TECHNION - ISRAEL INSTITUTE OF TECHNOLOGY

COMPUTATIONAL BIOLOGY (134141)
WINTER 2025-2026

Atom-Level Modulation of LogP for Analysis of Molecular Interaction in Drug Molecules

Authors:
Ofir Simhi (315908863)

February 6, 2026

Introduction

A central challenge in drug discovery is relating molecular structure to physicochemical properties that govern biological performance (e.g., solubility and membrane permeability). Because experimentally measuring such properties at scale is time-consuming and costly, computational approaches that can rapidly estimate key descriptors are valuable during early-stage molecular design.

A key descriptor is the octanol–water partition coefficient, $\log P$, which quantifies lipophilicity:

$$\log P = \log\left(\frac{C_{\text{oil}}}{C_{\text{water}}}\right)$$

Lipophilicity influences passive membrane permeability and is inversely related to aqueous solubility [1]. Consequently, drug-design heuristics such as Lipinski's Rule of Five treat $\log P$ as a practical constraint for oral bioavailability [2, 3].

Modern machine-learning and deep-learning models can predict molecular properties directly from structure with high accuracy [4], yet their predictions are often difficult to interpret. For $\log P$, interpretability is particularly important because lipophilicity is an emergent effect of context-dependent interactions with the environment: the same functional group can contribute differently depending on its molecular neighborhood.

In this work, we structure the prediction around an additive, atom-decomposed baseline to retain a transparent reference point. Rather than treating baseline error as noise, we treat it as structured signal: systematic, context-dependent deviations that can be learned.

This framing positions the work between classical fragment models (interpretable but rigid) and end-to-end graph neural networks (accurate but often opaque). We ask whether learning deviations from a chemically grounded baseline can yield corrections that are more interpretable than direct, end-to-end prediction while still improving $\log P$ accuracy. Specifically, we test the hypothesis that the learned context-dependent corrections are (i) spatially localized to specific functional regions, and (ii) chemically structured (e.g., concentrated around polar groups and their local environments) rather than diffuse or random.

Methods

Wildman–Crippen Atomic Fragment Model

As a reference for molecular lipophilicity, this work uses the Wildman–Crippen (WC) atomic fragment method to predict $\log P$ [5]. The model approximates lipophilicity as a sum of fixed, atom-typed contributions fitted to experimental data:

$$\log P_{\text{WC}} = \sum_{i=1}^{N} c_i,$$

where c_i is the contribution for atom i (assigned by an atom-typing scheme based on local chemical features).

Because atoms of the same type contribute identically across molecules, WC yields an atom-wise decomposition of $\log P$. Its main limitation is context-invariance: interaction-driven effects that change effective polarity (e.g., local environment, intramolecular hydrogen bonding, and solvent exposure) are not modeled explicitly. Here, WC is used as a deterministic baseline that is subsequently modulated in a context-dependent manner.

Dataset and Experimental LogP Values

Experimental $\log P$ values were taken from the OPERA dataset (curated from PHYSPROP) [6], comprising roughly 14,000 organic compounds.

Molecules are provided as 2D structures from which connectivity and atom/bond features are derived. Each molecule has a single experimental $\log P$ used as the regression target.

Molecular Graph Representation and Message Passing

Molecules are represented as graphs in which atoms are nodes and covalent bonds are edges. To incorporate molecular context beyond local atom types, we propagate information along bonds using a message-passing scheme inspired by CMPNN [7].

Each atom is initialized with a fixed (non-learned) feature vector encoding its local chemistry (e.g., type, degree, formal charge, hybridization, aromaticity, bonded hydrogens, and valence), and each bond is represented by features such as bond order and conjugation. Over multiple message-passing steps, atoms aggregate information from their bonded neighborhoods, yielding context-enriched atom embeddings.

At message-passing step t,

$$\mathbf{m}_i^{(t)} = \sum_{j \in \mathcal{N}(i)} \phi\Big(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{e}_{ij}\Big), \qquad \mathbf{h}_i^{(t)} = \psi\Big(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\Big),$$

where $\mathcal{N}(i)$ is the bonded neighborhood and $\mathbf{e}_{ij}$ are bond attributes.

The resulting embeddings provide a compact representation of local graph context that is subsequently used to predict per-atom modulation factors.

Learning Atom-Level LogP Modulation Factors

Context-enriched atom embeddings are used to learn per-atom scaling factors applied to the WC contributions. For each atom i, a Multi-Layer Perceptron (MLP) takes the concatenation of the atom embedding $\mathbf{h}_i$ and a pooled molecule embedding $\mathbf{g}$ and predicts a scalar factor f_i.

Training minimizes molecular-level mean squared error,

$$\mathcal{L} = \frac{1}{M} \sum_{j=1}^{M} \Big(\log P_{\text{corr}}^{(j)} - \log P_{\text{exp}}^{(j)}\Big)^2.$$

The corrected lipophilicity is

$$\log P_{\text{corr}} = \sum_{i=1}^{N} c_i \, f_i,$$

with

$$f_i = f_\theta([\mathbf{h}_i \,\|\, \mathbf{g}]),$$

where c_i is the fixed WC contribution.

This design preserves a WC-anchored atomic decomposition while allowing context-dependent reweighting, and it enables qualitative analysis by inspecting which atoms receive strong positive or negative modulation.

Model Architecture Overview

Figure 1 summarizes the context-aware atom-level modulation pipeline used throughout this work.

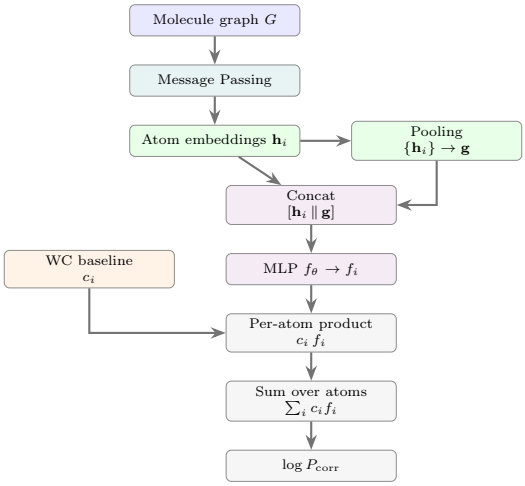


Figure 1: Architecture of the context-aware atom-level modulation model. Molecular graphs are featurized deterministically, context is propagated via non-parametric message passing, and a shared MLP predicts per-atom scaling factors applied to the Wildman–Crippen baseline before summation to obtain $\log P_{\text{corr}}$.

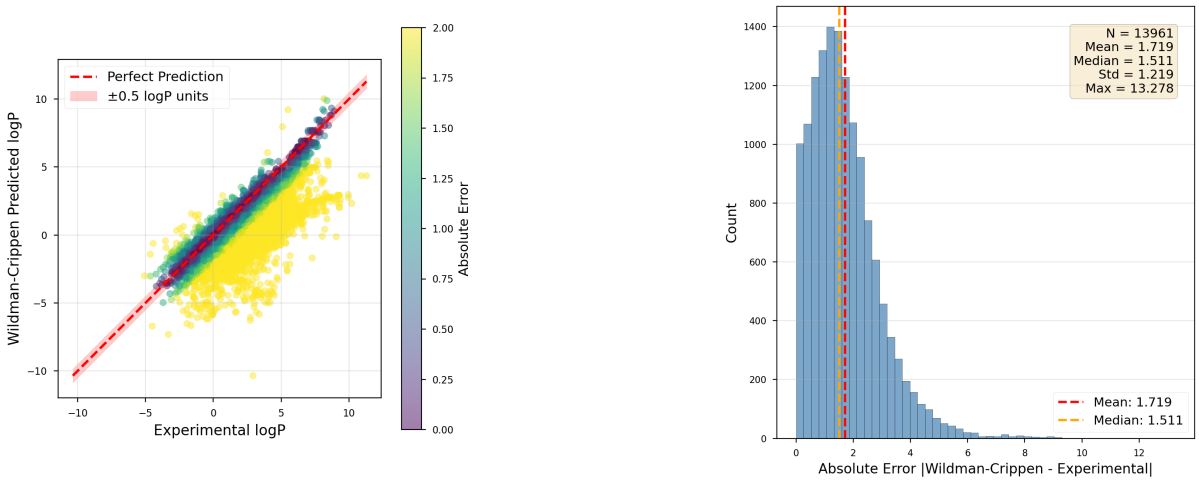
Results

The results are organized around the central methodological stance of this work: baseline error is treated as structured signal rather than noise. We first quantify performance relative to an interpretable fragment baseline, then evaluate whether the learned, context-dependent deviations improve prediction accuracy and provide an atom-level lens for interpretation.

Baseline LogP Prediction Using the Wildman–Crippen Model

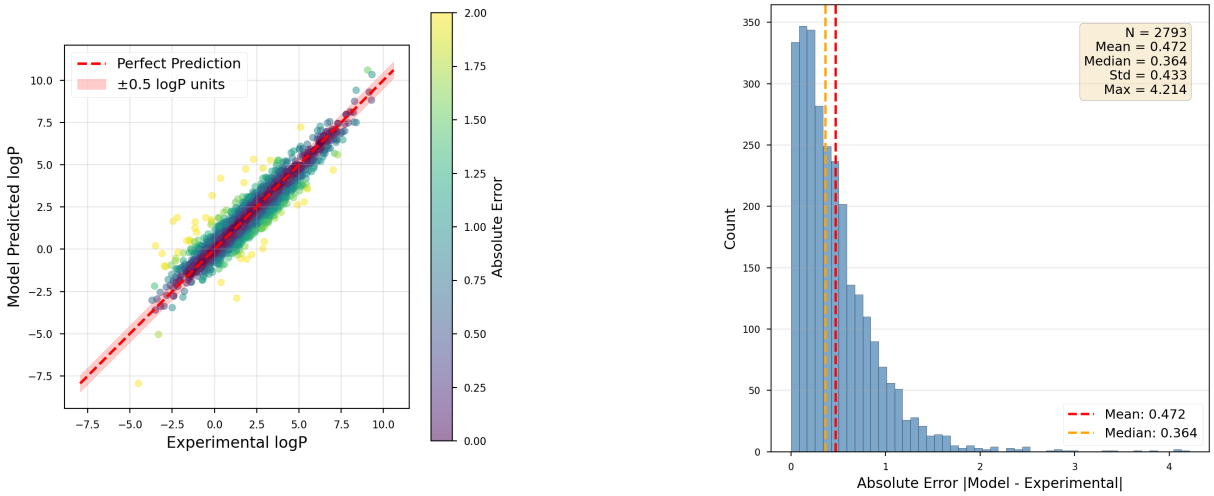
As a reference point, the predictive performance of the Wildman–Crippen (WC) atomic fragment model was evaluated against experimental $\log P$ values from the OPERA dataset. For each molecule, lipophilicity was computed as the sum of fixed atomic contributions.

Figure 2 shows predicted versus experimental $\log P$ values and the corresponding error distribution. While the WC model captures the overall trend in lipophilicity, substantial deviations are observed across a wide range of compounds.



(a) Predicted vs experimental $\log P$. (b) Error distribution.

Figure 2: Baseline performance of the Wildman–Crippen model on the OPERA dataset.



(a) Predicted vs experimental $\log P$. (b) Error distribution.

Figure 3: Performance of the context-aware atom-level modulation model on the OPERA dataset.

Context-Dependent Atom-Level Modulation Improves LogP Prediction

The full model augments the WC baseline with learned, context-dependent atom-level modulation factors derived from molecular graph representations.

Figure 3 shows predicted versus experimental $\log P$ values and the corresponding error distribution for the context-aware model. Predictions are more tightly concentrated around the diagonal relative to the WC baseline, indicating improved agreement with experiment.

The corrected residuals exhibit a narrower distribution and reduced tail mass compared to the baseline. Quantitatively, the full model achieves lower mean squared error (MSE) and mean absolute error (MAE) than the WC model, demonstrating the benefit of incorporating molecular context.

Ablation Studies: Role of Atomic Features and Molecular Context

To disentangle the contributions of local atomic features, propagated molecular context, and the WC prior, two ablation models were evaluated in addition to the baseline and full model.

In the *atoms-only* ablation, atom-level modulation factors were predicted using local atomic features only, without message passing. In the complementary *context-only* ablation, molecular graph representations were used to predict $\log P$ directly, without incorporating the WC atomic baseline.

Table 1 summarizes performance across all models. Both ablations substantially outperform the WC baseline, indicating that learned representations capture structure–property relationships absent from the fragment-based model. The atoms-only and context-only models achieve comparable accuracy, while the full model consistently yields the lowest error.

Table 1: Performance comparison of ablation models on the test set.

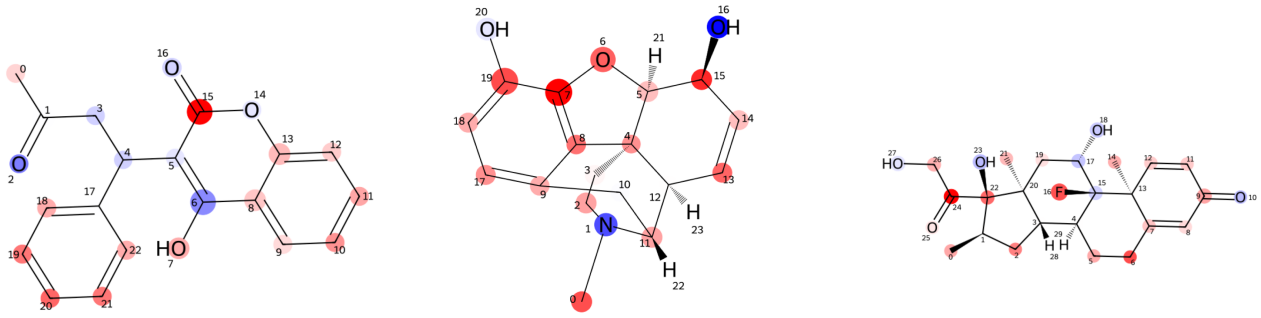
Model	RMSE	MAE
Wildman–Crippen baseline	2.09	1.73
Atoms-only ablation	1.01	0.77
Context-only ablation	1.06	0.79
Full model	0.64	0.47

These results indicate that atom-level modulation explains the majority of the improvement over the baseline, while propagated molecular context and the WC prior provide complementary refinements.

Atom-Level Modulation and Molecular Interpretability

To examine how molecular context influences individual atomic contributions, the learned atom-level modulation factors were analyzed at the level of individual molecules.

Representative case studies were used to evaluate whether the learned atom-level corrections reflect chemically plausible, interaction-driven effects, rather than only improving aggregate metrics. In particular, lipophilicity is shaped by context-dependent interactions (e.g., local polarity, hydrogen-bond patterns, and solvent exposure), so we examined whether the model's atom-level modulation highlights regions where such interactions are expected to influence effective $\log P$.



(a) Warfarin (b) Morphine (c) Dexamethasone

Figure 4: Test molecules used to qualitatively evaluate interpretability of the learned atom-level $\log P$ corrections.

Figure 4 shows three drug molecules selected for qualitative testing: Warfarin, Morphine, and Dexamethasone. Atom colors encode the change in atom-level contributions predicted by our model relative to the WC baseline (red: more hydrophobic; blue: more hydrophilic; intensity: magnitude). In all three cases, the predicted $\log P$ was closer to the experimental value than the WC baseline.

Warfarin (partial alignment). In warfarin (Figure 4a), the model assigns increased hydrophobic contributions across the aromatic ring systems (e.g., atoms 9–13 and 18–22), consistent with the largely nonpolar

character of these regions. The hydroxyl oxygen (atom 7) is also shifted toward increased hydrophobicity, consistent with a local interaction involving the 4-hydroxyl and the carbonyl oxygen (atom 16) [8]. However, the carbonyl oxygen (atom 16) remains hydrophilic, indicating that the model captures the polarity modulation primarily through the hydroxyl site rather than symmetrically across both interaction partners.

Morphine (partial alignment). In morphine (Figure 4b), the model increases hydrophobic contributions across much of the fused ring scaffold, and it assigns a more hydrophobic contribution to the ether oxygen (atom 6), suggesting partial recognition of local masking. In contrast, the hydroxyl oxygens (atoms 16 and 20) and the amine nitrogen (atom 1) remain strongly hydrophilic, indicating limited modulation at key polar sites often involved in context-dependent interactions [9].

Dexamethasone (weak alignment). In dexamethasone (Figure 4c), the model achieves its correction mainly by increasing hydrophobic contributions across the steroid core (e.g., atoms 5–9, 12, and 19) and the fluorine (atom 16). Several polar functional groups remain strongly hydrophilic (e.g., hydroxyl oxygens 18, 23, and 27; ketone oxygen 10), suggesting that the model captures the global trend but provides weaker atom-level evidence for interaction-driven polarity modulation.

Together, these examples illustrate the central theme of this work: an atom-decomposed baseline can be improved by learning context-dependent corrections, and the resulting per-atom modulations provide a practical lens for analyzing where the model attributes interaction-driven effects within a molecule.

Discussion

This work is best viewed as a methodological stance: instead of treating baseline error as noise, we treat systematic deviations from a chemically grounded baseline as structured signal to be learned and analyzed. Under this framing, the goal is not only improved $\log P$ prediction, but also a more interpretable representation of where context-dependent effects are attributed at the atom level.

Overall, the approach was partially successful: it consistently improved predictive accuracy over the Wildman–Crippen baseline, and the ablation results indicate that both local atom features and broader molecular context contribute meaningfully to the final prediction.

Beyond accuracy, we evaluated interpretability by inspecting atom-level modulation patterns on three representative drug molecules. Here the conclusions were more nuanced. The learned corrections often amplified hydrophobic regions in a chemically sensible way, and in some cases highlighted atoms plausibly involved in interaction-driven effects. However, the alignment with specific, experimentally supported interaction mechanisms was only partial, suggesting that improved numerical performance does not always translate into a mechanistically faithful decomposition at the atom level.

These limitations are consistent with the information available to the model: a 2D graph representation and local message passing cannot fully encode three-dimensional conformations, solvent-dependent reorganization, or transient hydrogen-bond networks that can strongly influence partitioning behavior. Incorporating 3D structure, conformational ensembles, or environment-aware features is therefore a promising direction for strengthening the link between learned atom-level corrections and physical interaction mechanisms.

Nevertheless, the results support the central framing of this report: even when the underlying predictor is a deep-learning "black box," anchoring predictions to an interpretable baseline makes the learned deviations informative. In practice, the per-atom modulation factors provide a useful analysis tool for generating and prioritizing hypotheses about which regions of a molecule drive context-dependent changes in effective lipophilicity.

Author Contributions

OS developed the code, designed and performed the computational experiments, analyzed the results, and wrote the manuscript.

Acknowledgments and Disclosure

This work was prepared with assistance from the AI platform OpenAI Prism [10]. The author takes full responsibility for the content of this manuscript, including all analyses, interpretations, and conclusions.

Appendix

The code used for this work can be found at https://github.com/OfirSimhi1612/logP_factor_learning.

References

[1] Alex Avdeef. *Absorption and Drug Development: Solubility, Permeability and Charge State*. Wiley, 2 edition, 2012.

[2] Christopher A. Lipinski. Experimental and computational approaches to estimate solubility and permeability in drug discovery. *Advanced Drug Delivery Reviews*, 23:3–25, 1997.

[3] Christopher A. Lipinski, Franco Lombardo, Beryl W. Dominy, and Paul J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 46:3–26, 2001.

[4] Keith T. Butler, Daniel W. Davies, Hugh Cartwright, Olexandr Isayev, and Aron Walsh. Machine learning for molecular and materials science. *Nature*, 559:547–555, 2018.

[5] Scott A. Wildman and Gordon M. Crippen. Prediction of physicochemical parameters by atomic contributions. *Journal of Chemical Information and Computer Sciences*, 39(5):868–873, 1999.

[6] Kamel Mansouri, Christopher M. Grulke, Richard S. Judson, and Antony J. Williams. Opera models for predicting physicochemical properties and environmental fate endpoints. *Journal of Cheminformatics*, 10 (1):10, 2018.

[7] Yuanyuan Song, Shufeng Zheng, Yutong Liu, Jiahai Wang, Ming Yang, and Jian Zhang. Communicative representation learning on attributed molecular graphs. *Proceedings of IJCAI*, 2020.

[8] Milan Remko, Ria Broer, and Anna Remková. A comparative study of the molecular structure, lipophilicity, solubility, acidity, absorption and polar surface area of coumarinic anticoagulants and direct thrombin inhibitors. *RSC Advances*, 4(16):8072–8084, 2014.

[9] Alaa Malik, Sándor Hosztafi, Anna Vincze, András Marton, Márta Kraszni, György T. Balogh, Béla Noszál, and Károly Mazák. Characterization of opioid agonist morphine derivatives with emphasis on medicinal chemistry. *ChemMedChem*, 20(4):e202400654, 2025. doi: 10.1002/cmdc.202400654.

[10] OpenAI. OpenAI Prism: Ai-assisted online LaTeX editor, 2026. URL https://prism.openai.com/. AI-assisted writing and editing support used in preparing this manuscript (accessed 2026-01-29).