displays near In-Line Attack Conformations in the Hammerhead Ribozyme Self-Cleavage Reaction. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 11077-11082. https://doi.org/10.1073/pnas.95.19.11077

[43] Rodrigues, C.H., Pires, D.E. and Ascher, D.B. (2018) DynaMut: Predicting the Impact of Mutations on Protein Conformation, Flexibility and Stability. *Nucleic Acids Research*, **46**, W350-W355. https://doi.org/10.1093/nar/gky300

[44] Velazquez-Campoy, A. (2019) What KD Value Should Be Considered as Strong for a Protein-Protein Interaction? https://www.researchgate.net/post/What_Kd_Value_should_be_considered_as_strong_for_a_protein-protein_interaction

[45] Ma, W., Yang, L. and He, L. (2018) Overview of the Detection Methods for Equilibrium Dissociation Constant KD of Drug-Receptor Interaction. *Journal of Pharma-*

ceutical Analysis, **8**, 147-152. https://doi.org/10.1016/j.jpha.2018.05.001

[46] Papke, B., Murarka, S., Vogel, H.A., Martín-Gago, P., Kovacevic, M., Truxius, D.C., *et al.* (2016) Identification of Pyrazolopyridazinones as PDEδ Inhibitors. *Nature Communications*, **7**, Article No. 11360. https://doi.org/10.1038/ncomms11360

[47] Ribeiro, R.P. (2022) What Is the Acceptance Value of RMSD? https://www.researchgate.net/post/What_is_the_acceptance_value_of_RMSD

[48] Williams, L.D. (2025) Molecular Interactions (Noncovalent Interactions). The Williams Lab. https://williams.chemistry.gatech.edu/structure/molecular_interactions/mol_int.html#E3

[49] Chen, X., Zhao, X., Xiong, Y., Liu, J. and Zhan, C. (2011) Fundamental Reaction Pathway and Free Energy Profile for Hydrolysis of Intracellular Second Messenger Adenosine 3',5'-Cyclic Monophosphate (Camp) Catalyzed by Phosphodiesterase-4. *The Journal of Physical Chemistry B*, **115**, 12208-12219. https://doi.org/10.1021/jp205509w

[50] Wilson, A.L. and White, W.E. (2001) *Ab Initio* Studies on Hexavalent Phosphorus Compounds. 2001 *ECBC Scientific Conference on Chemical and Biological Defense Research*, Hunt Valley, 6-8 March 2001, 1-11.

[51] Wang, J., Gu, J. and Leszczynski, J. (2008) Theoretical Modeling Study for the Phosphonylation Mechanisms of the Catalytic Triad of Acetylcholinesterase by Sarin. *The Journal of Physical Chemistry B*, **112**, 3485-3494. https://doi.org/10.1021/jp076974w

[52] Dey, S. (2015) A Review on Nucleophilic Substitution Reactions at P=O Substrates of Organophosphorous Compounds. *American Journal of Organic Chemistry*, **5**, 83-94.

[53] Leonard, D.A., Bonomo, R.A. and Powers, R.A. (2013) Class D β-Lactamases: A Reappraisal after Five Decades. *Accounts of Chemical Research*, **46**, 2407-2415. https://doi.org/10.1021/ar300327a

[54] Shin, S., Yun, Y.S., Koo, H.M., Kim, Y.S., Choi, K.Y. and Oh, B. (2003) Characterization of a Novel Ser-CisSer-Lys Catalytic Triad in Comparison with the Classical Ser-His-Asp Triad. *Journal of Biological Chemistry*, **278**, 24937-24943. https://doi.org/10.1074/jbc.m302156200

[55] Guthrie, J.P. (1979) Tautomerization Equilibria for Phosphorous Acid and Its Ethyl Esters, Free Energies of Formation of Phosphorous and Phosphonic Acids and Their Ethyl Esters, and pK_a Values for Ionization of the P-H Bond in Phosphonic Acid and Phosphonic Esters. *Canadian Journal of Chemistry*, **57**, 236-239. https://doi.org/10.1139/v79-039

[56] Milev, A.S., Kamali Kannangara, G.S., Ben-Nissan, B. and Wilson, M.A. (2004) Temperature Effects on a Hydroxyapatite Precursor Solution. *The Journal of Physical Chemistry B*, **108**, 5516-5521. https://doi.org/10.1021/jp035548h

[57] Tondo, D.W. and Pliego, J.R. (2004) Modeling Protic to Dipolar Aprotic Solvent Rate Acceleration and Leaving Group Effects in S_n2 Reactions: A Theoretical Study of the Reaction of Acetate Ion with Ethyl Halides in Aqueous and Dimethyl Sulfoxide Solutions. *The Journal of Physical Chemistry A*, **109**, 507-511. https://doi.org/10.1021/jp047386a

[58] Hung, L., Sanbonmatsu, K.Y., Williams, R.F. and Chen, J.C.H. (2025) Acetylcholinesterase: Structure, Dynamics, and Interactions with Organophosphorus Compounds. *Protein Science*, **34**, e70297. https://doi.org/10.1002/pro.70297