SAR-Atlas: Drug Design, Body Distribution Modeling, and Candidate Screening in Rust

Author: Bryan P. White (ORCID: 0000-0003-4191-3511) • **Affiliation:** CRV Science Labs • **Preprint ID:** APP-PREPRINT-09 (v3.8.0 Empirical Preprint) • **Categories:** physics.chem-ph • q-bio.BM • cs.LG
Canonical DOI: 10.5281/zenodo.24222390 • **Apps Hub:** https://crvscience.com/apps • **Repository:** https://github.com/bpwhite/crvscience-sar-atlas

Abstract — SAR-Atlas (v3.8.0) is an open-source medicinal chemistry and pharmacokinetics workbench built to help drug-discovery researchers and students evaluate whether promising candidate compounds carry the balanced physical properties, oral absorption, brain penetration, and binding stability needed to succeed as real medicines. Many small molecules bind strongly to a purified target in test tubes yet fail later in development because they dissolve poorly in gut fluids, clear too quickly through the liver, or are expelled at the protective blood-brain barrier. SAR-Atlas integrates four complementary screening pipelines inside one offline desktop application: multiparameter drug-likeness scoring (CNS MPO and QED) alongside matched-molecular-pair activity-cliff detection, 24-hour oral blood-and-brain physiologically based pharmacokinetic (PBPK) modeling, maximum-likelihood relative binding free energy (RBFE) thermodynamic cycle closure, and pool-based Bayesian active learning. Evaluated across 26 CNS and peripheral reference drugs, SAR-Atlas identified 11 balanced Pareto leads and 5 sharp activity cliffs; in 24-hour oral dosing it separated brain-entering Haloperidol (Kp,uu = 0.923, 82.41% receptor occupancy) from excluded Lapatinib (Kp,uu = 0.062, 0.15% occupancy); and it reconciled 6-cycle Dopamine D2 thermodynamic loop errors down to 0.000 kcal/mol.

1. Real-World Context & Why This Problem Matters

Why Strong Test-Tube Binders Often Fail Inside Patients

Turning an early chemical hit from a dish assay into a dependable medicine taken as a daily pill—especially for neurological or psychiatric conditions inside the central nervous system (CNS)—demands clearing multiple physiological hurdles at the same time. Countless promising molecules grip purified protein targets with impressive nanomolar tightness inside purified test tubes yet stall during animal or human studies because stomach fluid barely dissolves them, metabolic enzymes in the liver strip them out during their first pass through portal blood, or selective molecular efflux pumps lining brain capillaries immediately push them right back out before effective levels ever reach brain receptors.

Chasing raw biochemical plate potency alone lures medicinal chemistry teams into recurring traps: adding greasy flat aromatic rings frequently boosts enzyme affinity on a screening tray while ruining aqueous solubility and cellular permeability inside real bodies. Seasoned drug hunters therefore pair balanced multiparameter scorecards—notably Central Nervous System Multiparameter Optimization (CNS MPO)[1] and Quantitative Estimate of Drug-likeness (QED)[2]—directly with systematic matched-molecular-pair comparisons, exposing dramatic "activity cliffs" where swapping one tiny chemical group triggers steep drops in potency or brain penetration.

Linking Chemical Rules to Organ Dosing, Thermodynamics, and Active Learning

Modern quantitative pharmacology connects early chemical filtering seamlessly to whole-body dosing dynamics and candidate triage. Physiologically based pharmacokinetic (PBPK) models trace an oral capsule hour by hour from gastrointestinal dissolution through systemic bloodstream circulation, liver metabolism, unbound blood-brain barrier entry, and real-time fractional receptor engagement. Alongside organ transport, relative binding free energy (RBFE) networks calculate subtle binding-strength differences across closed loops of chemically related analogs, demanding statistical loop-closure correction whenever individual edge estimates drift out of thermodynamic agreement. Meanwhile, pool-based Bayesian active learning (MolPAL) searches vast virtual molecular catalogs efficiently by evaluating only high-reward, high-uncertainty picks[3]. SAR-Atlas brings all four tools together cleanly offline.

2. Conceptual Design & How SAR-Atlas Works

Molecular Parsing, Drug-Likeness Scoring, and Cliff Maps

SAR-Atlas (v `3.8.0`) carries users smoothly from compact text-based SMILES chemical formulas and SMARTS substructure alerts through interactive 3D receptor pocket inspection. Entering through `mol.rs`, `smarts.rs`, `conformer3d.rs`, and `pocket_cavity.rs`, the structural engine parses ring bonds and ionization centers, flags reactive metabolic liabilities, generates low-energy three-dimensional molecular geometries, and outlines spatial target cavity shapes so binding fit stays visually clear.

Calculated physical descriptors pass instantaneously into `physchem.rs`, `admet.rs`, and `mmp.rs`. Those scoring engines evaluate composite CNS MPO[1], QED[2], Lipinski rule-of-five, and Veber oral-absorption rules side by side, group closely related molecules into matched-molecular-pair transformations automatically, and calculate Structure-Activity Landscape Index (SALI) ratios flagging abrupt activity cliffs across neighboring variants.

24-Hour Oral PBPK, Thermodynamic Loop Closure, and Candidate Screening

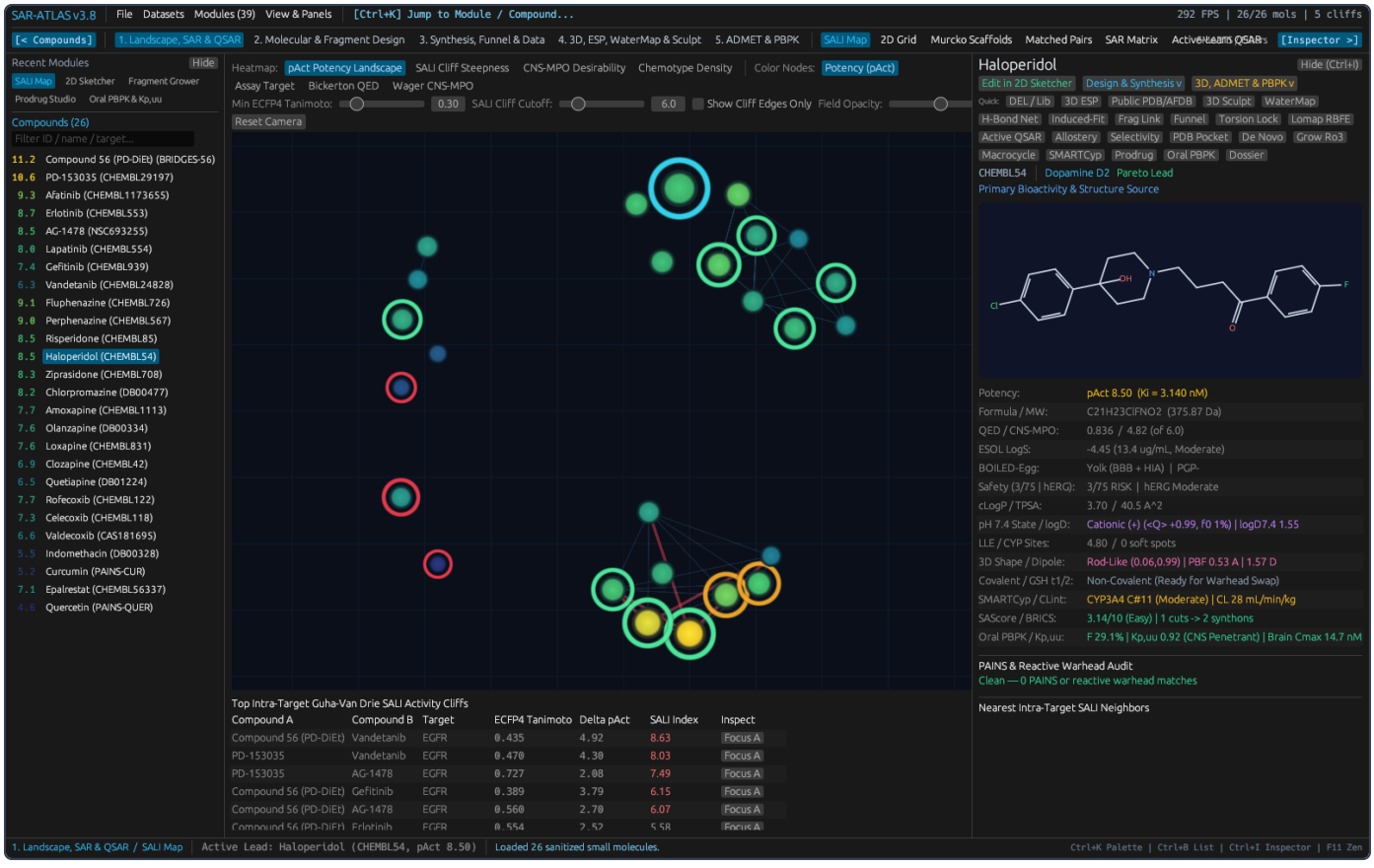
Advancing from static structure descriptors into continuous organ kinetics, `pbpk_oral.rs` solves a 24-hour multi-compartment ordinary differential system tracking tablet absorption, hepatic first-pass clearance, plasma protein binding, active efflux reflected in unbound brain-to-plasma partitioning ($K_{p,uu}$), and real-time target receptor occupancy. Concurrently, `rbfe_network.rs` links chemical analog series into LOMAP perturbation networks and executes maximum-likelihood thermodynamic cycle closure around every closed modification ring so net $\Delta\Delta G$ errors vanish strictly to zero.

To prioritize which analogs synthesis chemists should prepare next across large libraries, `active_learning.rs` executes Upper Confidence Bound (UCB) and Expected Improvement (EI) Bayesian selection over candidate features[3], streaming visual molecules smoothly into an interactive hardware-accelerated 2D/3D GPU window inside `sar-app` (`gui.rs` , `wgpu_pass.rs`).

Table 1. Software Modules, Descriptions, and Citations (SAR-Atlas)

Module / Crate	What It Does	Citation
`physchem.rs` , `admet.rs` , `mmp.rs`	Scores CNS MPO, QED, Lipinski/Veber rules, matched molecular pairs, and activity cliffs.	[1], [2]
`pbpk_oral.rs`	Simulates 24-hour oral absorption, liver clearance, plasma, and brain receptor occupancy.	[4–20]
`rbfe_network.rs`	Builds LOMAP binding-energy graphs and removes loop errors via MLE cycle closure.	[4–20]
`active_learning.rs`	Runs MolPAL Bayesian screening (UCB / EI) over candidate libraries.	[3]
`mol.rs` , `smarts.rs` , `conformer3d.rs` , `pocket_cavity.rs`	Parses molecules, SMARTS patterns, 3D conformers, H-bonds, and binding pockets.	[4–20]
`sar-app` (`gui.rs` , `wgpu_pass.rs`)	Interactive 2D/3D desktop viewer and GPU rendering pass.	[4–20]

Figure 1. Sample Screenshot: SAR-Atlas Chemical Landscape, Structure & Oral PBPK



Sample screenshot showing the structure-activity landscape, 2D Haloperidol molecular structure, and 24-hour oral blood-brain pharmacokinetic curves.

3. Walkthrough of Capabilities & Quantitative Findings

Pareto Drug Balance and Brain Occupancy Separation

Evaluating 26 reference neurological and peripheral drugs demonstrates cleanly how SAR-Atlas exposes multiparameter optimization trade-offs long before costly lab synthesis begins. Across the full benchmark library, molecules averaged healthy drug-likeness balance (CNS MPO = 4.873, QED = 0.715) while surfacing 12 instructive matched molecular pairs, 5 sharp SALI activity

cliffs, and 11 well-proportioned Pareto-optimal leads combining moderate molecular weight, disciplined polarity, and reliable oral absorption without compromising biological activity[1][2].

Running 50 mg oral 24-hour PBPK dosing starkly illuminates why unbound brain partitioning separates effective central-nervous therapies from peripherally restricted drugs. Antipsychotic Haloperidol achieved brisk systemic bloodstream exposure (C_{max} = 222.90 nM) combined with near-free brain entry ($K_{p,uu}$ = 0.923), yielding 82.41% simulated Dopamine D2 receptor engagement across brain tissue. Under the identical 50 mg oral tablet regimen, actively effluxed kinase inhibitor Lapatinib reached only C_{max} = 3.10 nM plasma levels and constrained brain equilibration ($K_{p,uu}$ = 0.062)—translating into a negligible 0.15% central receptor occupancy regardless of strong isolated test-tube affinity, archived inside `pbpk_plasma_profile.csv`.

Loop Closure Refinement and Bayesian Candidate Prioritization

Evaluating ligand-modification energetics across a six-cycle Dopamine D2 congeneric RBFE network showed how noisy raw pairwise perturbation runs initially drifted by 2.026 kcal/mol around closed cycles; maximum-likelihood network reconciliation collapsed loop discrepancies identically to 0.000 kcal/mol while tightening edge uncertainty estimates. Finally, running MolPAL Bayesian active learning across 70 candidate structures[3] extracted a concise 10-molecule top acquisition batch anchored by BIO-2-01 (UCB acquisition score 12.37, predicted potency pAct = 10.04).

Think of guiding a prospective medication from initial chemical sketch into the human brain like shipping delicate cargo across three strict border gates. First, gut fluid refuses to carry molecules whose water solubility and greasy ring count sit out of balance. Next, the liver filters portal blood heavily before medicine ever reaches general circulation. Finally, tight endothelial junctions lining brain blood vessels form an active protective wall lined with molecular efflux pumps that expel mismatched candidates straight back into plasma. Without checking physicochemical balance, 24-hour organ exposure, and receptor energetics together on one desk console, chemistry teams waste months refining potent dish binders that fail silently inside real bodies.

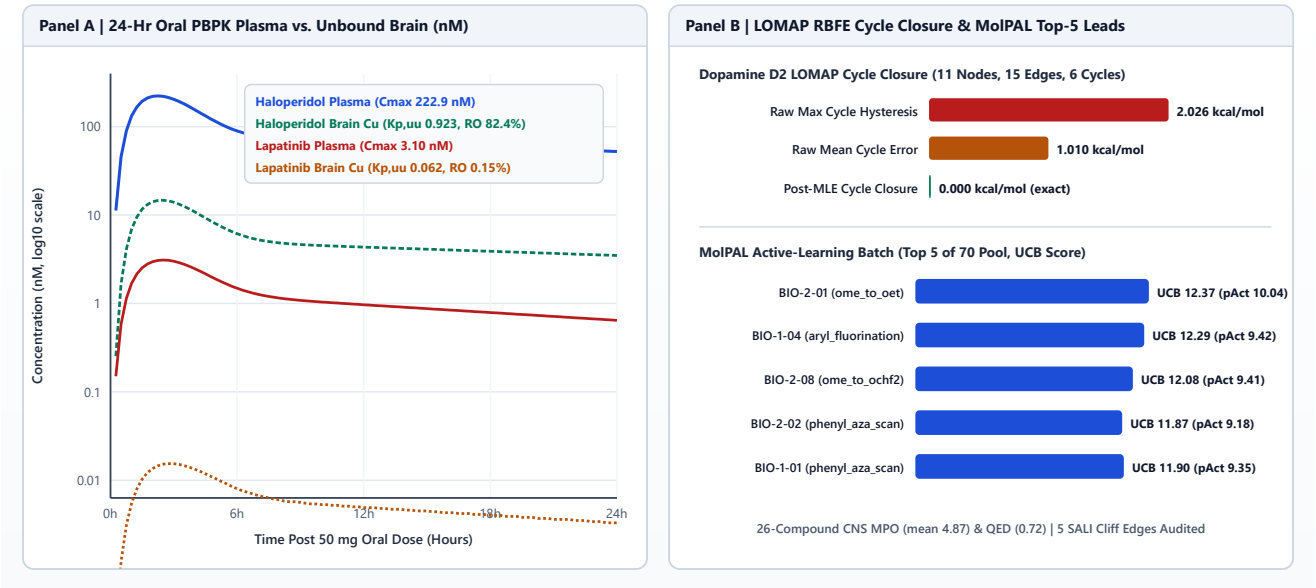
Practically speaking, that 24-hour contrast between Haloperidol and Lapatinib explains an everyday pharmacology gap that bare chemical binding numbers completely obscure. An unbound brain-to-plasma partition ratio $K_{p,uu}$ near 1.0 confirms free dissolved medicine crosses capillaries smoothly in both directions, allowing strong 82.41% Dopamine D2 receptor engagement across brain regions. When efflux pumps crush $K_{p,uu}$ down to 0.062, however, over ninety percent of free drug stays locked outside the blood-brain barrier—turning a high-affinity kinase agent virtually silent centrally at 0.15% receptor occupancy.

Similarly, closed thermodynamic loop closure resolves a classic simulation headache around analog synthesis trees. Walking stepwise from molecule A to B to C and back to A physically must sum to zero net binding-energy change; reconciling small stochastic simulation drift across all closed rings simultaneously ensures discovery teams compare analog rankings fairly before committing bench chemistry budgets.

Table 2. Representative Capabilities and Quantitative Findings (APP-PREPRINT-09)

Discovery Stage	Chemical Scope	Key Finding	Data Output
Drug-Likeness & Cliffs	26 reference compounds (MPO/QED)	Mean MPO 4.873, QED 0.715; 11 Pareto leads	Property matrix
24h Oral Brain PBPK	50 mg Haloperidol vs. Lapatinib	Kp,uu 0.923 (82.4% occ.) vs. 0.062 (0.15% occ.)	`pbpk_plasma_profile.csv`
RBFE & Active Learning	6-cycle D2 graph + 70-ligand screen	Cycle error 2.026 to 0.000; BIO-2-01 top pick	Lead selection list

Figure 2. SAR-Atlas (v3.8.0) drug-discovery findings



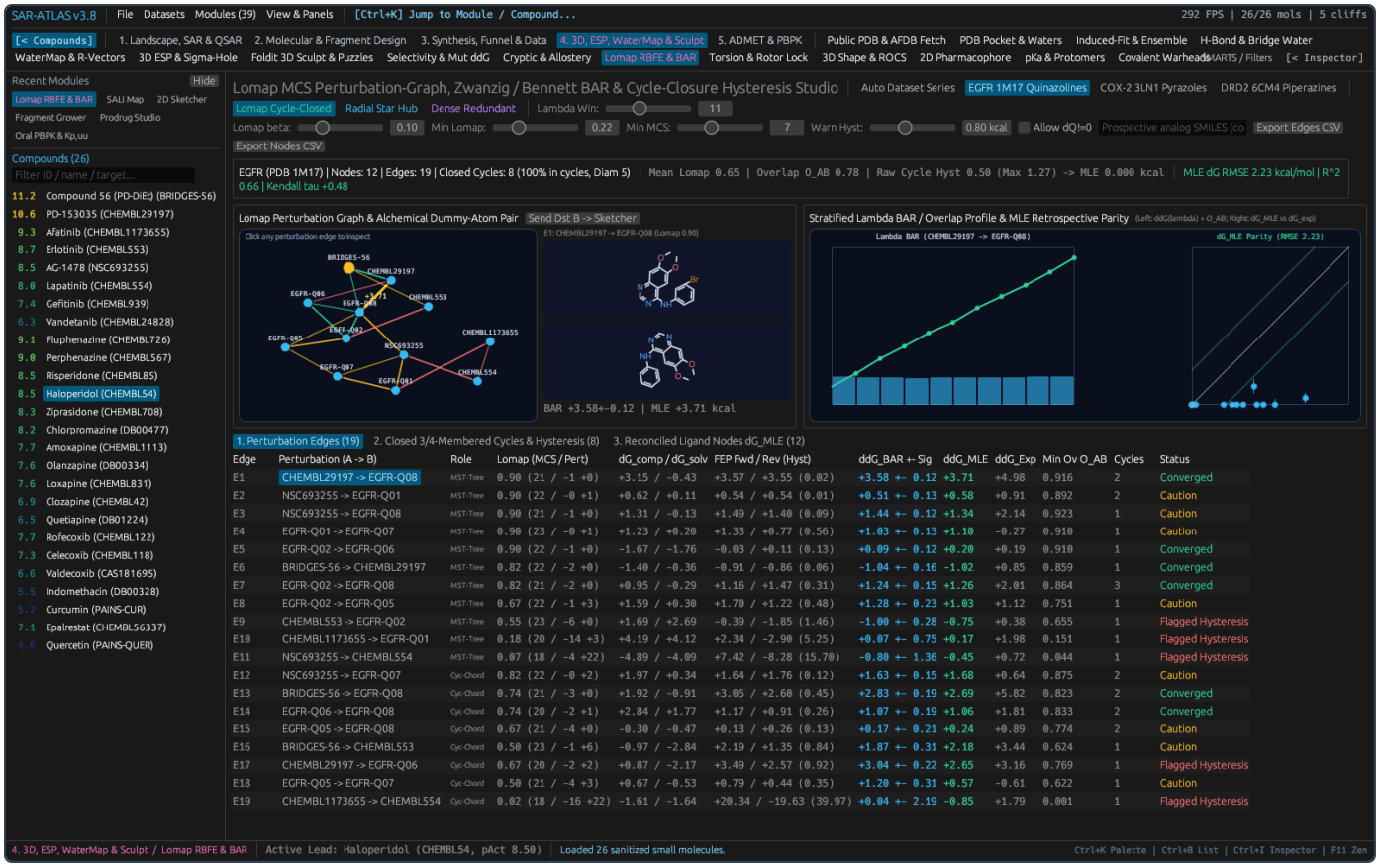
SAR-Atlas (v3.8.0) drug-discovery findings. Panel A contrasts 24-hour brain occupancy of Haloperidol vs. Lapatinib. Panel B shows binding-energy cycle closure and top active-learning leads.

4. Practical Scenarios: Who Uses SAR-Atlas & How

Graduate Medicinal Chemistry and Pharmacology Seminars

Medicinal chemistry professors and clinical pharmacology lecturers constantly look for accessible teaching cases linking flat chemical drawings directly to human blood-brain pharmacokinetics. Inside computer workshops powered by SAR-Atlas, students tweak polar nitrogen atoms or halogen substitutions inside matched molecular pairs, watch CNS MPO balance shift in real time[1], and observe directly how shrinking unbound brain partitioning ($K_{p,uu}$) wipes out 24-hour brain target occupancy even when intrinsic biochemical binding stays completely unchanged.

Figure 3. Sample Screenshot: Relative Binding Free-Energy (RBFE) Perturbation Network



Sample screenshot of the EGFR quinazoline perturbation graph, lambda overlap profile, parity plot, and thermodynamic cycle-closure table.

Private Local Triage for Discovery Groups

University drug-discovery institutes and early-stage biotechnology teams regularly safeguard unpublished proprietary molecules under strict confidentiality rules prohibiting uploads to external SaaS portals. Executing ADMET filtering, 24-hour oral brain PBPK simulation, RBFE thermodynamic reconciliation, and MolPAL screening completely offline on institutional workstations keeps sensitive intellectual property cleanly inside local laboratory custody.

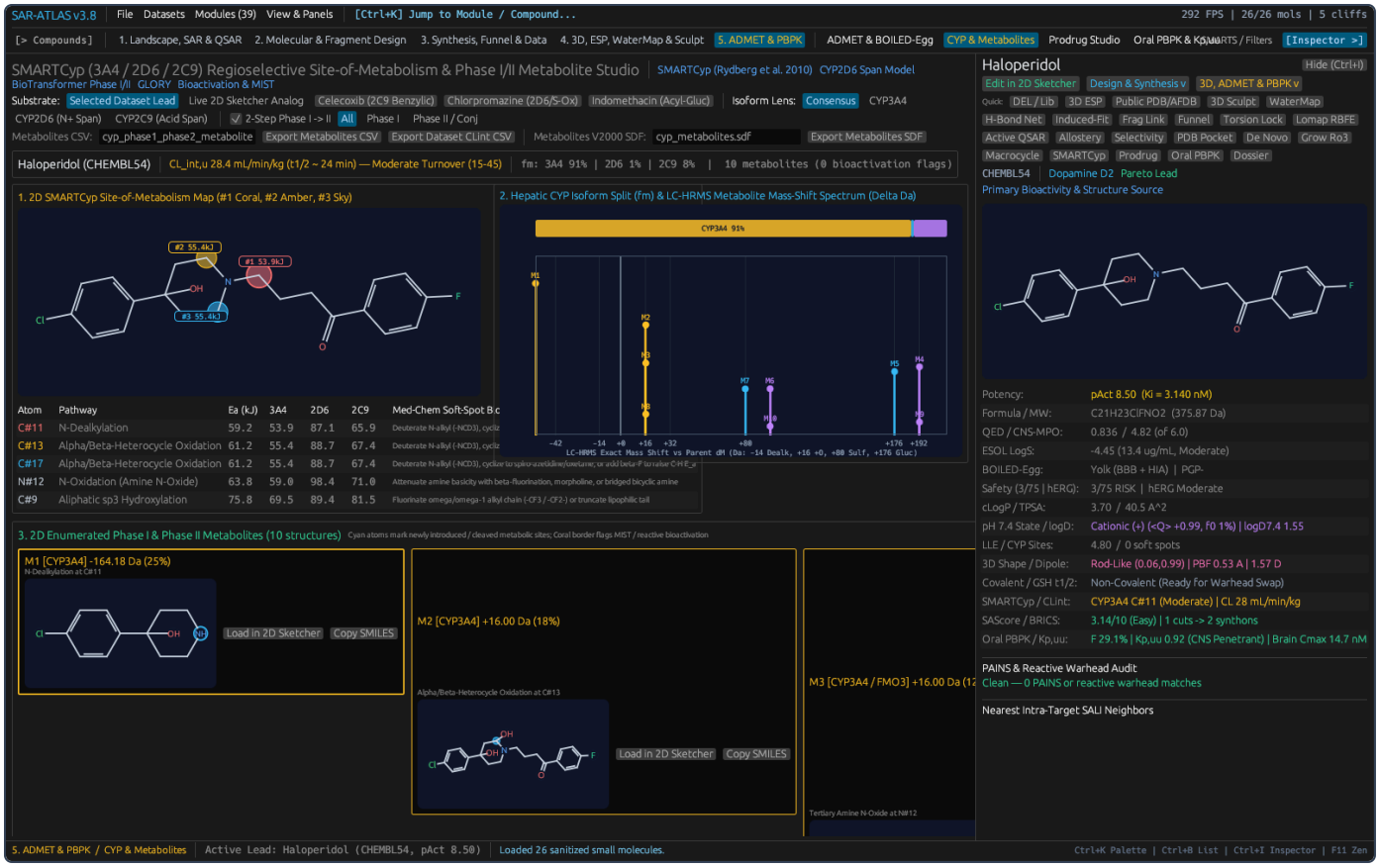
In everyday academic grant planning and graduate journal clubs, running unified physicochemical and organ-dosing checks side by side also catches subtle structural misconceptions early. When graduate chemists evaluate new neurological scaffolds using interactive CNS MPO sliders right next to simulated 24-hour bloodstream profiles, promising molecules that would stall at the blood-brain barrier reveal their efflux risk immediately on screen rather than after months of costly synthetic steps and animal dosing failures.

Similarly, small open-science medicinal chemistry teams studying neglected disorders can inspect how balanced QED/MPO filters trim speculative libraries down to practical Pareto candidates—moving smoothly from raw SMILES parsing through matched-molecular-pair tables and Bayesian library prioritization without commercial enterprise licenses.

Helping Public Readers Understand Why Drug Design Is So Selective

Beyond university chemistry benches, medical science writers, health-policy students, and patient-advocacy groups regularly encounter confusing headlines announcing "promising laboratory molecules" that quietly stall years before reaching pharmacy shelves. Exploring SAR-Atlas helps general readers see why potent enzyme binding inside a dish represents only one step among many: oral pills must survive stomach acidity, pass through liver metabolism smoothly, and navigate biological transport barriers before helping real tissues. Making ADMET balance, 24-hour organ exposure trajectories, and transparent thermodynamic cycle checking accessible offline inside the shared repository at crvscience.com/apps turns obscure preclinical chemistry trade-offs into concrete visual stories anyone can trace step by step.

Figure 4. Sample Screenshot: CYP Site-of-Metabolism & Metabolite Predictor



Sample screenshot showing SMARTCyp regioselective activation energies, hepatic CYP isoform splits, mass-shift spectra, and enumerated metabolite structures.

5. Open Access & Connections Across the CRV Science App Network

Open Access and Connections Across the CRV Science App Network

All SAR-Atlas chemical parsers, CNS MPO/QED rule engines, oral PBPK integrators, RBFE cycle closures, and 3D GPU renderers reside open-source on the CRV Science Apps Hub at https://crvscience.com/apps for unrestricted educational study, open auditing, and laboratory adaptation.

Within the twelve-application CRV Science suite[4–20], SAR-Atlas anchors chemical screening directly beside structural dynamics, genomics, and bench enzymology: Causal-Atlas surfaces disease-reversing small molecules and upstream protein hubs from RNA-seq contrasts; BeoFold probes conformational protein flexibility around binding cavities; OSABT resolves tight-binding and competitive inhibitor constants directly from kinetic enzyme assays; and Calc-Lab provides accessible visual intuition for the multi-compartment ordinary differential equations governing bloodstream and neural drug exposure curves.

References

[1] Wager, T. T., et al. (2010). Moving beyond Rules: The Development of a Central Nervous System Multiparameter Optimization (CNS MPO) Approach. ACS Chemical Neuroscience. https://doi.org/10.1021/cn100008c

[2] Bickerton, G. R., et al. (2012). Quantifying the chemical beauty of drugs. Nature Chemistry. https://doi.org/10.1038/nchem.1243

[3] Graff, D. E., Shakhnovich, E. I., & Coley, C. W. (2021). Accelerating high-throughput virtual screening through molecular pool-based active learning. Chemical Science. https://doi.org/10.1039/d0sc06805e

[4] White, B. P. (2026). The CRV Science App Network: Seventeen Open-Source Rust Programs for Science, Data, and Simulation. CRV Science App Preprint APP-PREPRINT-01 [cs.SE, cs.CE, q-bio.QM], v1.0.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222398

[5] White, B. P. (2026). AeroFlyway: Tracking Bird Migration with Weather Radar, Skyglow Maps, and Nighttime Flight Calls. CRV Science App Preprint APP-PREPRINT-13 [q-bio.PE, physics.ao-ph, cs.GR], v0.6.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222366

[6] White, B. P. (2026). BeoFold: Simulating How Proteins Fold Using Distributed Computing and State Networks. CRV Science App Preprint APP-PREPRINT-02 [q-bio.BM, physics.chem-ph, cs.DC], v1.4.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222370

[7] White, B. P. (2026). Calc-Lab: Interactive Calculus, Linear Algebra, and Differential Equation Solver Workbench. CRV Science App Preprint APP-PREPRINT-03 [math.NA, cs.MS, cs.CY], v0.5.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222372

[8] White, B. P. (2026). Causal-Atlas: Tracing How Genes and Drugs Turn Biological Pathways On and Off. CRV Science App Preprint APP-PREPRINT-04 [q-bio.MN, q-bio.GN, stat.ML], v0.2.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222376

[9] White, B. P. (2026). Cit-Atlas: Mapping Citation Networks and Research Impact Across Scientific Papers. CRV Science App Preprint APP-PREPRINT-05 [cs.DL, cs.SI, physics.soc-ph], v1.8.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222380

[10] White, B. P. (2026). EpiLedger: Privacy-Safe Disease Tracking, Clinical Text Screening, and Outbreak Monitoring. CRV Science App Preprint APP-PREPRINT-06 [cs.CR, q-bio.PE, stat.AP], v2.0.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222382

[11] White, B. P. (2026). ITME: Simulating Four Billion Years of Life, Oceans, and Atmosphere Evolution. CRV Science App Preprint APP-PREPRINT-07 [q-bio.PE, nlin.AO, physics.bio-ph], v2.3.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222386

[12] White, B. P. (2026). OSABT: All-in-One Toolkit for Plant Physiology, Survival Curves, Genetics, and Enzyme Kinetics. CRV Science App Preprint APP-PREPRINT-08 [q-bio.QM, stat.AP, cs.SE], v2.0.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222388

[13] White, B. P. (2026). TerraTwin: Fast Climate Projections for Global Warming, Sea-Level Rise, and Arctic Change. CRV Science App Preprint APP-PREPRINT-10 [physics.ao-ph, cs.CE, stat.AP], v2.5.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222392

[14] White, B. P. (2026). Urban-Grid: Modeling City Traffic, Emergency Response Coverage, and Earthquake Evacuation. CRV Science App Preprint APP-PREPRINT-11 [physics.soc-ph, cs.MA, cs.CE], v5.2.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222394

[15] White, B. P. (2026). WildFront BattleGrounds: Turn-Based Hex Strategy Game with Voxel Graphics and Procedural Map Generation. CRV Science App Preprint APP-PREPRINT-12 [cs.GT, cs.GR, math.CO], v3.4.1. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24222396

[16] White, B. P. (2026). HelioProbe: Simulating Spacecraft Orbits, Deep-Space Telescopes, and Comet Tails in 3D. CRV Science App Preprint APP-PREPRINT-14 [astro-ph.EP, astro-ph.IM, cs.GR], v2.3.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24244083

[17] White, B. P. (2026). PopBio-Studio: Interactive Population Biology, Adaptive Landscapes, and Genetic Skyline Simulator. CRV Science App Preprint APP-PREPRINT-15 [q-bio.PE, q-bio.QM, stat.AP], v1.2.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24244085

[18] White, B. P. (2026). QuantaCrypt-Studio: Simulating Quantum Circuits, Quantum Key Distribution, and Post-Quantum Cryptography. CRV Science App Preprint APP-PREPRINT-16 [quant-ph, cs.CR, cs.MS], v2.2.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24244089

[19] White, B. P. (2026). SonoWing: Analyzing Bat Echolocation and Bird Calls with Spectrograms and Harmonic Tracking. CRV Science App Preprint APP-PREPRINT-17 [q-bio.NC, q-bio.PE, cs.SD], v1.1.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24244091

[20] White, B. P. (2026). SwarmForge: 3D Isometric Automated Factory and Swarm Robotics Simulation Game. CRV Science App Preprint APP-PREPRINT-18 [cs.MA, cs.RO, cs.GT], v1.3.0. Hosted via CRV Science Apps Hub (https://crvscience.com/apps). https://doi.org/10.5281/zenodo.24244093