

Closed-Loop Reinforcement Learning in De Novo Molecular Design: Bridging Generative AI with Automated Chemical Synthesis

Michele Covington

In Association with Blue Marble University, 40 Hillsborough Street, Roseau, 00152

Commonwealth of Dominica

Email: [michele.covington\[at\]gmail.com](mailto:michele.covington[at]gmail.com)

Abstract: *Closed-loop reinforcement learning offers a promising approach to integrating generative molecular design with automated chemical synthesis and experimental evaluation. However, translating autonomous chemical optimization into pharmaceutical discovery remains challenging because biological measurements are often delayed, variable, and affected by assay limitations. This critical review and perspective examines current autonomous discovery platforms and evaluates the challenges associated with incorporating physical experimental feedback into molecular optimization. It proposes an uncertainty-aware, multi-objective reward architecture and a four-gate qualification framework covering operational synthesizability, analytical identity and purity, biological assay validity, and governed policy updating. Particular attention is given to distinguishing statistically censored measurements from pharmacologically confounded or invalid observations. The article also introduces supervised autonomy as a practical governance approach in which automated systems perform routine experimental operations while scientific experts retain responsibility for critical decisions. The proposed framework provides a foundation for improving the reliability, interpretability, and practical implementation of experimentally informed molecular discovery. Prospective validation remains necessary to establish its effectiveness.*

Keywords: Generative artificial intelligence; reinforcement learning; de novo molecular design; automated chemical synthesis; pharmaceutical developability.

1. Introduction

Small-molecule drug discovery requires navigating an immense chemical space estimated to exceed 10^{60} synthetically feasible structures [1]. Lead optimization demands the simultaneous harmonization of target potency, selectivity, physicochemical properties, absorption, distribution, metabolism, excretion, toxicity (ADMET), and synthetic manufacturability. Because physical high-throughput screening (HTS) samples only an infinitesimal fraction of this universe, generative artificial intelligence (AI) has emerged as a computational paradigm for de novo design. Deep generative architectures—including recurrent neural networks (RNNs), variational autoencoders (VAEs), graph neural networks (GNNs), transformers, and score-based diffusion models—learn underlying probabilistic representations of bioactive chemical space, sampling novel candidates designed to satisfy specific multi-parameter constraints [1].

Reinforcement learning (RL) couples molecular generation with sequential decision-making [2]. An agent builds a candidate structure step-by-step, receiving feedback from evaluation models. While early sequence-based RL biased generation toward virtual descriptors [2], unconstrained computational optimization frequently yields "synthetic fantasies"—molecules with high predicted scores that are impossible to synthesize, metabolically unstable, or deceptively scored by biased surrogate models [3,4].

The rapid development of automated chemical synthesis, continuous-flow robotics, and self-driving laboratories (SDLs) offers a physical mechanism to close the design–make–test–analyze–learn (DMTAL) cycle [5–8,23]. Computer-controlled liquid handlers, automated microfluidic

reactors, modular batch workcells, and online chromatographic instrumentation allow automated reaction execution, work-up, and purification. Autonomous platforms have demonstrated closed-loop optimization of photocatalytic reaction conditions [5], multi-property optical materials [6], and tool-coordinated organic synthesis driven by large language models (LLMs) [8].

Despite these advances, a fundamental distinction separates autonomous chemical reaction optimization from autonomous pharmaceutical drug discovery. Chemical reaction loops optimize physical or chemical properties (e.g., yield, selectivity, optical absorbance) that can be measured inline or within minutes to hours under abiotic, well-controlled conditions. Conversely, a pharmaceutical discovery loop must evaluate candidate molecules across living biological systems. Target binding kinetics, functional cellular responses, membrane permeability, and metabolic stability introduce multi-day incubation delays, plate-to-plate drift, well-specific spatial artifacts, and bounded or invalid readouts caused by low solubility, cytotoxicity, or assay interference.

The purpose of this paper is to examine how closed-loop RL can integrate delayed, noisy, censored, and confounded analytical and biological measurements so that physical feedback improves rather than destabilizes generative molecular optimization. The manuscript is organized around four contributions: an evidence-based critique of current autonomous platforms; an uncertainty-aware algorithmic reward formulation; an operational four-gate qualification framework for physical-to-computational translation; and a prospective supervised-autonomy governance model for pharmaceutical implementation.

2. Methods

A structured critical narrative review was used to synthesize evidence across pharmaceutical chemistry, cheminformatics, reinforcement learning, laboratory automation and robotics, analytical chemistry, and biological screening. The purpose was not to generate a pooled treatment effect or claim systematic-review completeness, but to critically evaluate where current closed-loop platforms achieve experimental closure and where pharmaceutical feedback remains methodologically fragile.

Comprehensive literature searches were conducted through September 2026 across PubMed, Crossref-indexed publisher databases, IEEE Xplore, and ChemRxiv. Boolean search strings integrated four intersecting concepts:

- 1) ("reinforcement learning" OR "generative molecular design" OR "de novo design")
- 2) ("closed loop" OR "self-driving laboratory" OR "autonomous discovery")
- 3) ("retrosynthesis" OR "automated synthesis" OR "robotic chemistry")
- 4) ("biological assay" OR "cell-based screening" OR "Z-factor" OR "interval censoring")
- 5) ("SiLA 2" OR "AnIML" OR "laboratory interoperability")

Eligible publications described: (a) generative molecular optimization linked to physical synthesis; (b) robotic experimental platforms demonstrating autonomous chemical optimization; (c) algorithmic strategies for delayed, noisy, or multi-objective RL; or (d) open laboratory communication and data standards. Studies focusing solely on theoretical molecular generation without synthesis considerations, or generic reaction optimization lacking multi-property medicinal chemistry relevance, were excluded. Preprints were incorporated selectively only when they established capabilities not yet documented in peer-reviewed literature

(e.g., recent end-to-end multi-target kinase campaigns [10]) and are explicitly appraised with appropriate caution.

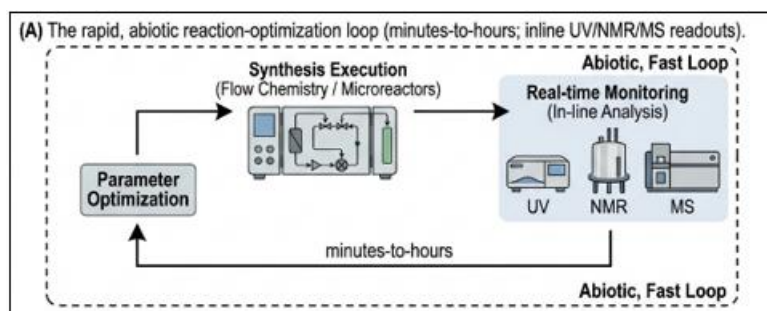
Evidence was synthesized qualitatively across technology readiness, hardware topology, feedback latency, biological validity, uncertainty handling, and degree of experimental loop closure. The review used transparent literature-search and evidence-selection principles but was not designed as a systematic review or formal PRISMA-ScR scoping review [20]. The framework and equations presented below are prospective technical formulations derived from the reviewed evidence rather than results of a new experimental validation study.

3. Results

3.1 Evidence-Based Critique of Existing Platforms: Chemical Execution vs. Stochastic Biological Feedback

A critical analysis of the literature reveals a significant disparity in technology readiness levels across the pharmaceutical closed loop. Molecular generation, retrosynthetic planning, and automated chemical synthesis have reached comparatively advanced levels of technical maturity, whereas analytical data reduction and biological screening remain important bottlenecks.

Generative algorithms can sample large numbers of candidate structures rapidly [21,22], and modern retrosynthetic neural networks can prioritize plausible reaction pathways and predict forward reaction outcomes [3,4]. Robotic platforms execute multi-step organic reactions, purify products via automated solid-phase extraction or preparative chromatography, and collect inline spectroscopy without human intervention [5,6]. However, these successes have occurred almost entirely within abiotic chemical domains where physical feedback is rapid and deterministic.



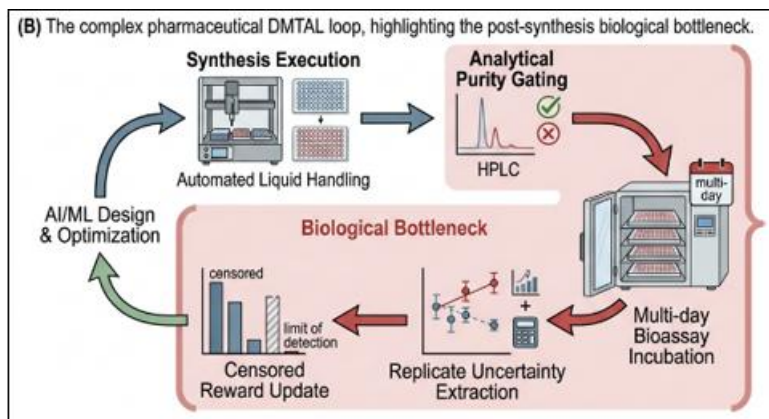


Figure 1: The Closed-Loop Pharmaceutical DMTAL Architecture vs. Reaction- Optimization Loops.

As illustrated in **Figure 1**, a pharmaceutical DMTAL loop diverges sharply from an abiotic reaction loop. In an automated reaction platform, the experimental unit is a reaction vessel whose chemical yield or conversion is measured directly upon reaction completion. In a pharmaceutical closed loop, the experimental unit is a living cellular or biochemical assay plate. The observed biological response is conditional not only on molecular structure, but also on compound solubility, liquid-handling precision, plate geometry, incubation duration, cell passage number, and serum protein binding.

Furthermore, biological readouts introduce multi-day feedback latencies. While computational generation occurs in milliseconds and chemical synthesis requires hours, cell-based functional assays and ADME/PK-related measurements can require substantially longer turnaround times, ranging from hours to days or longer depending on the endpoint and workflow. Current literature demonstrates that closed-loop platforms frequently compress or bypass these biological steps, using computational proxies rather than true physical biological feedback to guide the RL agent. When physical biological assays are integrated, published studies disproportionately report clean, successful endpoints, omitting failed reactions, insoluble compounds, and uninterpretable assay plates. Omitting these negative outcomes prevents the learning policy from mapping the true boundaries of drug-like chemical space.

3.2 Algorithmic Reward Formulation: Chemical-Space MDPs and Uncertainty-Aware Reward Shaping

Molecular de novo generation within an RL framework can be formalized as an episodic Markov decision process defined by the tuple:

$$\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma) \quad (1)$$

where state $s_t \in \mathcal{S}$ represents the intermediate molecular representation (e.g., graph fragment, SMILES token, or SELFIES string), action $a_t \in \mathcal{A}$ denotes the addition of an atom, bond, or building block, $\mathcal{P}(s_{t+1} | s_t, a_t)$ is the transition distribution, $\mathcal{R}(sT)$ is the terminal reward awarded upon completing a chemically valid molecule $m = sT$, and $\gamma \in (0, 1]$ is the discount factor. Intermediate step rewards $\mathcal{R}(s_t)$ for $t < T$ are zero, making the return sparse and delayed.

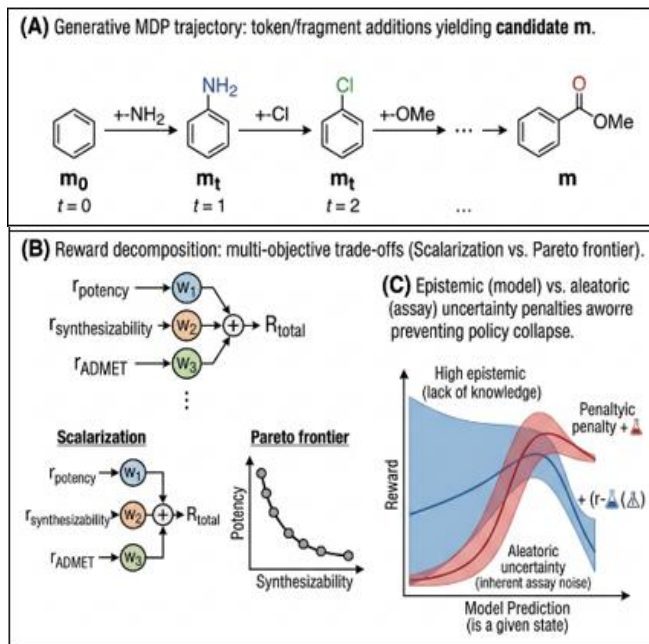


Figure 2: Algorithmic MDP Formulation and Uncertainty-Aware Reward Shaping

When terminal evaluation relies on real-world experimental feedback, a scalar reward dominated by predicted potency can promote undesirable policy behavior, including optimization toward poorly developable structures or exploitation of surrogate-model artifacts. As shown in Figure 2, a robust reward pipeline must balance multi-parameter optimization, synthetic feasibility, and experimental noise:

$$R(m) = \sum_{k=1}^K w_k f_k(m) - \lambda C_{\text{synth}}(m) - \mu \sigma_{\text{exp}}^2(m) \quad (2)$$

where $f_k(m) \in [0, 1]$ represents normalized desirability functions for potency, selectivity margins, and physicochemical properties; w_k denotes weighting coefficients ($\sum w_k = 1$); $C_{\text{synth}}(m)$ is an operational synthetic burden penalty; and $\sigma_{\text{exp}}^2(m)$ captures replicate-level experimental variance and predictive epistemic uncertainty.

Policy parameters θ are updated via policy gradient ascent:

$$\nabla_{\theta} J(\theta) = E_{\pi} \left[\sum_{t=0}^T \nabla_{\theta} \log \pi_{\theta}(a_t | s_t) \cdot (R(m) - B(s_t)) \right] \quad (3)$$

where $B(s_t)$ is a state-dependent baseline that reduces gradient variance [12].

While linear scalarization (Equation 2) produces a single scalar ranking, it implicitly assumes that an exceptional gain

in one endpoint can compensate for a fatal deficit in another (e.g., high affinity masking severe insolubility). In contrast, Pareto-based optimization identifies non-dominated candidates across orthogonal axes, preserving candidate diversity along the potency–developability frontier (Figure 2B). Furthermore, explicit uncertainty propagation (σ^2_{exp}) is critical: epistemic model uncertainty derived from Bayesian neural networks or Monte Carlo dropout [11] penalizes out-of-domain virtual designs, while aleatoric measurement variance penalizes unstable assay readouts (Figure 2C) [9].

Because automated physical synthesis is orders of magnitude slower than computational sampling, pure on-policy training is sample-inefficient [13]. Practical architectures must decouple generative exploration from physical execution: dense, computational surrogate rewards guide the generator within an off-policy replay buffer, while sparse, physical experimental measurements periodically recalibrate the surrogate models and baseline functions via reward decomposition algorithms [14,15].

3.3 Operational Four-Gate Qualification Framework: Retrosynthetic Gating, Analytical Verification, and Pharmacologic Data Integrity

A recurring limitation in computational drug design is the reliance on heuristic synthetic accessibility scores (e.g., SAScore, SYBA). These metrics evaluate structural complexity from historical literature frequencies but fail to reflect whether an automated platform can execute the route.

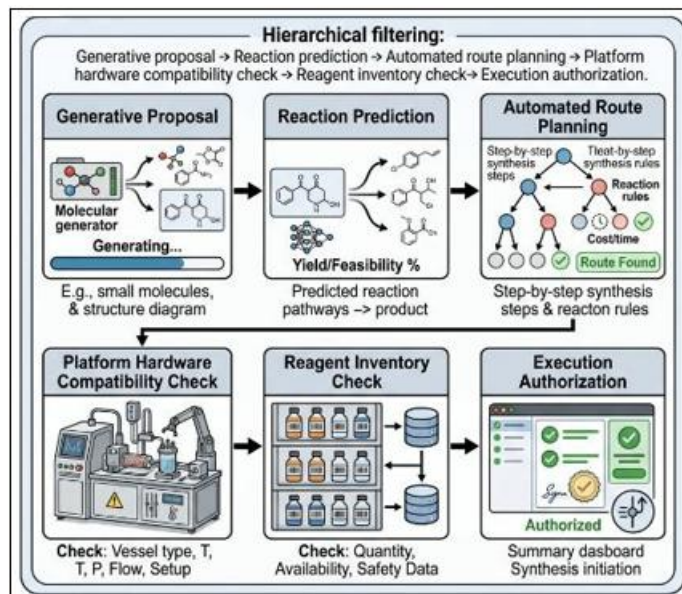


Figure 3: Operational Retrosynthetic Gating and Platform Execution

As detailed in Figure 3, practical execution requires an operational synthesizability gate can be expressed as:

$$\text{Gate}_{\text{synth}}(m) = 1 \text{ if } N_{\text{steps}} \leq N_{\text{max}} \wedge \hat{Y}_{\text{overall}} \geq Y_{\text{min}} \wedge H_{\text{hardware}}(m)=1 \wedge S_{\text{safety}}(m)=1; \text{ otherwise } 0 \quad (4)$$

Candidate molecules that require non-automated manual steps (e.g., cryogenic conditions, hazardous reagents, un-orchestrated flash chromatography) evaluate to $\text{Gate}_{\text{synth}}(m) = 0$. These candidates are filtered prior to robotic dispatch, returning an immediate synthetic penalty to the computational buffer.

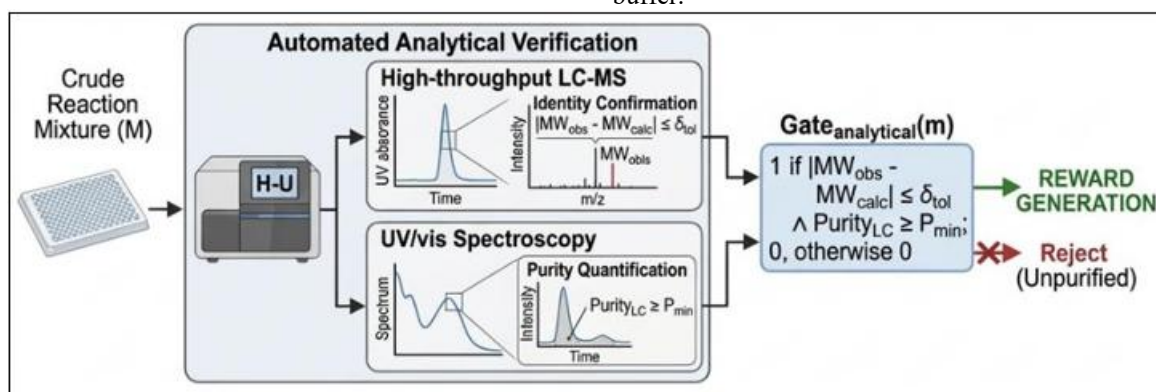


Figure 4A: Automated Analytical Verification

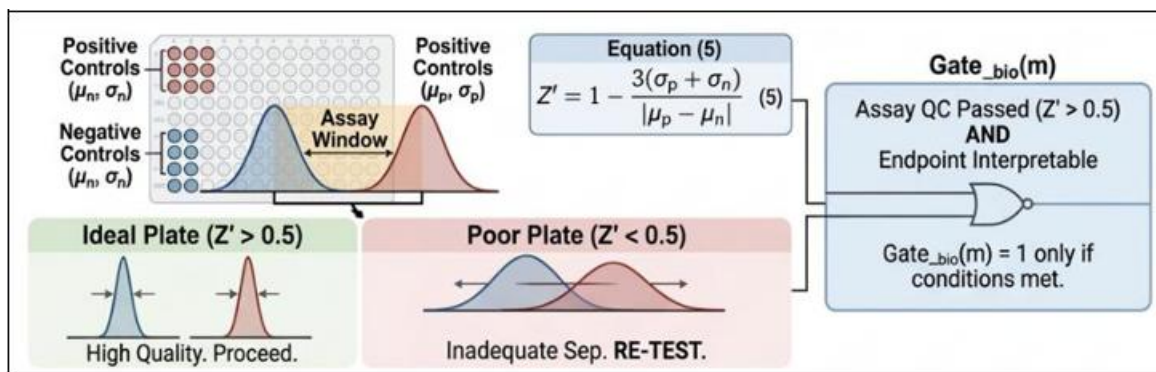


Figure 4B: Statistical Qualification of Biological Assays

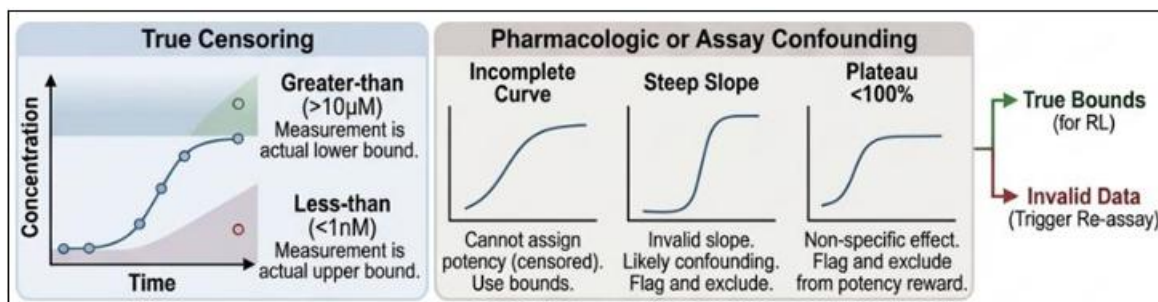


Figure 4C: Classifications of Biological Measurements

Following physical synthesis, crude reaction mixtures cannot be introduced directly into biological assays. They must pass through automated analytical verification (Figure 4A). Samples are subjected to high-throughput LC-MS and UV/vis spectroscopy to confirm identity and quantify purity:

$$\text{Gate}_{\text{anal}} \text{Y}_{\text{ti}} \text{C}_{\text{ai}}(\text{m}) = 1 \text{ if } |\text{MW}_{\text{ofs}} - \text{MW}_{\text{cal}}| \leq \delta_{\text{tol}} \wedge \text{Purity}_{\text{LC}} \geq \text{P}_{\text{min}}; \text{ otherwise } 0$$

Samples failing analytical purity ($\text{P}_{\text{min}} < 85\text{--}90\%$) are rejected. Rewarding an unpurified mixture misattributes biological activity to the target structure when it may stem from unreacted starting materials, catalysts, or off-target isomers.

Biological assay data must also undergo statistical qualification before conversion into an RL reward. Quantitative high-throughput screening using concentration-response titrations can reduce single-concentration artifacts [16]. Microplate quality is commonly assessed using the Z'-factor [17]:

$$Z' = 1 - 3(\sigma_p + \sigma_n) / |\mu_p - \mu_n| \quad (5)$$

where μ_p , σ_p and μ_n , σ_n are the means and standard deviations of positive and negative controls (Figure 4B). A plate with $Z' < 0.5$ indicates inadequate assay-window separation for many screening applications; such observations should generally trigger review or re-testing rather than automatic policy updating [17].

$\text{Gate}_{\text{bio}}(\text{m}) = 1$ only if assay QC passes and the endpoint is interpretable; censored values retain bounds, whereas confounded/invalid values are flagged and excluded from direct potency reward assignment.

Biological measurements should be classified according to the statistical information they actually contain. True censoring must be distinguished from pharmacologic or assay confounding (Figure 4C):

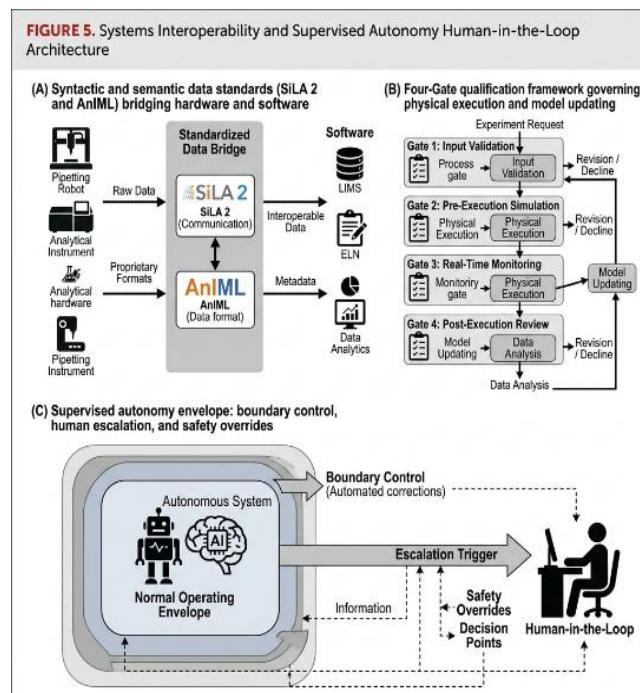
- 1) Right-censored potency (solubility or tested-concentration ceiling): when no 50% effect is reached before the maximum soluble or tested concentration, potency is bounded as $\text{IC}_{50} > \text{C}_{\text{max}}$ rather than assigned a precise inactive value.
- 2) Left-censored potency (assay floor): when potency exceeds the lower limit of quantification, the result is bounded as $\text{IC}_{50} < \text{LLOQ}$ rather than treated as an exact sub-threshold value.
- 3) Assay confounding or non-specific interference: cytotoxicity, colloidal aggregation, reactive or interference-prone chemotypes, fluorescence quenching, and related mechanisms can invalidate or confound the pharmacologic readout. These outcomes are not statistical censoring and should be encoded separately, investigated

with orthogonal assays where appropriate, and excluded from direct potency reward assignment unless a prespecified rule defines an interpretable adverse endpoint.

Treating censored, confounded, and invalid endpoints as a single “zero reward” erases mechanistically important information and can distort policy updates. Closed-loop platforms should retain censoring bounds for valid but limited potency estimates, while using structured reason codes for cytotoxicity, aggregation, assay interference, analytical failure, and other invalidating mechanisms. This separation allows the learning system to distinguish uncertainty about potency from evidence that the assay itself is not interpretable.

3.4 Evidence-Based Platform Benchmark and Hardware–Software Interoperability

To evaluate current technology readiness, representative autonomous and closed-loop platforms were analyzed across decision engines, synthesis hardware, analytical modalities, and biological loop closure (Table 1).



As synthesized in Table 1 and illustrated in Figure 5A, hardware–software integration requires multi-layer interoperability:

- **Syntactic Layer:** Open communication protocols, such as Standardization in Lab Automation (SiLA 2), provide microservice architectures enabling generative algorithms to control disparate physical devices (e.g., Tecan liquid handlers, Chemspeed batch reactors, Agilent LC-MS systems) [18].
- **Semantic Layer:** Standardized data representation formats, such as the Analytical Information Markup Language (AnIML), capture raw spectra, chromatograms, integration parameters, and instrument metadata [19].

Table 1: Comparative decision matrix for representative autonomous and closed-loop discovery platforms.

Platform	Decision Engine	Reaction Planning	Execution Hardware	Readout Modality & Typical Latency	Degree of Pharmaceutical Loop Closure
Burger et al. [5]	Bayesian optimization (active learning); not molecular RL	Not a retrosynthetic workflow (pre-formulated chemical components)	Mobile autonomous robot operating conventional batch instruments	Photocatalytic H_2 evolution; minutes to hours per experiment	Autonomous reaction optimization; lacks de novo molecular generation and biological assay feedback
Koscher et al. [6]	Multiproperty surrogate models with iterative batch selection	Integrated automated synthesis planning	Modular flow synthesis, automated purification, and inline analysis	Physical and optical property measurements (UV-Vis, NMR); iterative days	Autonomous molecular discovery; demonstrates multiproperty closed loop but focused on optoelectronic materials rather than bio-assays
Boiko et al. [8]	Large language model (LLM) agent for reasoning, planning, and tool use	Software-mediated reaction planning and documentation retrieval	Modular external hardware controlled via Python APIs	Standard organic transformations; hours per synthetic campaign	Autonomous experimental orchestration; lacks integrated biological evaluation and iterative policy gradient updates
Born et al. [10] (Preprint)	Generative molecular AI with iterative policy/surrogate updating	Machine-learning-assisted retrosynthetic planning integrated	Integrated liquid handlers, automated synthesis workcells, and plate readers	Kinase target inhibition assays (JAK family); multicycle workflow over several days	Near-complete pharmaceutical DMTA closed loop; demonstrates biological feedback guiding generative AI (pending peer review)

Existing platforms exhibit high maturity in chemical synthesis and physical property profiling (**Table 1**). However, platforms that successfully integrate live biological screening remain rare exceptions. The recent JAK kinase workflow by Born et al. [10] demonstrates an end-to-end pharmaceutical loop over multi-day iterations, but its preprint status emphasizes the need for independent replication and standardized community reporting metrics.

4. Discussion

4.1 Implications for Medicinal Chemistry and Lead Optimization

In conventional medicinal chemistry, lead optimization is an iterative process where chemists synthesize structural analogues to establish structure–activity relationships (SAR) while tracking property trade-offs (e.g., lipophilic efficiency, metabolic clearance). When an automated RL agent navigates this process, uncalibrated scalar rewards can drive the policy toward structural artifacts.

For instance, an RL policy maximizing binding affinity without operational solubility constraints will systematically increase lipophilicity and molecular weight, generating hydrophobic, flat aromatic chemotypes prone to non-specific binding and precipitation. By enforcing multi-objective Pareto frontiers (Figure 2B) and penalizing aleatoric assay variance (σ^2_{exp}), the RL policy is forced to prioritize molecules that maintain active, reproducible signals within physiological solubility limits.

4.2 Latency Asynchrony and Training Stability

A critical operational challenge is the latency mismatch between computational inference and biological testing. A generative model can sample 10^4 candidates in minutes, whereas synthesizing, purifying, and testing 20 molecules in a cell-based assay requires days to weeks.

If policy updates are strictly synchronous, costly automated robotic hardware sits idle while waiting for biological readouts. Conversely, if updates are fully asynchronous, the generative policy continuously updates its weights based on stale surrogate models, causing newer batches to drift away from the parameter space of pending physical experiments.

To mitigate this, closed-loop architectures must adopt a federated batch design: the generative policy explores local chemical space using frozen batch checkpoints, while an off-policy replay buffer stores validated physical returns. As physical biological data become available, return decomposition methods (e.g., RUDDER [15]) assign credit retroactively to early trajectory actions, stabilizing the gradient without halting synthesis.

4.3 Error Propagation Across the Four-Gate Architecture

The four-gate qualification framework (**Figure 5B**) prevents systemic error propagation:

Gate 1 (Operational Synthesizability): Enforces hardware compatibility, preventing robotic liquid handlers from attempting unfeasible transformations.

Gate 2 (Analytical Verification): Blocks impure or decomposed mixtures from biological evaluation, ensuring

that observed phenotypes are driven by the intended chemical structure.

Gate 3 (Biological Assay Qualification, Censoring & Confounding): Validates plate-level statistical robustness ($Z' \geq 0.5$ where appropriate), preserves true right- or left-censoring bounds, and separately flags cytotoxicity, aggregation, PAINS-related interference, or other non-specific assay confounders.

Gate 4 (Governed Updating): Propagates uncertainty into gradient baselines, preventing destabilizing updates from high-variance measurements.

Bypassing any single gate compromises the integrity of the loop. If analytical purity checks are omitted, the policy can reward false positives generated by reactive synthetic intermediates. If censoring bounds are discarded, the policy can learn from invented potency values; if cytotoxicity or assay interference is mislabeled as censoring, the policy can mistake non-specific effects for target-relevant pharmacology.

4.4 Prospective Governance: The Supervised Autonomy Model

Given the safety, regulatory, and cost constraints of pharmaceutical development, fully unattended autonomy is neither practical nor desirable. As depicted in **Figure 5C**, near-term closed-loop systems are most credibly deployed under a paradigm of *supervised autonomy*.

In this framework, robotic hardware and algorithmic models execute low-level, high-throughput workflows (e.g., route planning, liquid dispensing, LC-MS injection, image segmentation) within pre-approved boundary limits. Human pharmaceutical scientists retain high-level governance: defining multi-property desirability functions, establishing safety envelopes (e.g., excluding reactive or explosive intermediates), adjudicating anomalous or interval-censored data, and authorizing lead series advancement. This hybrid model prevents reward hacking and ensures that autonomous discovery remains aligned with clinical and regulatory requirements.

4.5 Limitations of Current Evidence and Future Directions

This review identified several critical reporting gaps across the literature. First, published studies rarely report negative results, such as failed synthetic steps, insoluble analogues, or uninterpretable assay plates. Without negative experimental datasets, computational models cannot learn the true boundaries of chemical synthesizability and biological assay tolerance.

Second, the definition of a "cycle" varies widely across publications, ranging from purely computational optimization iterations to complete synthesis-to-assay workflows. To enable meaningful cross-platform comparisons, the pharmaceutical automation community must establish standardized reporting benchmarks that document: (1) policy checkpoint versions, (2) automated synthesis success rates, (3) analytical purity thresholds, (4) raw and processed

microplate assay metrics (Z'), and (5) total wall-clock time and cost per compound.

Future research should focus on open benchmark environments containing raw analytical spectra, replicate-level cellular assays, explicit right- and left-censoring bounds, and separate annotations for cytotoxicity, aggregation, assay interference, and other invalidating mechanisms. Prospective head-to-head studies comparing RL against Bayesian optimization and expert medicinal chemistry design under fixed experimental budgets will be required to evaluate the real-world efficiency of autonomous closed-loop discovery.

5. Conclusion

Closed-loop reinforcement learning represents a promising framework for experimentally grounded pharmaceutical discovery, but its value depends on the quality of the information returned from physical experiments. A central challenge is the biological feedback gap: delayed and variable measurements must be distinguished from statistically censored observations and from pharmacologically confounded or invalid readouts before they are allowed to influence learning.

The proposed four-gate framework—operational synthesizability, analytical identity and purity verification, biological assay qualification with explicit separation of censoring from confounding, and governed policy updating—provides a prospective architecture for reducing error propagation from physical experiments into generative policy updates.

Under supervised autonomy, closed-loop reinforcement learning provides a testable framework for integrating computational molecular design with experimentally grounded decision-making. Its pharmaceutical value should ultimately be judged not by virtual throughput, but by prospective evidence that it improves the reproducibility, efficiency, and developability of experimentally validated candidates.

This paper is important because it establishes that automated chemical synthesis and inline analytics have reached comparatively advanced levels of maturity, whereas functional biological screening remains a major and comparatively underappreciated translational bottleneck.

References

- [1] Meyers J, Fabian B, Brown N. De novo molecular design and generative models. *Drug Discov Today*. 2021;26(11):2707–2715. doi:10.1016/j.drudis.2021.05.019.
- [2] Olivecrona M, Blaschke T, Engkvist O, Chen H. Molecular de-novo design through deep reinforcement learning. *J Cheminform*. 2017;9(1):48. doi:10.1186/s13321-017-0235-x.
- [3] Coley CW, Jin W, Rogers L, et al. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chem Sci*. 2019;10(2):370–377. doi:10.1039/C8SC04228D.

- [4] Schwaller P, Laino T, Gaudin T, et al. Molecular Transformer: A model for uncertainty-calibrated chemical reaction prediction. *ACS Cent Sci*. 2019;5(9):1572–1583. doi:10.1021/acscentsci.9b00576.
- [5] Burger B, Maffettone PM, Gusev VV, et al. A mobile robotic chemist. *Nature*. 2020;583(7815):237–241. doi:10.1038/s41586-020-2442-2.
- [6] Koscher BA, Canty R, McDonald M, et al. Autonomous, multiproperty-driven molecular discovery: From predictions to measurements and back. *Science*. 2023; 382(6677): eadi1407. doi:10.1126/science.adi1407.
- [7] Canty, R. B.; Abolhasani, M. The Past, Present and Future of Self-Driving Laboratories. *Nat. Rev. Chem*. 2026, 10 (8), 523–537. <https://doi.org/10.1038/s41570-026-00847-2>.
- [8] Boiko DA, MacKnight R, Kline B, Gomes G. Autonomous chemical research with large language models. *Nature*. 2023;624(7992):570–578. doi:10.1038/s41586-023-06792-0.
- [9] Korshunova M, Huang N, Capuzzi S, et al. Generative and reinforcement learning approaches for the automated de novo design of bioactive compounds. *Commun Chem*. 2022; 5: 129. doi:10.1038/s42004-022-00733-0.
- [10] Born J, Baldassari C, Grabocka D, et al. Toward fully autonomous closed-loop molecular discovery—A case study on JAK targets. *ChemRxiv*. 2026. Preprint. doi:10.26434/chemrxiv-2026-q7xdt.
- [11] Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. *Proc 33rd Int Conf Mach Learn (ICML)*. 2016; 48: 1050–1059.
- [12] Williams RJ. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach Learn*. 1992;8(3):229–256. doi:10.1007/BF00992696.
- [13] Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. Proximal policy optimization algorithms. *arXiv*. 2017; arXiv:1707.06347.
- [14] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning. *Nature*. 2015;518(7540):529–533. doi:10.1038/nature14236.
- [15] Arjona-Medina JA, Gillhofer M, Widrich M, et al. RUDDER: Return decomposition for delayed rewards. *Adv Neural Inf Process Syst (NeurIPS)*. 2019; 32: 13566–13577.
- [16] Inglese J, Auld DS, Jadhav A, et al. Quantitative high-throughput screening: A titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proc Natl Acad Sci USA*. 2006;103(31):11473–11478. doi:10.1073/pnas.0604348103.
- [17] Zhang JH, Chung TDY, Oldenburg KR. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J Biomol Screen*. 1999;4(2):67–73. doi:10.1177/108705719900400206.
- [18] Juchli, D. (2022). SiLA 2: The Next Generation Lab Automation Standard. In: Beutel, S., Lenk, F. (eds) Smart Biolabs of the Future. Advances in Biochemical Engineering/Biotechnology, vol 182. Springer, Cham. https://doi.org/10.1007/10_2022_204.
- [19] ASTM International. *ASTM E1947: Standard Specification for Analytical Information Markup Language (AnIML)*. West Conshohocken, PA: ASTM International; 2024.
- [20] Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Ann Intern Med*. 2018;169(7):467–473. doi:10.7326/M18-0850.
- [21] Du, Y., et al. “Machine Learning-Aided Generative Molecular Design.” *Nature Machine Intelligence*, 2024, doi:10.1038/s42256-024-00843-5.
- [22] Kanakala, Ganesh Chandan, et al. “Generative Artificial Intelligence for Small Molecule Drug Design.” *Current Opinion in Biotechnology*, vol. 89, 2024, article 103175. Elsevier, <https://doi.org/10.1016/j.copbio.2024.103175>.
- [23] Tom G, Schmid SP, Baird SG, et al. Self-Driving Laboratories for Chemistry and Materials Science. *Chem Rev*. 2024;124(16):9633–9732. doi:10.1021/acs.chemrev.4c00055.

Author Profile



Michele earned a Bachelor of Science in Nursing from the University of Pittsburgh and is pursuing a PhD in Pharmaceutical Science. She brings more than 20 years of experience in medical affairs, scientific communications, pharmaceutical leadership, clinical practice, and military experience. Across Hematology, Immunology, Respiratory, and Neuroscience, she has advanced strategic medical alignment, strengthened expert engagement, supported clinical development and global medical teams, and contributed to major congress and communication strategies for leading pharmaceutical companies. Her career began in clinical nursing and includes decorated service in the United States Air Force.