

ORIGINAL ARTICLES

STATISTICAL EVALUATION OF *IN SILICO* PHARMACOKINETIC PREDICTORS OF ISATIN-BASED COMPOUNDS

Kaloyan Mihalev¹, Ivelin Iliev¹, Svetlana Petkova², Svetlana Georgieva¹

¹*Department of Pharmaceutical Chemistry, Faculty of Pharmacy,
Medical University of Varna, Bulgaria*

²*Dr. Ivan Bogorov Vocational High School of Economics and Technology, Varna, Bulgaria*

ABSTRACT

This study presents a comprehensive *in silico* pharmacokinetic analysis of 30 novel isatin derivatives designed as potential antibacterial agents. Physicochemical and ADMET parameters were predicted using SwissADME and ADMETlab 2.0, and the relationships between key descriptors were evaluated by Pearson correlation analysis and multiple linear regression. All compounds complied with Lipinski's Rule of Five. Lipophilicity (logP) showed a strong positive correlation with the volume of distribution (VD_{ss}; $r = 0.810$, $p < 0.001$) and a moderate positive correlation with plasma protein binding ($r = 0.480$, $p = 0.007$), while no significant dependence of systemic clearance on logP or plasma protein binding was observed. Multiple regression analysis identified VD_{ss} as the dominant determinant of elimination half-life ($\beta = -1.163$, $p < 0.001$), indicating that distribution-driven processes govern pharmacokinetic variability within this scaffold. The findings provide a rational basis for the optimization of isatin-based compounds with improved ADMET profiles.

Keywords: *isatin derivatives, in silico pharmacokinetics, ADMET, lipophilicity, Pearson correlation, drug design*

INTRODUCTION

Antimicrobial resistance (AMR) represents one of the most pressing challenges in contemporary medicine, with multidrug-resistant bacterial pathogens increasingly undermining the clinical efficacy of established antibiotic classes (1). The World Health Organization has identified AMR as a global health priority, estimating that resistant infections already account for hundreds of thousands of deaths annually, a burden projected to rise substantially over the coming decades (2). In this context, the discovery and development of structurally novel antibacterial

agents with distinct mechanisms of action is of paramount importance (3). Heterocyclic scaffolds with demonstrated broad-spectrum activity are particularly attractive starting points, as they offer opportunities for systematic chemical diversification to optimize both biological potency and pharmacokinetic behavior (4).

Isatin (1H-indole-2,3-dione) is a naturally occurring bicyclic compound found in various plant species and produced endogenously in humans from tryptophan metabolism (5). Its unique indole-2,3-dione core confers remarkable chemical versatility, enabling facile functionalization at the N-1 position, C-3 carbonyl, and multiple ring positions (6). This structural plasticity has been extensively exploited in medicinal chemistry, resulting in a broad library of derivatives with documented antibacterial, antifungal, antiviral, anticonvulsant, anticancer, and anti-inflammatory activities (7–12). Several isatin-containing compounds have progressed to clinical evaluation, most notably isatin-based analogues active

Address for correspondence:

Kaloyan Mihalev
Faculty of Pharmacy
Medical University of Varna
84 Tzar Osvoboditel Blvd
9002 Varna, Bulgaria
e-mail: kaloyan.mihalev@mu-varna.bg

Received: November 2, 2025

Accepted: December 14, 2025



against HIV integrase and influenza neuraminidase, further validating this scaffold as a privileged structure in drug discovery (13). The ability to modulate biological activity through targeted structural modifications—particularly N-alkylation and halogenation at the C-5 or C-6 positions—makes isatin derivatives especially amenable to systematic structure-activity and structure-pharmacokinetic relationship studies.

A critical yet frequently underappreciated determinant of clinical success in antibacterial drug development is the pharmacokinetic profile of a candidate compound. Even molecules demonstrating potent *in vitro* activity may fail *in vivo* if they cannot achieve and maintain adequate drug concentrations at the site of infection (14). Key pharmacokinetic parameters—including volume of distribution (VDss), systemic clearance (CL), plasma half-life ($t_{1/2}$), and plasma protein binding (PPB)—collectively govern the absorption, distribution, metabolism, and elimination (ADME) behavior of a compound in the biological milieu. High plasma protein binding reduces the free drug fraction available for antimicrobial action, while unfavorable distribution characteristics may limit tissue penetration in compartments where pathogens reside (15). Rapid systemic clearance shortens the dosing interval and may necessitate higher or more frequent doses, with consequent implications for tolerability and patient compliance (16). Understanding how structural features translate into pharmacokinetic properties is therefore essential for the rational optimization of antibacterial candidates (17).

In recent years, *in silico* approaches have become indispensable tools in early-stage drug discovery, enabling the rapid and cost-effective prediction of ADMET properties prior to synthesis (18). Computational platforms such as SwissADME and ADMETlab 2.0 leverage machine learning models trained on large experimental datasets to provide reliable estimates of physicochemical descriptors, oral absorption, tissue distribution, metabolic liability, and excretion parameters for virtual compound libraries (19). These tools allow researchers to identify and eliminate pharmacokinetically unfavorable candidates at the design stage—a strategy commonly referred to as *ADMET filtering*—thereby reducing attrition rates in later development phases (20). The va-

lidity of such *in silico* predictions has been confirmed in multiple benchmark studies, and their integration into medicinal chemistry workflows is now considered best practice by regulatory and industry standards alike (21). Importantly, for novel compounds that have not yet been synthesized, *in silico* ADMET profiling represents the only feasible approach to assess pharmacokinetic viability and guide structural prioritization (22).

Classical pharmacokinetic theory posits that physicochemical properties, particularly lipophilicity (logP) and topological polar surface area (TPSA), are primary determinants of membrane permeability, tissue distribution, and systemic clearance (23,24). Compounds with higher TPSA, indicative of greater hydrophilicity, are generally expected to exhibit a lower volume of distribution due to limited tissue penetration and preferential retention in the plasma compartment (25). Conversely, lipophilic compounds with higher logP values typically demonstrate greater tissue partitioning, resulting in a larger VDss. However, the quantitative relationships between these descriptors and pharmacokinetic parameters can vary substantially across different chemical series, and their predictive value for a given scaffold must be established empirically. Additional factors—including active transport processes, CYP-mediated metabolism, and plasma protein binding—may also significantly modulate the *in vivo* pharmacokinetic behavior in ways that are not straightforwardly predicted from lipophilicity alone (26).

Despite the recognized pharmacological potential of isatin derivatives, systematic *in silico* pharmacokinetic analyses of structurally diversified isatin series—encompassing both N-1 alkyl substitution and C-6 halogenation—remain scarce in the literature. A quantitative understanding of how these two key structural variables jointly influence ADMET parameters would provide a rational framework for the prioritization of candidates prior to synthesis and biological evaluation. To address this gap, the present study conducted a comprehensive *in silico* pharmacokinetic analysis of 30 novel isatin derivatives with potential antibacterial activity, applying Pearson correlation analysis and multiple linear regression to characterize the relationships between physicochemical descriptors and predicted pharmacokinetic parameters.

AIM AND OBJECTIVES

The aim of this study was to characterize the *in silico* pharmacokinetic profile of 30 novel isatin derivatives and to establish quantitative structure-pharmacokinetic relationships. The specific objectives were: (i) to predict and compare physicochemical and ADMET parameters across six N-1 substitution patterns and five C-6 halogenation states; (ii) to quantify bivariate correlations between logP and VDss, CL, $t_{1/2}$, and PPB using Pearson analysis; (iii) to determine the relative contributions of VDss and CL to half-life variability by multiple linear regression; and (iv) to translate the identified relationships into drug design recommendations for scaffold optimization.

MATERIALS AND METHODS

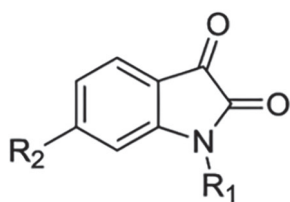
Compound Selection

Thirty novel isatin derivatives, designated compounds 1–30, were selected for this study. These compounds were designed *in silico* as potential antibacterial agents and have not yet been synthesized. Their structures were represented using SMILES notation, ensuring compatibility with all employed computational platforms.

The compound library was organized into six series based on the N-1 substituent of the isatin core.

tyl—of incrementally increasing chain length and branching, enabling assessment of both chain extension and branching effects on ADMET parameters. The terminal chlorine atom in these substituents was included as a potential pharmacophoric element and to allow future derivatization. Series 5 incorporated a 3-methylbutyl chain, an isosteric non-halogenated analogue of Series 4, included to decouple the contribution of terminal halogenation from chain length effects on lipophilicity and distribution.

Within each series, the first compound represented the non-halogenated parent structure at the C-6 position, while the remaining four analogues were systematically halogenated with fluorine, chlorine, bromine, and iodine, respectively. This halogen walk at C-6 was designed to exploit the well-established bioisosteric relationship between halogens and hydrogen, while simultaneously introducing a graded increase in atomic radius, lipophilicity, and polarizability. Collectively, the combinatorial 6×5 matrix encompassing 30 compounds provided sufficient structural diversity to enable robust quantitative structure-pharmacokinetic relationship analysis across a meaningful range of physicochemical space. The general structure of the investigated isatin derivatives (1H-indole-2,3-dione) and the complete substituent library are presented in Fig. 1.



Series	R1	R2	Compounds
0	-H	-H, -F, -Cl, -Br, -I	1-5
1	-CH ₂ CH ₂ Cl	-H, -F, -Cl, -Br, -I	6-10
2	-(CH ₂) ₃ Cl	-H, -F, -Cl, -Br, -I	11-15
3	-CH ₂ CH(CH ₃)CH ₂ Cl	-H, -F, -Cl, -Br, -I	16-20
4	-(CH ₂) ₄ Cl	-H, -F, -Cl, -Br, -I	21-25
5	-CH ₂ CH ₂ CH(CH ₃) ₂	-H, -F, -Cl, -Br, -I	26-30

Fig. 1. General structure and combinatorial substituent matrix of the studied 1H-indole-2,3-dione (isatin) derivatives (6×5 library; $n = 30$ compounds).

Series 0 comprised the non-N-alkylated parent structures, retaining the -NH functionality and serving as the pharmacokinetic reference group. The rationale for N-alkylation in the remaining series was twofold: first, to progressively increase lipophilicity and steric bulk at the N-1 position in a controlled manner; and second, to eliminate the -NH hydrogen bond donor, thereby reducing polarity and TPSA. Series 1–4 incorporated haloalkyl chains—2-chloroethyl, 3-chloropropyl, 2-methyl-3-chloropropyl, and 4-chlorobu-

Computational Tools

Physicochemical and pharmacokinetic parameters were predicted using two complementary and widely validated *in silico* platforms: SwissADME and ADMETlab 2.0 (27,28). The combined use of both tools was deliberate, as the two platforms employ distinct underlying models and training datasets, thereby providing greater confidence in predictions where

results converge and enabling identification of potential uncertainties where they diverge.

SwissADME (Swiss Institute of Bioinformatics, Lausanne, Switzerland) is a freely accessible web tool that employs established physicochemical calculation algorithms and machine learning-based models for ADME property prediction. In the present study, SwissADME was used to calculate molecular weight (MW), consensus logP (cLogP, computed as the arithmetic mean of five independent logP prediction algorithms: iLOGP, XLOGP3, WLOGP, MLOGP, and SILICOS-IT), topological polar surface area (TPSA), number of hydrogen bond acceptors (NHA), number of hydrogen bond donors (NHD), and compliance with Lipinski's Rule of Five. The use of consensus logP was preferred over any single algorithm, as it reduces the influence of model-specific systematic errors and has been shown to provide more robust lipophilicity estimates across structurally diverse compound sets.

ADMETlab 2.0 is a comprehensive deep learning-based platform trained on large curated experimental datasets, designed to predict a broad spectrum of ADMET endpoints. In the present study, ADMETlab 2.0 was used to predict the following parameters: volume of distribution at steady state (VDss, log L/kg), systemic clearance (CL, mL/min/kg), plasma half-life ($t_{1/2}$, h), plasma protein binding (PPB, %), human intestinal absorption (HIA, %), Caco-2 permeability (log cm/s), P-glycoprotein (P-gp) substrate and inhibitor status, blood-brain barrier (BBB) permeability, aqueous solubility (LogS), and CYP isoform interaction profiles (CYP1A2, CYP2C9, CYP2C19, CYP2D6, and CYP3A4 substrate and inhibitor status). The predictive performance of ADMETlab 2.0 has been benchmarked across multiple independent test sets and shown to achieve satisfactory accuracy for the endpoints employed in this study.

Data Analysis

All predicted pharmacokinetic and physicochemical data were compiled in Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) and subsequently imported into IBM SPSS Statistics (version 27.0) for statistical analysis.

Prior to correlation analysis, the distributional properties of all continuous variables were assessed

using the Shapiro-Wilk test for normality ($n=30$). Variables exhibiting significant departures from normality were noted; however, given that Pearson's correlation coefficient is relatively robust to moderate non-normality at sample sizes above 25, bivariate correlations were computed using Pearson's r for all variable pairs to maintain consistency of interpretation (29). Pearson's r was calculated to quantify the strength and direction of linear associations between logP and each of the pharmacokinetic parameters (VDss, CL, $t_{1/2}$, PPB), as well as between selected pharmacokinetic parameters (CL- $t_{1/2}$ and VDss- $t_{1/2}$). The coefficient of determination ($R^2 = r^2$) was computed for each pair to express the proportion of variance in the dependent variable explained by the predictor. Statistical significance was evaluated using a two-tailed t-test with $n-2$ degrees of freedom, applying a significance threshold of $p < 0.05$.

To investigate the combined influence of distribution and elimination on half-life, a standard multiple linear regression (MLR) model was constructed with $t_{1/2}$ as the dependent variable and VDss and CL as independent predictors (30). Prior to model construction, the assumption of no excessive multicollinearity between the two predictors was verified by computing the variance inflation factor (VIF); VIF values below 5 were considered acceptable. Model fit was assessed using the coefficient of determination (R^2) and the adjusted R^2 , the latter correcting for the number of predictors included. Overall model significance was evaluated by ANOVA F-test. Individual predictor contributions were assessed by the standardized regression coefficients (β) and their associated p -values, with significance defined at $p < 0.05$. Regression diagnostics were performed to verify the key assumptions of the MLR model, including inspection of residual plots for linearity and homoscedasticity, and examination of the normal probability plot of standardized residuals for distributional adequacy.

RESULTS AND DISCUSSION

Physicochemical and Pharmacokinetic Properties

The predicted physicochemical and ADMET properties of the 30 investigated isatin derivatives are presented in Table 1 and Table 2. The compound library encompasses six structural series defined by

the N-1 substituent (Series 0-5), each containing five members differing in the nature of the C-6 halogen substituent (H, F, Cl, Br, I). This factorial design en-

ables the independent assessment of N-1 alkylation effects and C-6 halogenation effects on the pharmacokinetic profile of the isatin scaffold. The discussion

Table 1. Physicochemical and absorption-related properties of the investigated isatin derivatives.

Nº	C6	N1	MW	NHA	NHD	TPSA	logP	Lipinski	Caco2 permeability	P-gp interaction	HIA, %
1	H	-H	147.03	3	1	46.17	0.768	yes	-4.316	I	<30
2	F	-H	165.02	3	1	46.17	1.138	yes	-4.402	I	<30
3	Cl	-H	180.99	3	1	46.17	1.817	yes	-4.318	I	<30
4	Br	-H	224.94	3	1	46.17	1.953	yes	-4.708	I	<30
5	I	-H	272.93	3	1	46.17	2.211	yes	-4.463	I	<30
6	H	2-chloroethyl	209.02	3	0	37.38	1.301	yes	-4.234	I	<30
7	F	2-chloroethyl	227.01	3	0	37.38	1.482	yes	-4.392	I	<30
8	Cl	2-chloroethyl	242.99	3	0	37.38	2.177	yes	-4.369	I	<30
9	Br	2-chloroethyl	286.93	3	0	37.38	2.339	yes	-4.632	I	<30
10	I	2-chloroethyl	334.92	3	0	37.38	2.672	yes	-4.403	I	<30
11	H	3-chloropropyl	223.04	3	0	37.38	1.963	yes	-4.494	I	<30
12	F	3-chloropropyl	241.03	3	0	37.38	1.951	yes	-4.516	I	<30
13	Cl	3-chloropropyl	257.0	3	0	37.38	2.555	yes	-4.542	I	<30
14	Br	3-chloropropyl	300.95	3	0	37.38	2.699	yes	-4.673	I	<30
15	I	3-chloropropyl	348.94	3	0	37.38	3.040	yes	-4.571	I	<30
16	H	2-methyl-3-chloro-propyl	236.06	3	0	37.38	2.511	yes	-4.65	I	<30
17	F	2-methyl-3-chloro-propyl	255.05	3	0	37.38	2.600	yes	-4.659	I	<30
18	Cl	2-methyl-3-chloro-propyl	271.02	3	0	37.38	3.031	yes	-4.637	I	<30
19	Br	2-methyl-3-chloro-propyl	314.97	3	0	37.38	3.154	yes	-4.709	I	<30
20	I	2-methyl-3-chloro-propyl	362.95	3	0	37.38	3.369	yes	-4.678	I	<30
21	H	4-chloro-butyl	237.06	3	0	37.38	2.299	yes	-4.629	I	<30
22	F	4-chloro-butyl	255.05	3	0	37.38	2.226	yes	-4.568	I	<30
23	Cl	4-chloro-butyl	271.02	3	0	37.38	2.965	yes	-4.605	I	<30
24	Br	4-chloro-butyl	314.97	3	0	37.38	3.150	yes	-4.703	I	<30
25	I	4-chloro-butyl	362.95	3	0	37.38	3.393	yes	-4.644	I	<30
26	H	3-methyl-butyl	217.11	3	0	37.38	2.965	yes	-4.781	I	<30
27	F	3-methyl-butyl	235.1	3	0	37.38	3.099	yes	-4.722	I	<30
28	Cl	3-methyl-butyl	251.07	3	0	37.38	3.492	yes	-4.754	I	<30
29	Br	3-methyl-butyl	295.02	3	0	37.38	3.641	yes	-4.734	I	<30
30	I	3-methyl-butyl	343.01	3	0	37.38	3.859	yes	-4.777	I	<30

below addresses physicochemical drug-likeness, absorption-related parameters, distribution and plasma binding, and elimination characteristics in turn.

Molecular weights ranged from 147.13 g/mol (Compound 1, the non-substituted parent isatin) to 362.95 g/mol (compounds 25 and 30, the iodo-substituted members of Series 4 and 5), reflecting the pro-

Table 2. Predicted distribution, metabolic, and elimination parameters of the investigated isatin derivatives.

Nº	C6	N1	PPB, %	VD _{ss}	BBB	1A2	2C19	2C9	2D6	3A4	Cl	t _{1/2}
1	H	H	95.00	-0.269	-	D	N	N	N	N	6.944	1.739
2	F	H	97.8	-0.146	-	D	N	N	N	S	4.971	1.205
3	Cl	H	98.7	-0.227	-	D	N	N	N	S	4.172	1.053
4	Br	H	98.4	0.03	-	I	N	N	N	N	4.869	1.157
5	I	H	98.2	0.163	-	D	N	N	N	S	3.215	0.887
6	H	2-chloroethyl	86.7	-0.017	-	D	N	I	I	I	7.183	1.157
7	F	2-chloroethyl	96.1	-0.048	-	D	D	I	N	I	6.028	0.899
8	Cl	2-chloroethyl	98.2	-0.141	-	D	S	N	N	S	5.233	0.788
9	Br	2-chloroethyl	97.8	0.008	-	D	I	I	I	I	5.906	0.944
10	I	2-chloroethyl	97.8	0.155	-	D	S	S	I	S	4.171	0.677
11	H	3-chloropropyl	95.7	-0.003	-	D	I	I	I	I	8.580	0.874
12	F	3-chloropropyl	98.3	-0.028	-	D	S	I	N	I	7.764	0.724
13	Cl	3-chloropropyl	98.6	-0.069	-	D	D	N	N	D	6.985	0.573
14	Br	3-chloropropyl	98.4	0.041	-	D	I	I	I	I	7.466	0.791
15	I	3-chloropropyl	98.2	0.197	-	D	S	D	I	D	5.512	0.499
16	H	2-methyl-3-chloro-propyl	92.7	0.118	-	S	S	I	N	D	7.929	0.556
17	F	2-methyl-3-chloro-propyl	98.3	0.054	-	S	S	I	N	D	7.499	0.493
18	Cl	2-methyl-3-chloro-propyl	98.9	-0.030	-	D	S	N	N	D	6.882	0.387
19	Br	2-methyl-3-chloro-propyl	98.2	0.120	-	D	D	I	N	D	7.342	0.590
20	I	2-methyl-3-chloro-propyl	97.9	0.284	-	D	S	I	N	D	5.725	0.441
21	H	4-chloro-butyl	98.5	-0.113	-	D	I	I	I	I	6.959	0.346
22	F	4-chloro-butyl	98.6	0.084	-	D	I	I	N	D	6.802	0.264
23	Cl	4-chloro-butyl	98.8	-0.018	-	D	I	S	N	D	6.515	0.279
24	Br	4-chloro-butyl	98.7	0.123	-	D	I	I	I	I	6.714	0.337
25	I	4-chloro-butyl	98.5	0.332	-	D	D	D	I	D	5.519	0.362
26	H	3-methyl-butyl	98.4	0.229	-	D	D	D	I	I	8.769	0.462
27	F	3-methyl-butyl	98.6	0.242	-	D	D	D	I	I	8.108	0.296
28	Cl	3-methyl-butyl	99.0	0.245	-	D	D	D	N	D	7.587	0.257
29	Br	3-methyl-butyl	98.6	0.283	-	I	D	D	I	I	8.067	0.406
30	I	3-methyl-butyl	98.4	0.572	-	D	D	D	I	D	6.473	0.352

gressive increase in molecular mass introduced by both N-alkylation and C-6 halogenation. The non-alkylated Series 0 members (compounds 1–5) have the lowest MW values (147–273 g/mol), while the N-alkylated series ranged from approximately 209 to 363 g/mol. All 30 compounds complied with Lipinski's Rule of Five, confirming that N-alkylation and halogenation as applied in this library do not compromise fundamental drug-likeness criteria. Notably, all N-alkylated derivatives lack an -NH hydrogen bond donor (NHD = 0), in contrast to the Series 0 parent compounds (NHD = 1), a structural change that reduces TPSA from 46.17 Å² in Series 0 to 37.38 Å² in Series 1–5. This TPSA reduction upon N-alkylation is pharmacokinetically meaningful, as it places all N-alkylated derivatives comfortably below the 60 Å² threshold associated with optimal passive transcellular permeability, while the Series 0 members, although also below this threshold, retain a marginally higher polar character.

Lipophilicity (logP) values increased monotonically within each series as a function of C-6 halogenation, progressing from the non-halogenated parent to the iodo-substituted analogue, consistent with the known order of lipophilic contribution of halogens (F < Cl < Br < I). Across series, logP also increased with N-1 alkyl chain length and branching, with Series 0 showing the lowest values (0.768–2.211) and Series 5 the highest (2.965–3.859). Notably, the iodo-substituted member of Series 5 (Compound 30) exhibited the highest logP in the library (3.859), while the non-substituted parent isatin (Compound 1) had the lowest (0.768). This systematic variation in logP across the 6 × 5 matrix confirms that the compound design effectively covers a broad and continuous lipophilicity range, which is a prerequisite for meaningful quantitative structure-pharmacokinetic correlation analysis. All logP values remained below 5, consistent with Lipinski's upper limit, indicating that lipophilicity does not constitute a drug-likeness liability for any member of the series. Topological polar surface area (TPSA) was bimodally distributed, taking values of 46.17 Å² for Series 0 and 37.38 Å² for Series 1–5, both well below the 90 Å² threshold generally associated with adequate oral absorption. The narrow and discrete TPSA distribution reflects the limited structural diversity in the polar region of these compounds and suggests that TPSA is unlikely to be

a useful discriminator of pharmacokinetic behavior within this series.

Predicted Caco-2 permeability values ranged from approximately -4.78 to -4.23 log cm/s across the series, indicating moderate intestinal permeability for all compounds. The values showed a modest inverse trend with logP, with the most lipophilic compounds (Series 5, iodine-substituted) exhibiting slightly lower Caco-2 scores, potentially reflecting an upper-lipophilicity ceiling effect where increased partitioning into the lipid bilayer reduces net transcellular flux. All 30 compounds were predicted to act as P-glycoprotein (P-gp) inhibitors rather than substrates, a finding consistent with the presence of aromatic ring systems and lipophilic N-substituents that are characteristic features of many reported P-gp inhibitors. P-gp inhibitory activity may have dual implications: it could facilitate the intestinal absorption of co-administered P-gp substrates, but it may also raise concerns regarding drug-drug interaction potential that would require evaluation in further studies. Human intestinal absorption (HIA) was predicted to be below 30% for all derivatives, suggesting limited passive oral absorption despite the generally favorable Caco-2 permeability values. This apparent discrepancy is pharmacokinetically consistent with dissolution-limited absorption: the relatively low aqueous solubility (LogS) predicted for these moderately lipophilic compounds constrains the amount of dissolved drug available for intestinal uptake, irrespective of membrane permeability. This biopharmaceutical bottleneck represents a key optimization target for future analogues within this scaffold class.

Plasma protein binding (PPB) was consistently high across the series, ranging from 86.7% (compound 6, the non-halogenated N-(2-chloroethyl) derivative) to 99.0% (compound 28, the 6-chloro-N-(3-methylbutyl) derivative). The lowest PPB value across the entire library was observed for compound 6 (86.7%), the non-halogenated member of Series 1, which represents a notable outlier within the N-alkylated subseries. Among the non-alkylated Series 0 compounds, PPB ranged from 95.0% (Compound 1) to 98.7% (compound 3), values that are not markedly lower than those of the N-alkylated series. The anomalously low PPB of compound 6 does not therefore reflect a systematic Series 0 effect, but rather a compound-specific deviation whose structural basis

warrants further investigation. Collectively, the PPB data do not support a consistent trend of reduced protein binding in the absence of N-alkyl substitution. The high and narrowly distributed PPB values across most of the series indicate that the free plasma fraction (f_u) is consistently low (typically 1-5%), which has important pharmacodynamic implications: only the unbound fraction is available to interact with bacterial targets, and therefore *in vivo* efficacy will depend on achieving adequate free drug concentrations at the site of infection. The predicted VDss showed considerably greater variability than PPB, ranging from $-0.269 \log \text{L/kg}$ (Compound 1) to $0.572 \log \text{L/kg}$ (Compound 30), corresponding to approximately 0.54 to 3.73 L/kg in absolute terms. Negative $\log \text{VDss}$ values (i.e., $\text{VDss} < 1 \text{ L/kg}$) were most prevalent among Series 0 and the lower-lipophilicity members of Series 1–2, indicating predominant distribution within the plasma and extracellular water compartments. Positive $\log \text{VDss}$ values ($\text{VDss} > 1 \text{ L/kg}$), suggesting more extensive tissue distribution, were associated with the higher-lipophilicity members of Series 3–5, particularly the bromo- and iodo-substituted analogues. This pattern is consistent with lipophilicity-driven tissue partitioning and directly supports the correlation analysis presented in the following section.

Predicted systemic CL values ranged from 3.215 mL/min/kg (Compound 5, the 6-iodo Series 0 member) to 8.769 mL/min/kg (Compound 26, the non-halogenated N-(3-methylbutyl) derivative), consistent with moderate to high hepatic clearance. For reference, hepatic blood flow in humans is approximately 21 mL/min/kg, placing most of the studied derivatives in the intermediate-to-high extraction ratio range, which implies significant first-pass metabolism and reinforces the limited predicted oral bioavailability noted above. The elimination half-life ($t_{1/2}$) was notably short across the series, varying from 0.257 h (Compound 28) to 1.739 h (Compound 1). The longest $t_{1/2}$ values were observed in Series 0, which also exhibited the lowest $\log P$ values and lowest VDss, suggesting that limited tissue distribution contributes to prolonged systemic residence in this subseries. Conversely, the most lipophilic compounds (Series 5) showed the shortest $t_{1/2}$ despite their higher VDss values, a finding that is consistent with the pharmacokinetic relationship $t_{1/2} = (0.693 \times \text{VDss}) /$

CL when considered in conjunction with the concurrent increase in CL observed across the series. Although CL does not correlate significantly with $\log P$ at the level of the full dataset ($r = +0.213$, $p = 0.258$), the absolute CL values for Series 5 members (6.5–8.8 mL/min/kg) are among the highest in the library, and the proportional increase in CL relative to VDss in this subseries is sufficient to prevent the expected prolongation of $t_{1/2}$. The net result is that the increase in distribution volume is not accompanied by a commensurate extension of systemic exposure, so that $t_{1/2}$ remains short or decreases despite higher VDss—a pattern governed by the ratio VDss/CL rather than by either parameter alone.

Regarding metabolic liability, the predicted CYP interaction profiles were heterogeneous across the series, reflecting the structural diversity introduced by N-1 alkylation and C-6 halogenation. CYP1A2 was the most consistently implicated isoform, with the majority of compounds predicted as substrates or inhibitors, consistent with the planar aromatic indolone core which is a known structural feature associated with CYP1A2 recognition. CYP2C19 and CYP3A4 interactions were variable and appeared to correlate broadly with increasing N-1 alkyl chain length: the more lipophilic Series 3–5 members showed a higher frequency of predicted substrate or inhibitor activity toward these isoforms, which represent the principal drug-metabolizing enzymes in the liver. CYP2C9 and CYP2D6 interactions were less consistently predicted and did not follow a clear structural trend, suggesting that these interactions may be governed by specific binding pocket complementarity rather than lipophilicity alone. The heterogeneous CYP profiles observed across the series are consistent with the absence of a significant $\log P$ -CL correlation identified in the statistical analysis, supporting the conclusion that metabolic clearance in this series is governed primarily by structure-specific enzyme recognition rather than by bulk hydrophobicity. All 30 compounds were predicted to be BBB-impermeable, as ADMETlab 2.0 classified each derivative as a BBB non-penetrant—a finding consistent with the combination of moderate molecular weight, moderate-to-high lipophilicity, and predicted P-gp inhibitory activity across the series. Although BBB penetration is not a target for antibacterial agents addressing systemic infections, the

absence of predicted CNS penetration is a favorable safety attribute that reduces the risk of centrally mediated adverse effects.

Correlation Analysis

Pearson bivariate correlation analysis was performed between the primary lipophilicity descriptor (logP) and four pharmacokinetic parameters (PPB, VD_{ss}, CL, and $t_{1/2}$), as well as between selected inter-parameter pairs (VD_{ss}- $t_{1/2}$ and CL- $t_{1/2}$). A significance threshold of $p < 0.05$ (two-tailed, $df = 28$) was applied throughout. The complete correlation matrix is presented in Table 3.

Table 3. Pearson bivariate correlation matrix for logP, PPB, VD_{ss}, CL, and $t_{1/2}$ across the 30 investigated isatin derivatives ($n = 30$, $df = 28$, two-tailed).

Variable Pair	<i>r</i>	<i>R</i> ²	<i>t</i> (28)	<i>p</i>
logP-PPB	+0.480	0.231	+2.897	0.007
logP-VD _{ss}	+0.810	0.656	+7.307	< 0.001
logP-CL	+0.213	0.046	+1.155	0.258 (ns)
logP-$t_{1/2}$	-0.843	0.711	-8.300	< 0.001
VD _{ss} - $t_{1/2}$	-0.621	0.386	-4.191	< 0.001
CL- $t_{1/2}$	-0.312	0.097	-1.736	0.093 (ns)
PPB-CL	-0.171	0.029	-0.919	0.366 (ns)

ns = not significant ($p \geq 0.05$, two-tailed, $df = 28$). Bold/highlighted row indicates the strongest bivariate association in the dataset.

A moderate and statistically significant positive correlation was found between logP and plasma protein binding ($r = +0.480$, $R^2 = 0.231$, $t(28) = 2.897$, $p = 0.007$; Fig. 2). Lipophilicity accounts for 23.1% of the variance in PPB, indicating that hydrophobic interactions with serum albumin and other plasma proteins make a meaningful but partial contribution to binding. The remaining variance reflects additional structural and electronic determinants—including hydrogen-bond donor/acceptor capacity and molecular size—that govern complementarity with protein binding sites independently of bulk lipophilicity. The observed correlation strength is likely an underestimate of the true logP-PPB relationship: the uniformly high PPB values across the series (86.7–99.0%, predominantly above 95%) impose a near-

ceiling constraint that compresses the detectable dynamic range and attenuates Pearson r .

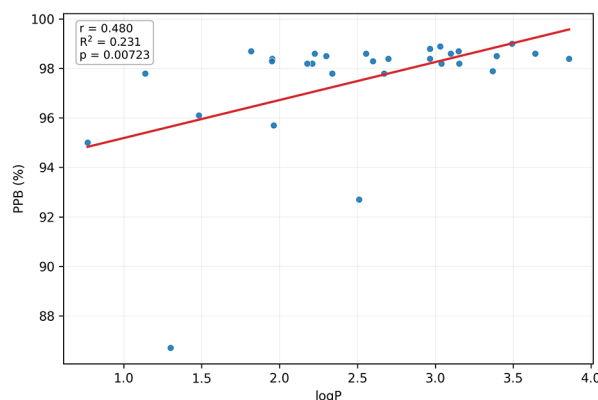


Fig. 2. Scatter plot of plasma protein binding (PPB, %) as a function of logP for the 30 investigated isatin derivatives ($r = +0.480$, $R^2 = 0.231$, $p = 0.007$).

The strongest association involving a physico-chemical descriptor and a distribution parameter was the highly significant positive correlation between logP and VD_{ss} ($r = +0.810$, $R^2 = 0.656$, $t(28) = 7.307$, $p < 0.001$; Fig. 3). Lipophilicity alone explains 65.6% of the variance in volume of distribution, establishing hydrophobicity as the primary structural driver of tissue partitioning within this series. This relationship is mechanistically grounded in the thermodynamic preference of lipophilic compounds for partitioning into tissue lipids, phospholipid membranes, and intracellular compartments, reducing plasma concentration relative to the total body burden and thereby increasing the apparent VD_{ss}. The systematic increase in VD_{ss} from Series 0 to Series 5, closely tracking the increase in logP driven by progressive N-alkylation and C-6 halogenation, provides direct structural evidence for this mechanism operating across the entire compound library. The large t -statistic (7.307) underscores the robustness of the association.

In marked contrast, no significant correlation was detected between logP and systemic clearance ($r = +0.213$, $R^2 = 0.046$, $t(28) = 1.155$, $p = 0.258$). Lipophilicity explains only 4.6% of the variance in CL, demonstrating that hydrophobicity is not a meaningful predictor of metabolic elimination in this series. This is notable because a naive application of classical ADMET theory would predict that more

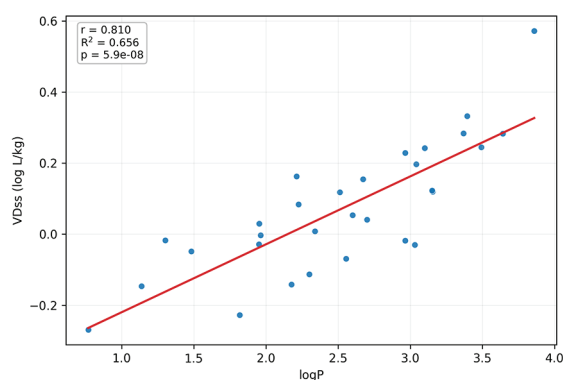


Fig. 3. Correlation between lipophilicity ($\log P$) and volume of distribution at steady state (VD_{ss} , L/kg) for the 30 investigated isatin derivatives ($r = +0.810$, $R^2 = 0.656$, $p < 0.001$).

lipophilic compounds undergo more extensive hepatic metabolism. The absence of this expected relationship implies that CYP-mediated clearance of the studied isatins is governed predominantly by structural complementarity with specific enzyme active sites rather than by passive hydrophobic partitioning—a conclusion consistent with the heterogeneous and structure-specific CYP interaction profiles described in the preceding section. The weakly positive direction of the correlation (rather than the intuitively expected negative sign) reflects the net balance of competing $\log P$ effects on VD_{ss} and CL within the pharmacokinetic equation $t_{1/2} = (0.693 \times VD_{ss})/CL$.

A key finding of this analysis—not highlighted in previous reports on this scaffold class—is the strong negative correlation between $\log P$ and elimination half-life ($r = -0.843$, $R^2 = 0.711$, $t(28) = -8.300$, $p < 0.001$; Fig. 4). This represents the single strongest bivariate association in the entire dataset, outperforming even the $\log P$ - VD_{ss} correlation in explanatory power. Lipophilicity alone accounts for 71.1% of the variance in $t_{1/2}$, directly linking a readily modifiable structural property to the clinically most relevant pharmacokinetic endpoint. The mechanistic basis is indirect: increasing $\log P$ raises VD_{ss} (via the confirmed $\log P$ - VD_{ss} relationship), and the resulting larger distribution volume shortens half-life because clearance does not increase proportionally—as expressed by the pharmacokinetic identity $t_{1/2} = (0.693 \times VD_{ss})/CL$. Practically, this finding imposes a direct constraint on lipophilicity optimization: increasing $\log P$ through N-alkyl chain exten-

sion or C-6 heavy halogenation will systematically shorten predicted systemic exposure, a trade-off that must be explicitly considered in any structure-guided campaign targeting this scaffold.

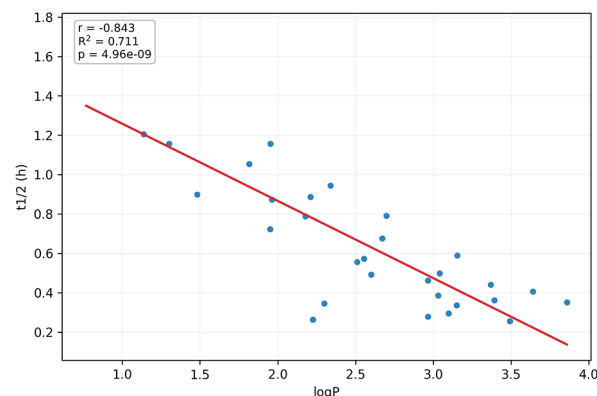


Fig. 4. Correlation between lipophilicity ($\log P$) and elimination half-life ($t_{1/2}$, h) for the 30 investigated isatin derivatives ($r = -0.843$, $R^2 = 0.711$, $p < 0.001$).

Among inter-pharmacokinetic parameter correlations, a significant moderate negative association was observed between VD_{ss} and $t_{1/2}$ ($r = -0.621$, $R^2 = 0.386$, $t(28) = -4.191$, $p < 0.001$). Volume of distribution explains 38.6% of the variance in half-life, confirming that distribution-related processes contribute substantially to elimination kinetics. The negative direction reflects the dominance of the VD_{ss} numerator in the $t_{1/2}$ equation operating at near-constant CL : as VD_{ss} increases, $t_{1/2}$ does not lengthen proportionally because CL also rises modestly with $\log P$, resulting in a net shortening of half-life for more widely distributed compounds. In contrast, the correlation between CL and $t_{1/2}$ was negative but did not reach significance ($r = -0.312$, $R^2 = 0.097$, $p = 0.093$), indicating that clearance variability contributes only marginally to half-life differences across the series. The coefficient of variation of CL across the 30 compounds is 21.0% (mean = 6.53 mL/min/kg, $SD = 1.37$), compared with 48.3% for VD_{ss} expressed in absolute units (mean = 1.29 L/kg, $SD = 0.62$), providing a quantitative explanation for why CL does not emerge as the dominant determinant of $t_{1/2}$ in this dataset. Finally, no significant association was found between PPB and CL ($r = -0.171$, $R^2 = 0.029$, $p = 0.366$). Given that protein binding is uniformly high and near-saturating across the series, the free drug fraction (f_u) is

too narrowly distributed to modulate hepatic extraction in a detectable manner. Taken together, the correlation analysis establishes logP as the central physicochemical axis of pharmacokinetic variability in this isatin library, with strong direct effects on VDss and $t_{1/2}$, and no significant effect on CL.

Multiple Regression Analysis

To investigate the combined influence of distribution and elimination on half-life variability, a standard MLR model was constructed with $t_{1/2}$ (h) as the dependent variable and VDss (log L/kg) and CL (mL/min/kg) as independent predictors. Prior to model construction, the assumption of no excessive multicollinearity was verified: the Pearson correlation between VDss and CL was negligible ($r = 0.132$), yielding a variance inflation factor (VIF = 1.018) well below the commonly applied threshold of 5, confirming that the two predictors carry largely independent information.

The overall model was statistically significant ($F(2,27) = 10.581$, $p = 0.0004$) and explained a moderate proportion of the variance in $t_{1/2}$ ($R^2 = 0.439$, $\text{adj } R^2 = 0.398$), indicating that approximately 44% of the half-life variability across the 30 derivatives is captured by the combination of VDss and CL. The fitted regression equation is: $t_{1/2} = 1.1399 + (-1.163 \times \text{VDss}) + (-0.061 \times \text{CL})$. Full regression coefficients, standard errors, and confidence intervals are presented in Table 4.

a more effective linear proxy for the underlying non-linear distribution-elimination interplay. This does not diminish the mechanistic primacy of VDss as a determinant of $t_{1/2}$; rather, it reflects a scale-dependent limitation of linear regression when applied to log-transformed pharmacokinetic parameters.

Among the two predictors, VDss was the sole statistically significant determinant of $t_{1/2}$ ($B = -1.163$, $SE = 0.287$, $\beta = -0.590$, $t(27) = -4.059$, $p = 0.0004$; 95% CI [-1.751, -0.575]). The negative sign of the VDss coefficient ($B = -1.163$) in the MLR model warrants explicit clarification. In classical pharmacokinetic theory, increasing VDss prolongs $t_{1/2}$ according to the identity $t_{1/2} = (0.693 \times \text{VDss})/\text{CL}$. However, the present MLR operates on independently predicted parameters from ADMETlab 2.0, in which VDss, CL, and $t_{1/2}$ are generated by separate machine learning models that are not constrained to satisfy this pharmacokinetic identity. The negative empirical regression coefficient therefore reflects the statistical covariance structure of the predicted dataset rather than a mechanistic pharmacokinetic relationship: compounds with higher predicted VDss (the more lipophilic Series 3–5 members) were simultaneously predicted with shorter $t_{1/2}$ by the ADMETlab model, and the MLR captures this empirical pattern. This decoupling of predicted parameters from the theoretical pharmacokinetic equation is an inherent limitation of ML-based ADMET platforms and rein-

Table 4. Multiple linear regression coefficients for the model $t_{1/2} = f(\text{VDss}, \text{CL})$.

Predictor	B	SE	β (std)	t(27)	p	95% CI for B
Intercept	1.1399	0.2491	–	4.576	< 0.001	[0.629, 1.651]
VDss	-1.1629	0.2865	-0.590	-4.059	< 0.001	[-1.751, -0.575]
CL	-0.0606	0.0377	-0.234	-1.611	0.119 (ns)	[-0.138, 0.017]

B = unstandardized regression coefficient; *SE* = standard error; β (std) = standardized coefficient; ns = not significant ($p \geq 0.05$). Bold row indicates the statistically significant predictor. Model fit: $R^2 = 0.439$, $\text{adj } R^2 = 0.398$, $F(2,27) = 10.581$, $p = 0.0004$; $VIF = 1.018$ for both predictors.

The apparently lower explanatory power of the MLR model ($R^2 = 0.439$) relative to the simple logP– $t_{1/2}$ correlation ($R^2 = 0.711$) is attributable to the log-linear nature of the VDss scale: because VDss is expressed as log L/kg, the MLR captures only the linear component of the VDss– $t_{1/2}$ relationship, whereas logP—being linearly related to log VDss—acts as

forces the caveat that *in silico* predictions should not be interpreted as internally consistent mechanistic pharmacokinetic models. The standardized coefficient ($\beta = -0.590$) expresses this in units of standard deviation: a one-SD increase in VDss is associated with a 0.59-SD decrease in $t_{1/2}$, confirming that dis-

tribution volume is the dominant pharmacokinetic driver within this MLR framework.

In contrast, CL did not contribute significantly to the model ($B = -0.061$, $SE = 0.038$, $\beta = -0.234$, $t(27) = -1.611$, $p = 0.119$; 95% CI $[-0.138, 0.017]$). The 95% confidence interval for the CL coefficient includes zero, further confirming the lack of a reliable independent CL- $t_{1/2}$ relationship when VDss is held constant. Notably, the standardized coefficient for CL ($\beta = -0.234$) is approximately 2.5-fold smaller than that for VDss ($\beta = -0.590$), quantifying the relative dominance of distribution over elimination in governing half-life in this series.

Residual diagnostics confirmed adequate model assumptions. The Shapiro-Wilk test applied to the standardized residuals showed no significant departure from normality ($W = 0.9817$, $p = 0.869$), and inspection of the residual plot indicated approximate homoscedasticity across the predicted range. The largest residual was observed for Compound 1 (the non-substituted parent isatin, $t_{1/2} = 1.739$ h), which represents a structural outlier relative to the N-alkylated series and is expected to exhibit pharmacokinetic behavior less well captured by a model based primarily on lipophilicity-driven distribution. Taken together, the MLR results confirm that distribution volume is the principal determinant of half-life variability in this isatin library, accounting for substantially greater variance than clearance when both predictors are considered simultaneously.

Mechanistic Insights

The totality of the statistical evidence converges on a coherent mechanistic picture: lipophilicity ($\log P$) is the central physicochemical axis governing the pharmacokinetic behavior of these isatin derivatives, exerting its influence primarily through distribution rather than elimination. The $\log P$ -VDss correlation ($r = +0.810$, $R^2 = 0.656$) demonstrates that increasing N-alkyl chain length and C-6 halogenation progressively drive tissue partitioning by enhancing the thermodynamic affinity of compounds for phospholipid membranes and intracellular lipid compartments. This lipophilicity-driven tissue distribution then propagates downstream: the confirmed $\log P$ - $t_{1/2}$ association ($r = -0.843$, $R^2 = 0.711$)—the strongest bivariate relationship in the dataset—is mechanistically explained as an indirect effect me-

diated through VDss. As $\log P$ increases, VDss increases, and because CL does not rise proportionally, the pharmacokinetic identity $t_{1/2} = (0.693 \times \text{VDss})/\text{CL}$ predicts a net shortening of half-life. This indirect path— $\log P \rightarrow \text{VDss} \rightarrow t_{1/2}$ —is quantitatively confirmed by the MLR model, in which VDss (standardized $\beta = -0.590$, $p < 0.001$) dominates over CL ($\beta = -0.234$, $p = 0.119$) as a determinant of $t_{1/2}$.

The $\log P$ -PPB correlation ($r = +0.480$, $R^2 = 0.231$) indicates that hydrophobic interactions make a meaningful contribution to plasma protein binding, most likely through binding to the hydrophobic pocket of serum albumin (site II), which preferentially accommodates lipophilic aromatic compounds of the type represented by the isatin scaffold. However, the near-ceiling PPB values across most of the series (predominantly 95–99%) indicate that protein binding is approaching saturation at moderate-to-high $\log P$ values, creating a biochemical ceiling that limits further PPB increases regardless of additional lipophilicity gains. This saturation effect mechanistically explains the attenuation of the $\log P$ -PPB correlation relative to $\log P$ -VDss: while both parameters respond to lipophilicity, VDss remains responsive across the full $\log P$ range of the series (0.768–3.859), whereas PPB is effectively constrained above approximately $\log P > 1.5$.

The absence of significant $\log P$ -CL ($r = +0.213$, $p = 0.258$) and PPB-CL ($r = -0.171$, $p = 0.366$) correlations reveals that metabolic elimination in this series is decoupled from bulk physicochemical properties. This behavior is mechanistically consistent with the heterogeneous CYP interaction profiles predicted across the library: the dominant involvement of CYP1A2, driven by the planar aromatic indolone core present in all compounds, imposes a structurally invariant metabolic component that does not scale with lipophilicity. Meanwhile, the variable engagement of CYP2C19 and CYP3A4 in more lipophilic series members likely reflects substrate recognition through shape complementarity rather than hydrophobic surface area. Notably, the weakly positive sign of the $\log P$ -CL correlation—counterintuitive under the classical model where lipophilic compounds are cleared more rapidly—may reflect a compensatory effect: in the most lipophilic compounds, high VDss reduces the hepatic drug concentration avail-

able for metabolism, partially offsetting any lipophilicity-driven increase in intrinsic clearance.

Taken together, the data support a distribution-centric pharmacokinetic model for this isatin scaffold: logP is the primary structural handle controlling both tissue partitioning and, indirectly, systemic exposure duration, while metabolic clearance operates as a secondary, enzyme-recognition-governed process largely independent of lipophilicity. This mechanistic framework has direct implications for scaffold optimization, as discussed in the following section.

Drug Design Implications

The present findings provide a quantitative framework for the structure-guided optimization of isatin-based antibacterial candidates. The most actionable recommendation concerns lipophilicity management. Given the strong logP–VDss relationship ($R^2 = 0.656$) and the even stronger inverse logP– $t_{1/2}$ association ($R^2 = 0.711$), lipophilicity functions simultaneously as a driver of tissue distribution and as a determinant of systemic exposure duration. The data indicate that an optimal logP window of approximately 1.5–2.5 would balance adequate tissue partitioning (VDss > 1 L/kg, favoring tissue-over-plasma distribution) with a sufficiently prolonged half-life to allow clinically relevant dosing intervals. Within the present series, this window corresponds broadly to the fluorine- and chlorine-substituted members of Series 1–3 (compounds 7–9, 12–14, 17–18), which combine moderate logP with the most favorable predicted $t_{1/2}$ values in the N-alkylated subseries. Structural strategies to remain within this window include: (i) preferring shorter N-alkyl chains (2-chloroethyl or 3-chloropropyl over 4-chlorobutyl or 3-methylbutyl); and (ii) selecting fluorine or chlorine at C-6 rather than bromine or iodine, which push logP above 3.0 and markedly shorten $t_{1/2}$.

Oral bioavailability represents the principal biopharmaceutical challenge for this scaffold. All 30 compounds were predicted with HIA below 30%, a limitation attributed to dissolution-limited absorption rather than inadequate membrane permeability. Improving aqueous solubility is therefore the priority strategy for enhancing oral exposure. Practical approaches include: introduction of ionizable groups at the N-1 position—for example, short aminoal-

kyl chains (dimethylaminoethyl, morpholinoethyl) which increase solubility while preserving moderate lipophilicity; replacement of terminal halogens in Series 1–4 with polar bioisosteres such as hydroxymethyl or methoxyethyl groups; and co-crystal or salt formation strategies for the neutral, poorly soluble members of Series 0. Any solubility-improving modification should be evaluated against its effect on logP to ensure that the target logP window of 1.5–2.5 is maintained.

Metabolic stability is the second major optimization target. The decoupling of CL from logP confirms that lipophilicity reduction alone will not improve metabolic stability in this series; instead, targeted structural changes are required. CYP1A2 engagement, driven by the planar indolone core, may be attenuated by introducing sp^3 carbon centers that reduce planarity—for example, spirocyclic substitution at C-3 or saturated heterocyclic N-1 substituents. CYP2C19 and CYP3A4 liability in the more lipophilic series members may be addressed by reducing the overall hydrophobic surface area of the N-1 substituent or by introducing fluorine at metabolically labile positions on the alkyl chain to block oxidative metabolism. Importantly, because CYP interaction profiles are predicted to be heterogeneous across the series, metabolic liability should be evaluated compound-by-compound in follow-up *in vitro* microsomal stability studies, rather than addressed through a single series-wide structural change.

High and near-saturating plasma protein binding (86.7–99.0%) is a structural property of this scaffold that is unlikely to be substantially reduced without major changes to the isatin core. The mechanistic analysis indicates that PPB is primarily driven by hydrophobic albumin binding (site II) and is near-saturating at logP > 1.5. Rather than targeting PPB reduction directly, the more productive strategy is to ensure that the free drug fraction—even at $f_u \approx 1$ –5%—is sufficient to achieve the minimum inhibitory concentration at the target site. This requires tight integration of PK predictions with pharmacodynamic data (MIC values against target pathogens) to define free-drug PK/PD indices. Finally, the confirmed P-gp inhibitory activity across the entire series should be monitored carefully in any future *in vivo* or clinical studies, as it may generate drug-drug interaction liability with co-administered P-gp sub-

strates—a particularly relevant concern in the clinical management of multi-drug-resistant bacterial infections where polypharmacy is common.

Limitations

Several limitations of the present study must be acknowledged in interpreting the findings. First, and most fundamentally, all pharmacokinetic and physicochemical parameters were derived exclusively from computational predictions using SwissADME and ADMETlab 2.0. While both platforms employ machine learning models trained on large experimental datasets and have been independently benchmarked, *in silico* predictions remain approximations subject to systematic model biases, particularly for structural classes underrepresented in the training data. The isatin scaffold with simultaneous N-1 alkylation and C-6 halogenation may represent such a case, and the quantitative accuracy of individual parameter predictions—especially VDss and CL, which are notoriously difficult to predict from structure alone—cannot be guaranteed without experimental validation. All conclusions regarding absolute pharmacokinetic values should therefore be interpreted as indicative rather than definitive.

Second, the statistical analyses are constrained by the sample size of $n = 30$. Although this is sufficient to achieve adequate statistical power for the observed effect sizes—the strong correlations ($r > 0.8$) are highly robust even at $n = 30$ —the moderate correlations ($\log P$ –PPB, $r = 0.480$; VDss– $t_{1/2}$, $r = -0.621$) carry wider confidence intervals and their precise magnitude should be interpreted with appropriate caution. Additionally, the entire compound set is confined to a single chemical scaffold varying along only two structural dimensions (N-1 substituent and C-6 halogen). This restricted chemical space, while advantageous for controlling confounding structural variables, limits the generalizability of the quantitative structure-pharmacokinetic relationships identified here to the isatin scaffold class; extrapolation to structurally distinct compound series would require independent validation.

Third, Pearson correlation analysis assumes linear relationships between variables. While linearity is a reasonable approximation for the $\log P$ ranges explored here (0.768–3.859), some pharmacokinetic parameters—particularly PPB, which exhibits

sigmoidal rather than linear dependence on $\log P$ in broader structural datasets—may show non-linear behavior that Pearson r underestimates. The near-ceiling effect observed for PPB in this series is consistent with this concern. Non-linear regression or rank-based correlation methods (Spearman ρ) may provide complementary insights and are recommended for future analyses incorporating a broader $\log P$ range. Furthermore, the MLR model, while satisfying the key statistical assumptions (normality of residuals, absence of multicollinearity), explains only 43.9% of the variance in $t_{1/2}$, indicating that approximately 56% of half-life variability arises from sources not captured by VDss and CL alone—including potentially unmodelled transporter-mediated processes, tissue-specific binding, or model prediction errors in the input parameters.

Fourth, the present analysis does not account for active transport processes, which are known to modulate the distribution and elimination of many drug classes but are poorly captured by current *in silico* ADMET platforms. The predicted P-gp inhibitory activity across all 30 compounds suggests that P-gp-mediated efflux at intestinal, blood-brain barrier, and renal tubular epithelial membranes may significantly influence the actual *in vivo* distribution and elimination profiles relative to the *in silico* predictions. Similarly, the predicted CYP interaction profiles represent qualitative classifications (substrate/inhibitor yes/no) rather than quantitative kinetic parameters, limiting their utility for predicting absolute *in vivo* clearance. These limitations underscore the necessity of experimental follow-up, including *in vitro* microsomal stability assays, Caco-2 bidirectional permeability assays, plasma protein binding measurements by equilibrium dialysis, and ultimately *in vivo* pharmacokinetic studies in appropriate animal models, before any of the present compounds are advanced in the drug discovery pipeline.

CONCLUSION

This study presents the first systematic *in silico* pharmacokinetic analysis of a 6×5 factorial library of 30 novel isatin derivatives designed as potential antibacterial agents, combining Pearson correlation analysis, multiple linear regression, and ADMET profiling via SwissADME and ADMETlab 2.0. All 30 compounds complied with Lipinski's Rule of

Five, confirming the drug-likeness of the scaffold across the full range of N-1 alkyl and C-6 halogen substitution patterns explored.

Lipophilicity (logP) emerged as the central physicochemical determinant of pharmacokinetic variability within this series. A strong positive correlation between logP and VD_{ss} ($r = +0.810$, $R^2 = 0.656$, $p < 0.001$) established that tissue partitioning is primarily driven by hydrophobicity, increasing systematically with N-alkyl chain length and C-6 halogenation. A strong negative correlation between logP and elimination half-life, not previously reported for this scaffold class ($r = -0.843$, $R^2 = 0.711$, $p < 0.001$)—the single strongest bivariate association in the dataset—was identified and mechanistically attributed to the indirect path $\log P \rightarrow \text{VD}_{ss} \rightarrow t_{1/2}$, as expressed by the pharmacokinetic identity $t_{1/2} = (0.693 \times \text{VD}_{ss})/\text{CL}$. A moderate positive correlation between logP and plasma protein binding ($r = +0.480$, $R^2 = 0.231$, $p = 0.007$) was also confirmed, though attenuated by a near-ceiling effect in PPB values across most of the series.

Multiple linear regression confirmed VD_{ss} as the dominant independent predictor of $t_{1/2}$ ($\beta = -0.590$, $p < 0.001$), while systemic clearance did not contribute significantly ($\beta = -0.234$, $p = 0.119$). The absence of significant logP–CL and PPB–CL correlations demonstrates that metabolic elimination in this series is governed by structure-specific enzyme recognition rather than bulk lipophilicity, consistent with the heterogeneous CYP interaction profiles predicted across the library. Collectively, these results support a distribution-centric pharmacokinetic model for the isatin scaffold, in which logP-driven tissue partitioning constitutes the primary determinant of systemic exposure, while clearance operates as a secondary, structurally encoded process.

From a drug design perspective, the findings define a target logP window of approximately 1.5–2.5, within which adequate tissue distribution and clinically acceptable half-life can be simultaneously achieved. Oral bioavailability represents the principal unresolved challenge for this scaffold, attributable to dissolution-limited rather than permeability-limited absorption, and should be addressed through solubility-improving structural modifications rather than further lipophilicity adjustment. The present work establishes a quantitative structure-pharma-

cokinetic relationship framework that can guide the prioritization and structural optimization of isatin-based antibacterial candidates prior to synthesis, and provides a rational basis for the design of next-generation analogues with improved ADMET profiles. Experimental validation through *in vitro* assays and *in vivo* pharmacokinetic studies in appropriate animal models remains an essential next step before advancing any members of this series in the drug discovery pipeline.

REFERENCES

1. Nazir A, Nazir A, Zuhair V, Aman S, Sadiq SUR, Hasan AH, et al. The Global Challenge of Antimicrobial Resistance: Mechanisms, Case Studies, and Mitigation Approaches. *Health Sci Rep*. 2025 Jul 23;8(7):e71077. doi: 10.1002/hsr2.71077.
2. Bertagnolio S, Dobрева Z, Centner CM, Oлару ID, Donà D, Burzo S, et al. WHO global research priorities for antimicrobial resistance in human health. *Lancet Microbe*. 2024 Nov;5(11):100902. doi: 10.1016/S2666-5247(24)00134-4.
3. Zhang Y, Zheng M, Wang Z, Liu Z, Chen S, Li X, et al. Discovery of novel antibacterial agent for the infected wound treatment: all-hydrocarbon stapling optimization of LL-37. *Theranostics*. 2024 Jan 20;14(3):1181-1194. doi: 10.7150/thno.87916.
4. Süßmuth RD, Kulike-Koczula M, Gao P, Kosol S. Fighting Antimicrobial Resistance: Innovative Drugs in Antibacterial Research. *Angew Chem Int Ed Engl*. 2025 Mar 3;64(10):e202414325. doi: 10.1002/anie.202414325.
5. Varun, Sonam, Kakkar R. Isatin and its derivatives: a survey of recent syntheses, reactions, and applications. *Medchemcomm*. 2019 Jan 15;10(3):351-368. doi: 10.1039/c8md00585k.
6. Kumar G, Singh NP, Kumar K. Recent Advancement of Synthesis of Isatins as a Versatile Pharmacophore: A review. *Drug Res (Stuttg)*. 2021 Mar;71(3):115-121. doi: 10.1055/a-1238-2639.
7. Liu B, Jiang D, Hu G. The Antibacterial Activity of Isatin Hybrids. *Curr Top Med Chem*. 2022;22(1):25-40. doi: 10.2174/156802662166621116090456.
8. Saleem M, Ahmad M, Mehmood RF, Shaheen A, Qaiser S, Niaz SI, et al. Synthesis of Isatin Derivatives Exhibiting Antibacterial, Antifungal and Cytotoxic Activities. *Curr Org Synth*. 2022 Aug

- 6;19(6):748-756. doi: 10.2174/1570179419666220128091313.
9. De Moraes Gomes PAT, Pena LJ, Leite ACL. Isatin Derivatives and Their Antiviral Properties Against Arboviruses: A Review. *Mini Rev Med Chem*. 2019;19(1):56-62. doi: 10.2174/1389557518666180424093305.
 10. Cheke RS, Firke SD, Patil RR, Bari SB. ISATIN: New Hope Against Convulsion. *Cent Nerv Syst Agents Med Chem*. 2018;18(2):76-101. doi: 10.2174/1871524917666171113124112.
 11. Ding Z, Zhou M, Zeng C. Recent advances in isatin hybrids as potential anticancer agents. *Arch Pharm (Weinheim)*. 2020 Mar;353(3):e1900367. doi: 10.1002/ardp.201900367.
 12. Sharma PK, Balwani S, Mathur D, Malhotra S, Singh BK, Prasad AK, et al. Synthesis and anti-inflammatory activity evaluation of novel triazolyl-isatin hybrids. *J Enzyme Inhib Med Chem*. 2016 Dec;31(6):1520-6. doi: 10.3109/14756366.2016.1151015.
 13. Cheke RS, Patil VM, Firke SD, Ambhore JP, Ansari IA, Patel HM, et al. Therapeutic Outcomes of Isatin and Its Derivatives against Multiple Diseases: Recent Developments in Drug Discovery. *Pharmaceuticals (Basel)*. 2022 Feb 22;15(3):272. doi: 10.3390/ph15030272.
 14. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? *Acta Pharm Sin B*. 2022 Jul;12(7):3049-3062. doi: 10.1016/j.apsb.2022.02.002.
 15. Burian A, Wagner C, Stanek J, Manafi M, Böhmendorfer M, Jäger W, et al. Plasma protein binding may reduce antimicrobial activity by preventing intra-bacterial uptake of antibiotics, for example clindamycin. *J Antimicrob Chemother*. 2011 Jan;66(1):134-7. doi: 10.1093/jac/dkq400.
 16. Yang L, Sunzel M, Xu P, Edeki T, Wilson D, Li J, et al. Evaluation of the pharmacokinetics and safety of single and multiple ceftaroline fosamil infusions in healthy Chinese and Western subjects. *Int J Clin Pharmacol Ther*. 2015 Aug;53(8):681-91. doi: 10.5414/CP202343.
 17. Wang S, Dong G, Sheng C. Structural simplification: an efficient strategy in lead optimization. *Acta Pharm Sin B*. 2019 Sep;9(5):880-901. doi: 10.1016/j.apsb.2019.05.004.
 18. Wu F, Zhou Y, Li L, Shen X, Chen G, Wang X, et al. Computational Approaches in Preclinical Studies on Drug Discovery and Development. *Front Chem*. 2020 Sep 11;8:726. doi: 10.3389/fchem.2020.00726.
 19. Venkataraman M, Rao GC, Madavareddi JK, Madadi SR. Leveraging machine learning models in evaluating ADMET properties for drug discovery and development. *ADMET DMPK*. 2025 Jun 7;13(3):2772. doi: 10.5599/admet.2772.
 20. Poroikov VV. Komp'uternoe konstruirovaniye lekarstv: ot poiska novykh farmakologicheskikh veshchestv do sistemnoi farmakologii [Computer-aided drug design: from discovery of novel pharmaceutical agents to systems pharmacology]. *Biomed Khim*. 2020 Jan;66(1):30-41. Russian. doi: 10.18097/PBMC20206601030.
 21. Marques L, Costa B, Pereira M, Silva A, Santos J, Saldanha L, et al. Advancing Precision Medicine: A Review of Innovative In Silico Approaches for Drug Development, Clinical Pharmacology and Personalized Healthcare. *Pharmaceutics*. 2024 Feb 27;16(3):332. doi: 10.3390/pharmaceutics16030332.
 22. Yongye AB, Medina-Franco JL. Systematic characterization of structure-activity relationships and ADMET compliance: a case study. *Drug Discov Today*. 2013 Aug;18(15-16):732-9. doi: 10.1016/j.drudis.2013.04.002.
 23. Othman B, Beigh S, Albanghali MA, Sindi AAA, Shanawaz MA, Ibahim MAEM, et al. Comprehensive pharmacokinetic profiling and molecular docking analysis of natural bioactive compounds targeting oncogenic biomarkers in breast cancer. *Sci Rep*. 2025 Feb 13;15(1):5426. doi: 10.1038/s41598-024-84401-4.
 24. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. 2002 Jun 6;45(12):2615-23. doi: 10.1021/jm020017n.
 25. Holt K, Nagar S, Korzekwa K. Methods to Predict Volume of Distribution. *Curr Pharmacol Rep*. 2019 Oct;5(5):391-399. doi: 10.1007/s40495-019-00186-5.
 26. Wu J, Lorusso PM, Matherly LH, Li J. Implications of plasma protein binding for pharmacokinetics and pharmacodynamics of the γ -secretase inhibitor RO4929097. *Clin Cancer Res*. 2012 Apr 1;18(7):2066-79. doi: 10.1158/1078-0432.CCR-11-2684.
 27. 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 Mar 3;7:42717. doi: 10.1038/srep42717.
28. Xiong G, Wu Z, Yi J, Fu L, Yang Z, Hsieh C, et al. ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res.* 2021 Jul 2;49(W1):W5-W14. doi: 10.1093/nar/gkab255.
29. Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med.* 2018 Aug 7;18(3):91-3. doi: 10.1016/j.tjem.2018.08.001.
30. Olsen AA, McLaughlin JE, Harpe SE. Using multiple linear regression in pharmacy education scholarship. *Curr Pharm Teach Learn.* 2020 Oct;12(10):1258-68. doi: 10.1016/j.cptl.2020.05.017.