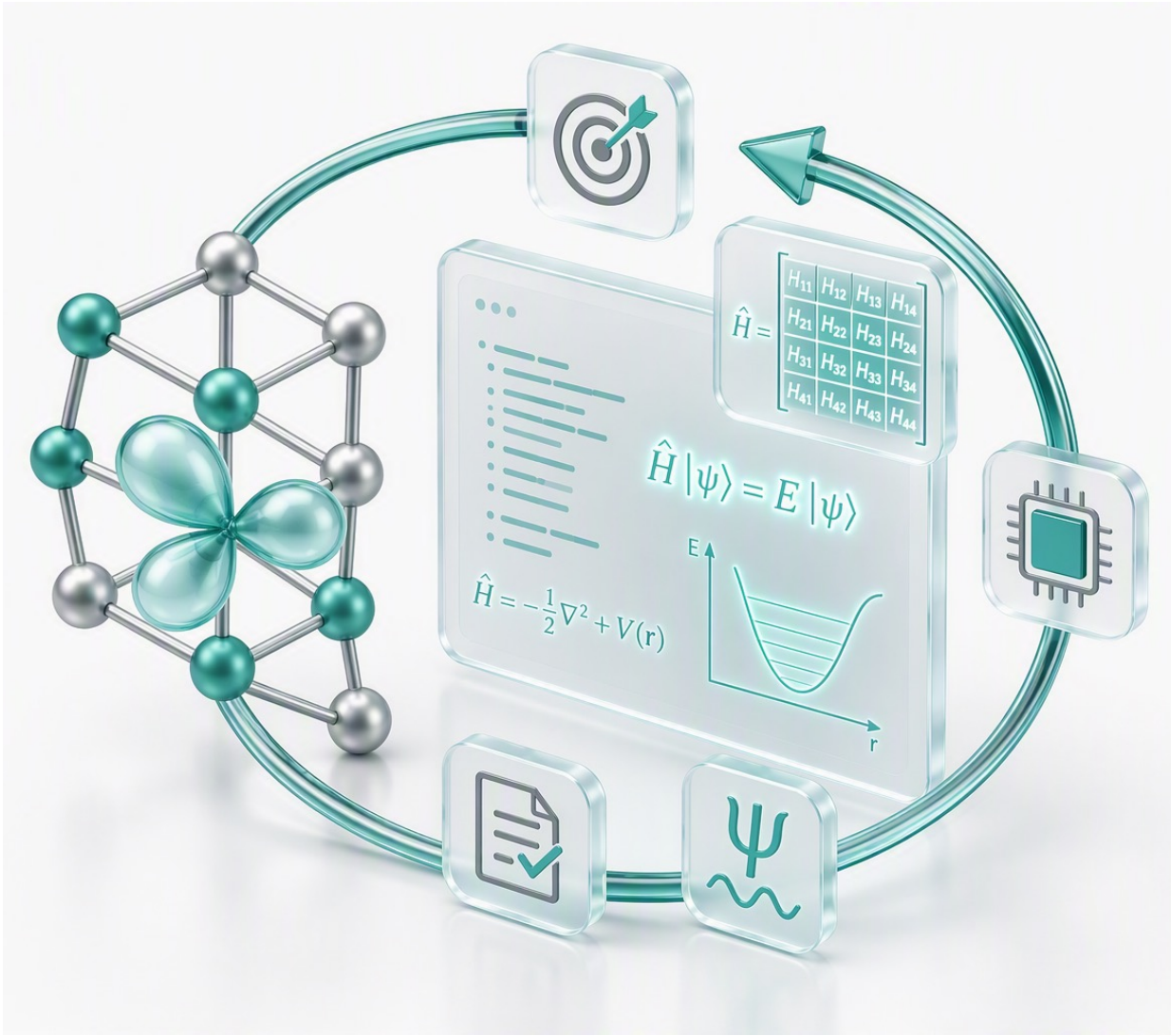
中国科学院大学 University of Chinese Academy of Sciences
前沿交叉科学学院 School of Advanced Interdisciplinary Sciences

Form Skills To Loop, Auto-Research, LLM+Wiki And Other Agentic Infra For A New Way of Materials Discovery

Yuan Jiao

School of Advanced Interdisciplinary Sciences, SAIS

University of Chinese Academy of Sciences



■ Form Skills Engineering to Loop Engineering

■ Auto-Research, LLM+Wiki

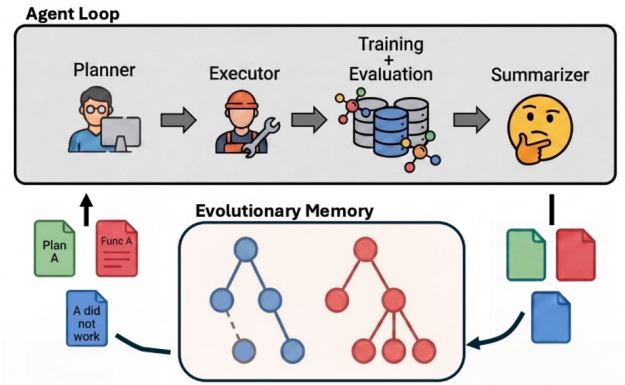
■ Auto-Research Agent for Materials

Github.com/QPengGroup

■ Auto-Research Agent for Materials Discovery

● **Functional**

- Computation Task Management
- Organic Materials Screening
- Simple Computational Error Correction
- Only Skills Based



● **Outlook**

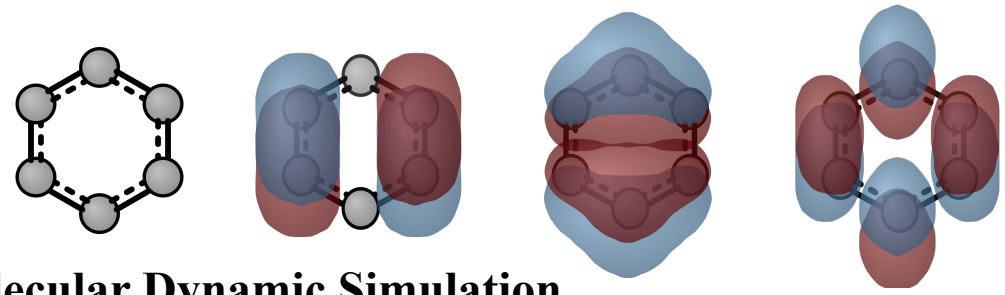
☐ **Long-Term(next vision):**

- Standard of HPC for Both Human user and Agent user
- Git tools for Code Management for Both Human and Agent
- LLM-Wiki for Knowledge Data Set

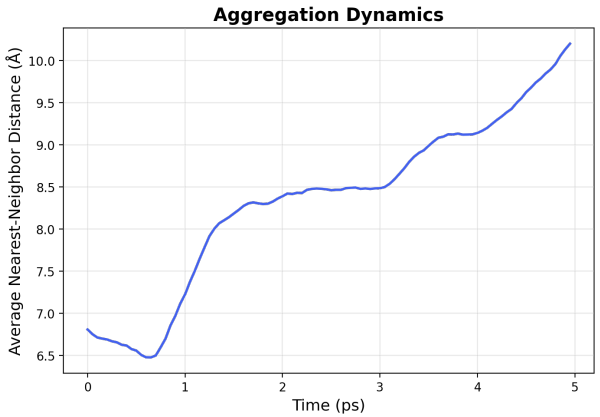
☐ **Short-Term(pre-release):**

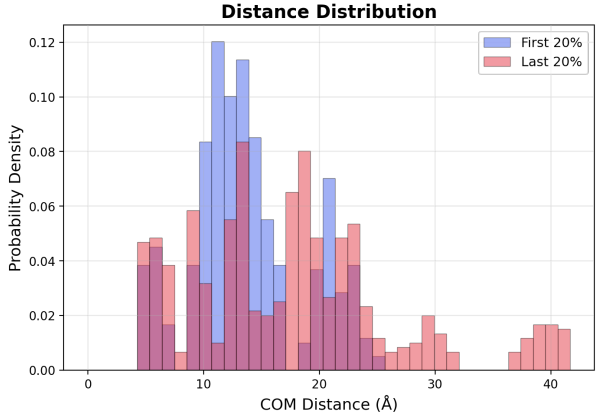
- Auto-Research for New Materials Discovery
- Experience and Knowledge form Auto-Research Loop
- Standard Workflow in Materials Discovery

■ Form Skills Engineering to Loop Engineering

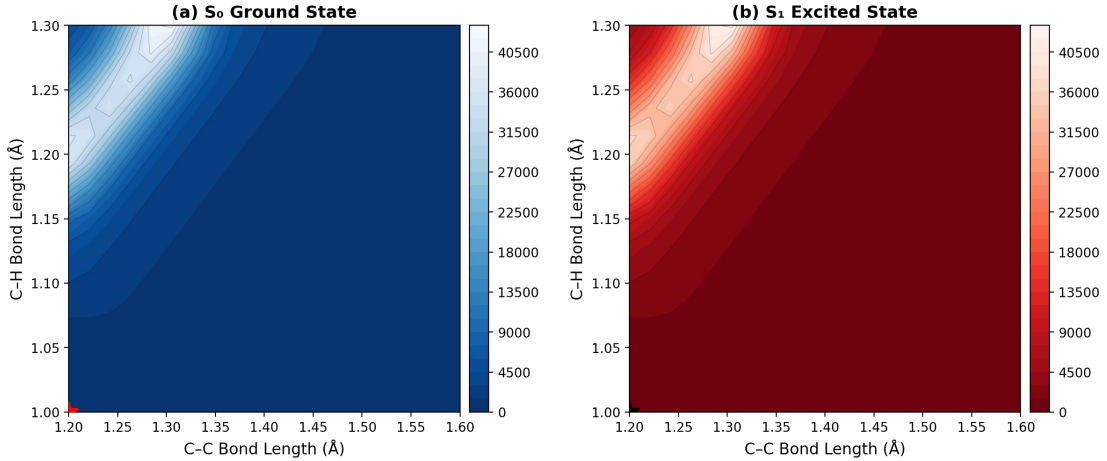


Molecular Dynamic Simulation

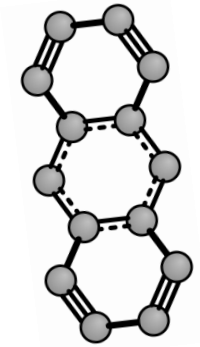




Potential Energy Surface

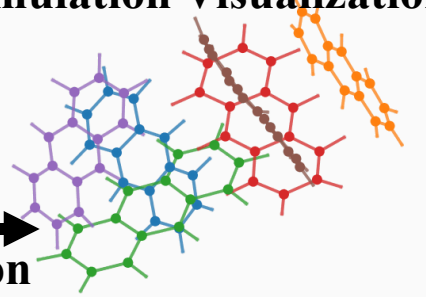


Molecular Dynamic Simulation Visualization

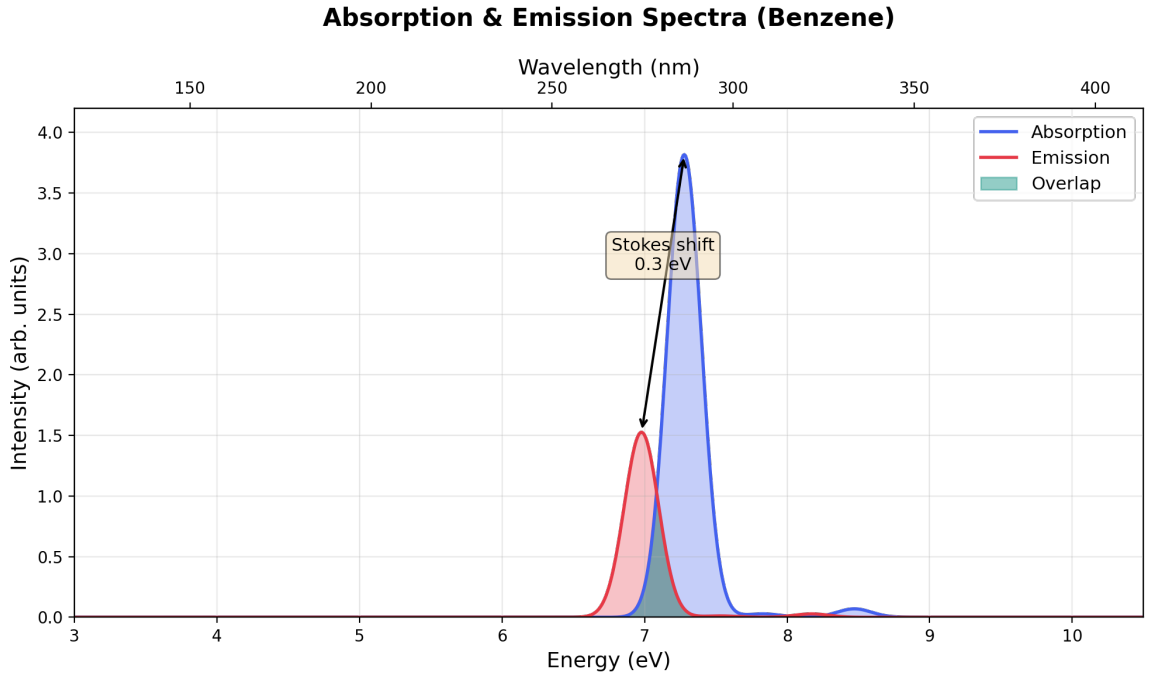


Random

Distribution

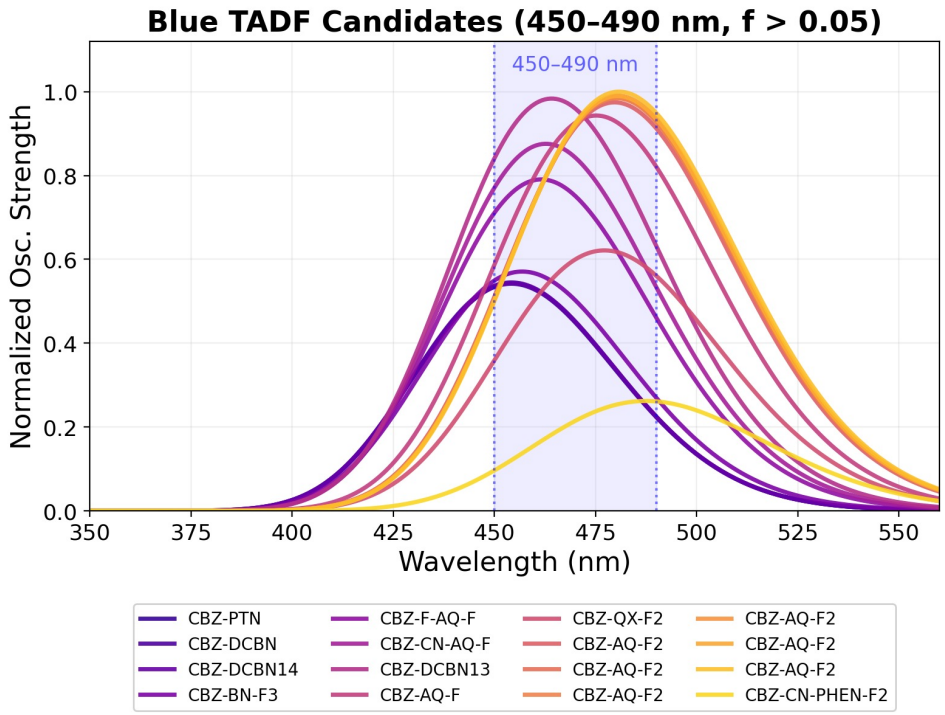


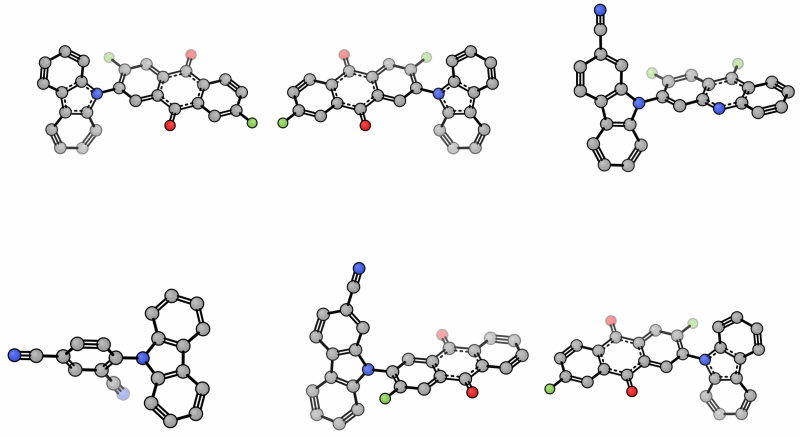
Absorption and Emission Spectra(Error !)



■ Auto-Research Agent for Materials Discovery

```
Stage 0: Initialization
  └ screening_workflow_initializer.py
     └ Confirm: topology, sample count (default 10k), interaction mode

Stage 1: Candidate Generation
  ├ build_da_topology_library.py  (D-A, D-A-D, A-D-A, D-pi-A...)
  └ SMILES → .xyz structure generation (export to manifest.csv)

Stage 2: xTB Pre-screening
  ├ Input: manifest.csv (idx, name, xyz_path)
  └ run_xtb_batch_manifest.py → xtb_progress.csv
      Required columns: idx, name, xyz_path, status, detail,
                        total_energy_eh, homo_lumo_gap_ev, normal_termination
  ⚠ If xTB ok=0 → STOP and diagnose

Stage 2B: TDDFT-xTB Wavelength Filter (optional)
  └ run_tddft_xtb_filter.py
      Filter: emission in target range, f ≥ threshold

Stage 3: TDDFT Validation (remote)
  ├ S0 optimization
  ├ Excited-state optimization (S1 for TADF/Fluorescence; T1 for Phosphorescence)
  ├ Vertical emission
  └ Gate: ΔE_ST < 0.10 eV, f ≥ 0.05, k_TADF > 10⁴ s⁻¹

Stage 4: Elite Candidates → downstream analysis
```





■ Auto-Research Agent for Materials Discovery

github.com/QPengGroup/ChemAgent

OpenChem
~

› Hello! Please introduce yourself

Thinking: The user is asking me to introduce myself. According to my system prompt, I should call myself OpenChem and describe myself as an interactive computational photophysics and organic optoelectronic-materials research agent.

你好！我是 **OpenChem**，一个专注于计算光物理与有机光电材料研究的交互式智能体。

我的核心研究方向包括：

- **有机光电材料**：TADF（热活化延迟荧光）、OLED 发光材料、激子动力学等
- **激发态机制**：吸收/发射光谱、内转换（IC）、系间窜越（ISC）、反向系间窜越（RISC）、电荷转移态等
- **分子设计与筛选**：高通量虚拟筛选、光物理性质预测、结构-性能关系分析
- **计算工作流**：基态 DFT → TDDFT 激发态计算 → 自旋轨道耦合/非辐射速率 → 速率常数提取（如 MOMAP 光物理服务适配器）

我也熟悉你在做的研究工作——等变图神经网络、range-separated hybrid 泛函优化、Hamiltonian 预测——并且可以协助：

- 计算化学输入文件的准备与验证（Gaussian、ORCA、Q-Chem、ABACUS 等）
- 结果解析与 Jablonski 图构建
- 文献与数据驱动的光物理机制解读
- 本地 PySCF 光物理工作流的执行（openchem_photophysics）

有什么光物理或材料计算方面的问题，随时可以开始！😊

■ Auto-Research Agent for Materials Discovery

一、关键能量数据 (ωB97X-D/def2-TZVP)

```
S₁ ─── 5.016 eV   (247 nm)   f = 0.0444   ππ* (HOMO→LUMO)
  │                    ΔE_ST = 0.931 eV = 7508 cm⁻¹
T₁ ─── 4.085 eV   (304 nm)   <S²> = 2.01   UKS
  │
S₀ ─── 0.000 eV
```

二、辐射速率 k_r (荧光)

参数	值
振子强度 f	0.0444
k_r	**7.46 × 10^8 s^{-1}**
τ_r (自然辐射寿命)	**1.34 ns**

> *^{1}Lb 态是弱允许的（对称性近禁阻），但由于激发能高（∝E^2），辐射速率仍然可观。*

三、系间窜越速率 k_ISC (S_1→T_1)

ΔE_ST = 0.93 eV　← 很大！TADF 要求 <0.1 eV

SOC 情景	SOC (cm^{-1})	λ=0.20 eV 时 k_ISC	结论
$^1ππ^*$→$^3nπ^*$ (El-Sayed 允许)	3	~7 × 10^{-18} s^{-1}	可忽略
$^1ππ^*$→$^3ππ^*$ (禁阻)	0.3	~7 × 10^{-20} s^{-1}	可忽略
N 原子极限 (73 cm^{-1})	73	~4 × 10^{-15} s^{-1}	可忽略

> *ISC 被 ΔE_ST = 0.93 eV 的 Marcus 势垒 (>1.4 eV) 完全抑制。苯胺不是 TADF 分子。*

四、内转换速率 k_IC (S_1→S_0)

能量间隙定律估算受制于大 ΔE 的指数衰减。实验上，苯胺的荧光寿命为：

- 气相 τ_f ≈ 4–10 ns → k_total ≈ 1.0–2.5 × 10^8 s^{-1}
- 与我们的 k_r = 0.75 × 10^8 s^{-1} 对比：
  - **k_nr(实验) ≈ 0.3–1.8 × 10^8 s^{-1}**
  - **荧光量子产率 Φ_f ≈ 0.3–0.75 (气相)**

五、磷光速率 k_ph (T_1→S_0)

参数	典型值
T_1 能量	4.09 eV (32950 cm^{-1})
k_ph (无重原子)	0.1–10 s^{-1}
τ_ph	0.1–10 s
实际 Φ_ph	<10^{-3} (极弱)

六、JablonskiGraph – 速率版

```
          k_r = 7.5×10⁸ s⁻¹  (τ=1.3 ns)
S₁  ──────────────────────────────→  S₀ + hν (247 nm)
 │
 │  k_ISC ≈ 0 (ΔE_ST=0.93 eV, 被 Marcus 势垒阻断)
 │
 ↓          k_ph ~ 0.1-10 s⁻¹
T₁  ──────────────────────────────→  S₀ + hν' (极弱)
 │
 │  k_nr(IC) ≈ 1-2×10⁸ s⁻¹ (实验估计)
 │
 ↓
S₀
```

■ Auto-Research Agent for Materials Discovery

✅ 苯胺光物理研究 – 完整报告已生成

报告保存在 /Users/jiaoyuan/aniline_report/REPORT.md。

📁 项目产出总览

```
aniline_report/
├── REPORT.md                ← 完整研究报告（本文）
├── wb97xd/
│   ├── results_wb97xd_tzvp.json    ωB97X-D 激发态数据
│   └── run.log                     计算日志
├── solvent/
│   ├── results.json                PCM 溶剂效应数据
│   └── run.log
├── rates/
│   └── run.log                     速率计算日志
└── aniline_photophysics_out/
    └── results_corrected.json      B3LYP/STO-3G 预览
```

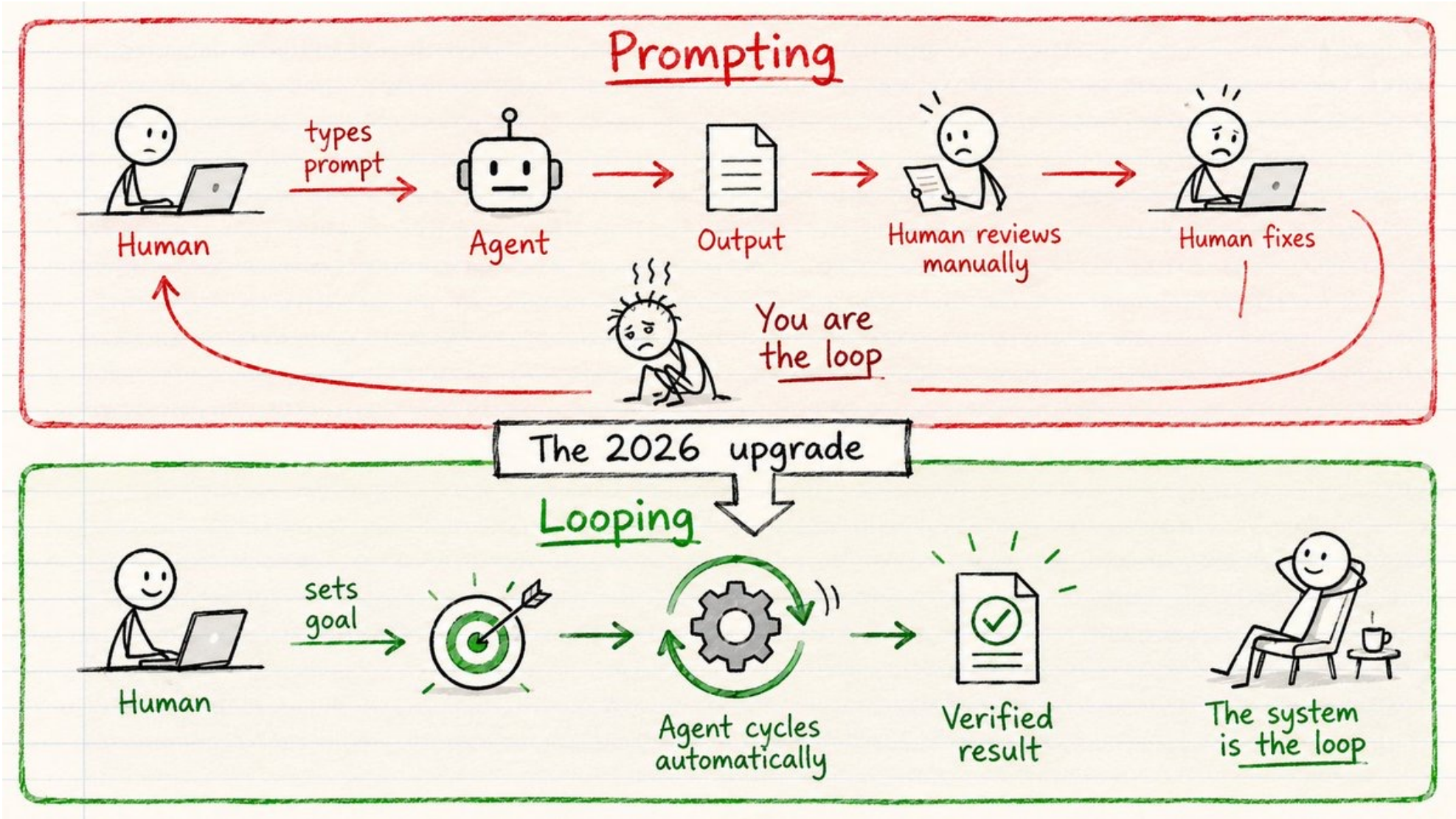
📊 三阶段研究总结

阶段	内容	关键结果
1. 激发态	B3LYP/STO-3G → ωB97X-D + CAM-B3LYP/TZVP	S_1=5.02 eV (^{1}Lb), S_2=5.92 eV (^{1}La), Δ vs exp ~0.8 eV
2. 溶剂	LR-PCM: Gas → cyclohexane → MeCN → H_2O	PCM 红移 ~0.07 eV; H 键效应缺失导致与实验蓝移符号相反
3. 速率	k_r, k_ISC, k_IC, k_ph	τ_r=1.3 ns, ISC 被 ΔE_ST=0.93 eV 阻断, IC 主导非辐射

🎯 核心结论

> *苯胺 $S_1 \to T_1$ 间隙为 0.93 eV, ISC 在动力学上完全阻断, 不适合独立作为 TADF 分子。但作为富电子给体, 可构建 D-A 型 TADF 候选物——这是下一步分子设计的自然延伸。*

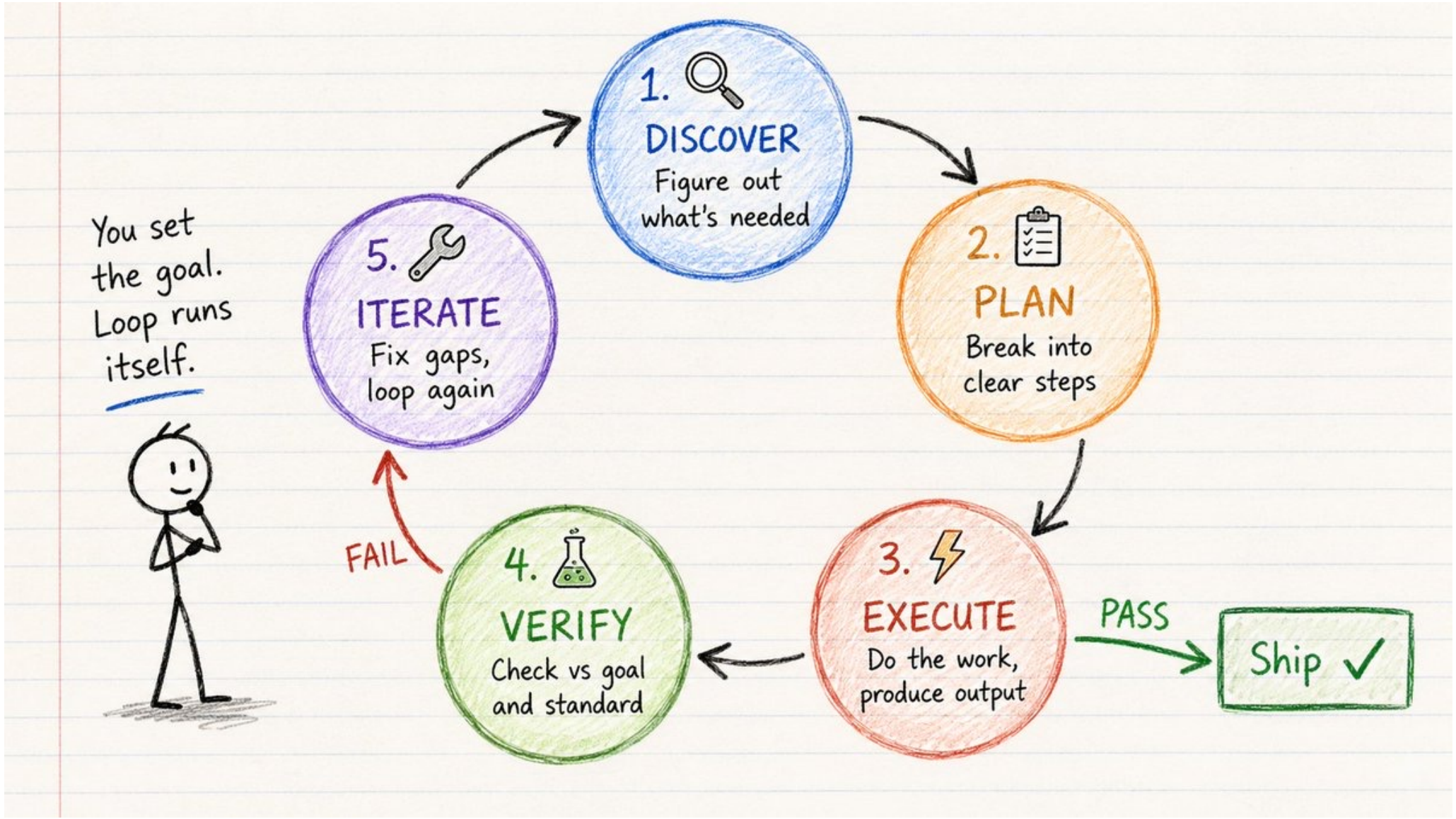
■ Form Skills Engineering to Loop Engineering



Ref: x.com/sairahul1

■ Auto-Research and LLM+Wiki

github.com/karpathy/autoresearch



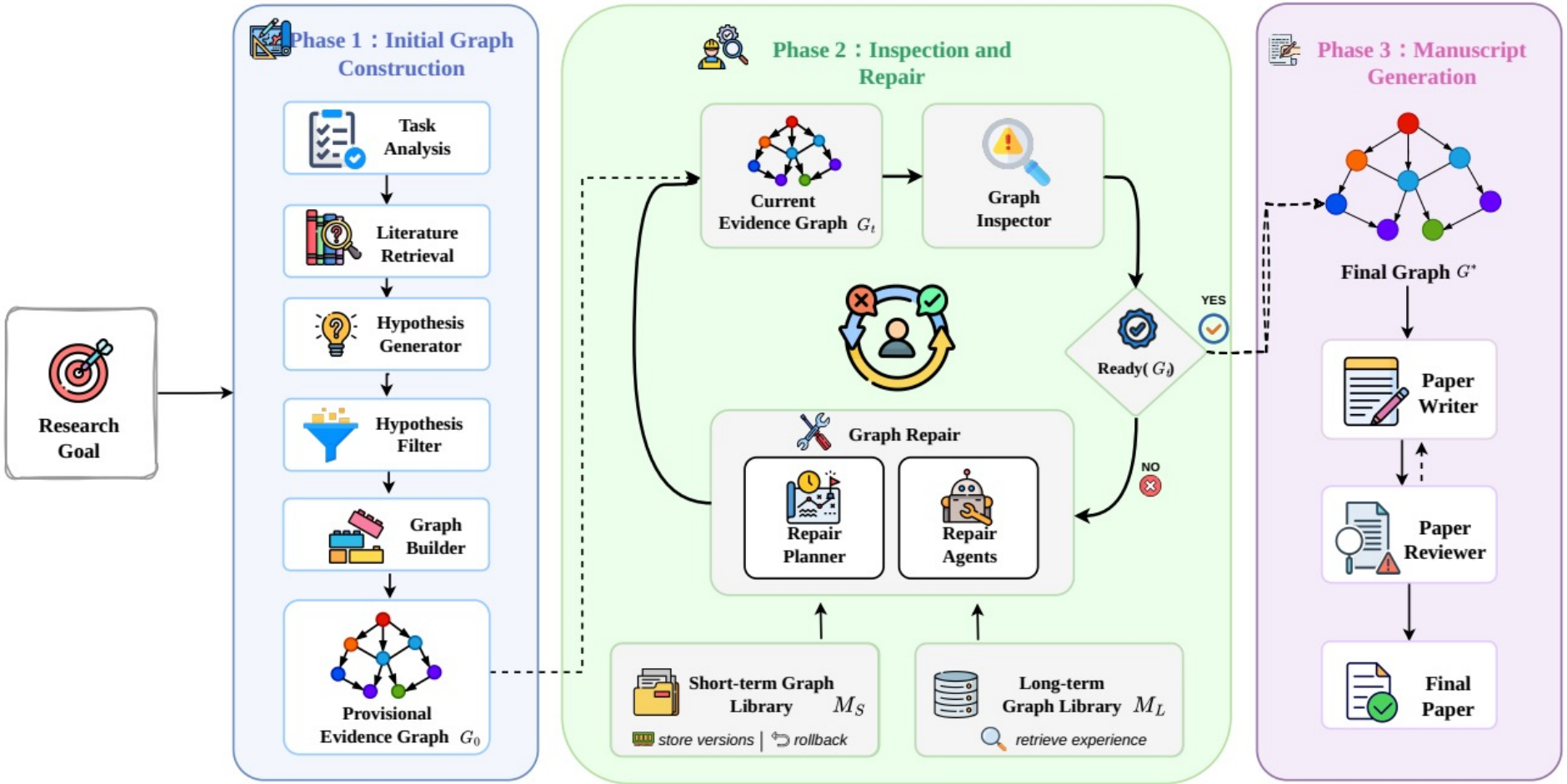
■ Auto-Research and LLM+Wiki

EviGraph: Evidence-Guided Autonomous Research Agents

Zhenjiang Ren[1,2], Ruiji Li[1], Xujing Zhang[3], Ziliang Pang[1], Shuo Ren[1]*, Jiajun Zhang[1,2,4]*

[1]Institute of Automation, Chinese Academy of Sciences
[2]School of Artificial Intelligence, University of Chinese Academy of Sciences
[3]Hong Kong Baptist University [4]Wuhan AI Research
{renzhenjiang2024, liruiji2026, ziliang.pang, shuo.ren}@ia.ac.cn,
24267368@life.hkbu.edu.hk, jjzhang@nlpr.ia.ac.cn

Arxiv: 2608.04738v2



■ Auto-Research and LLM+Wiki

■ Auto-Research Agent for Materials and Methods

Automated Discovery of Algorithms for Molecular Electronic Structure Calculations Using Physics-Informed Program Synthesis

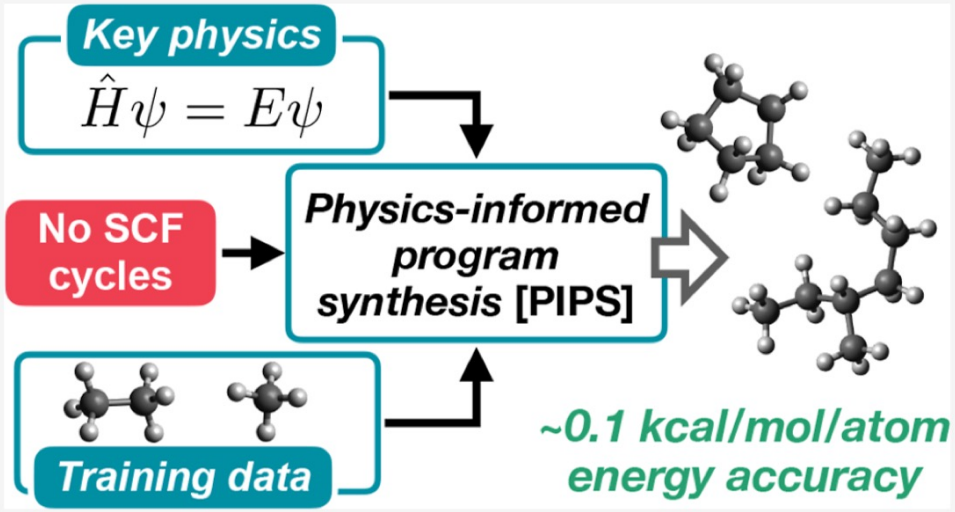
Kyle Acheson, Rastislav Turanyi, and Scott Habershon*

Cite This: *J. Am. Chem. Soc.* 2026, 148, 11991−12000

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We demonstrate a physics-informed program synthesis (PIPS) approach that can be used to identify entirely new algorithms that approximate the results of single-reference electronic structure approaches like Hartree−Fock (HF) and density-functional theory (DFT)—but without *any* self-consistent field iterations at all. Our PIPS strategy exploits the fact that the eigenvectors of the Fock matrix **F** (or Kohn−Sham matrix **K**) are the same as the eigenvectors of a broad class of matrix functions, $f(\mathbf{F})$. As a result, PIPS can be used to seek matrices **M** that yield the same molecular orbital coefficients as converged HF or DFT calculations. We demonstrate this approach by generating new algorithms that accurately predict total energies for a series of heterodiatomic molecules (LiCl, LiF, NaCl, NaF) and C_1−C_4 hydrocarbons; further simulations of C_8−C_{20} alkane species demonstrate further transferability and efficiency of the resulting algorithms. We obtain novel algorithms that can reproduce HF or DFT energies to within 0.1 kcal/mol/atom while requiring only a single matrix-diagonalization operation, rather than an iterative self-consistent field convergence. The approach demonstrated here could be similarly applied to more complex wave function ansatze, opening an interesting optimization-based pathway to identifying accurate yet efficient algorithms for molecular quantum chemistry.



■ Auto-Research Agent for Materials and Methods

➢ 1: Neural Network Atom Basis Designing and Training

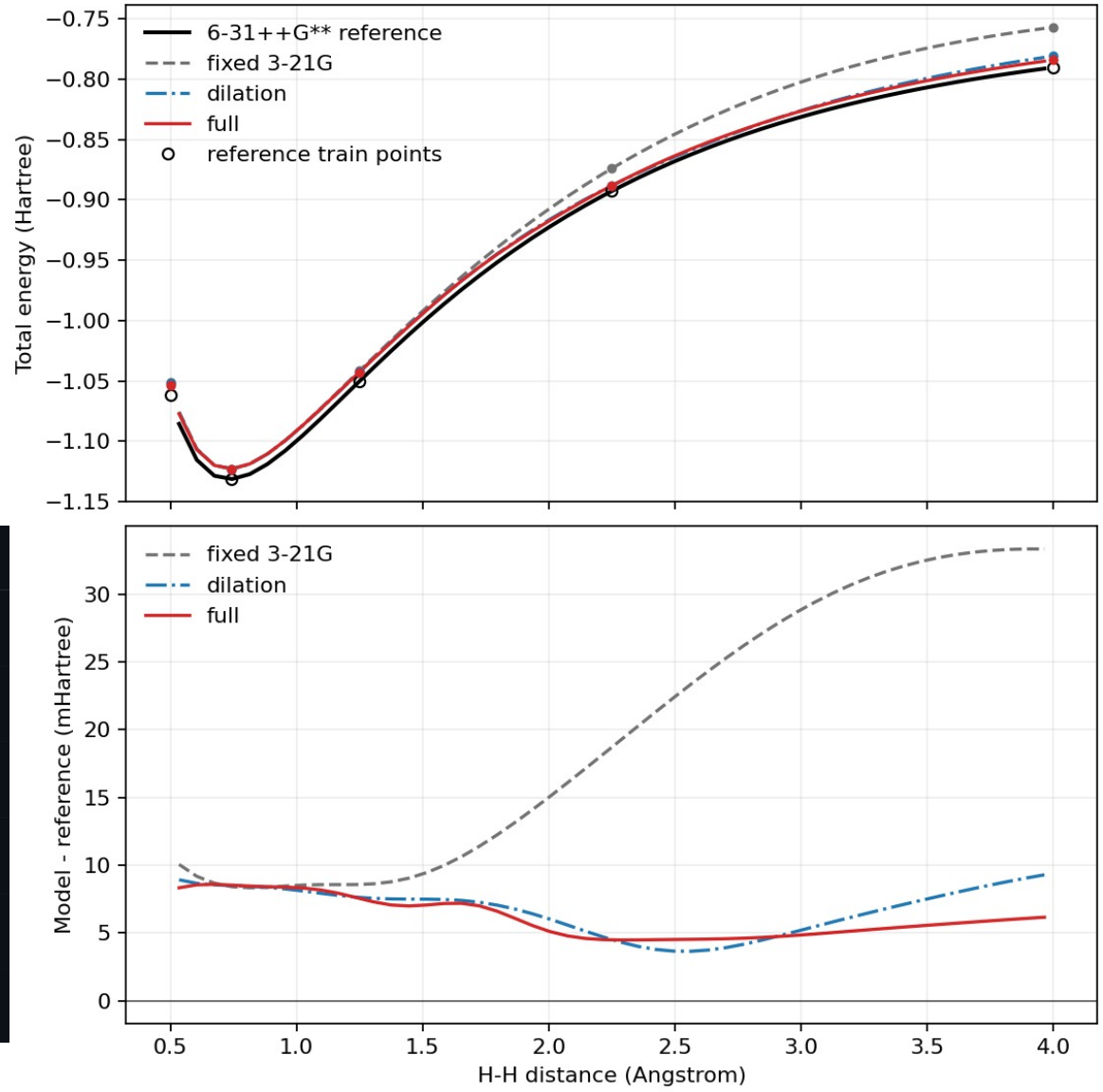
Step 1: Literature Searching

Step 2: Neural Network Architecture Designing

Step 3: Working in the Loop

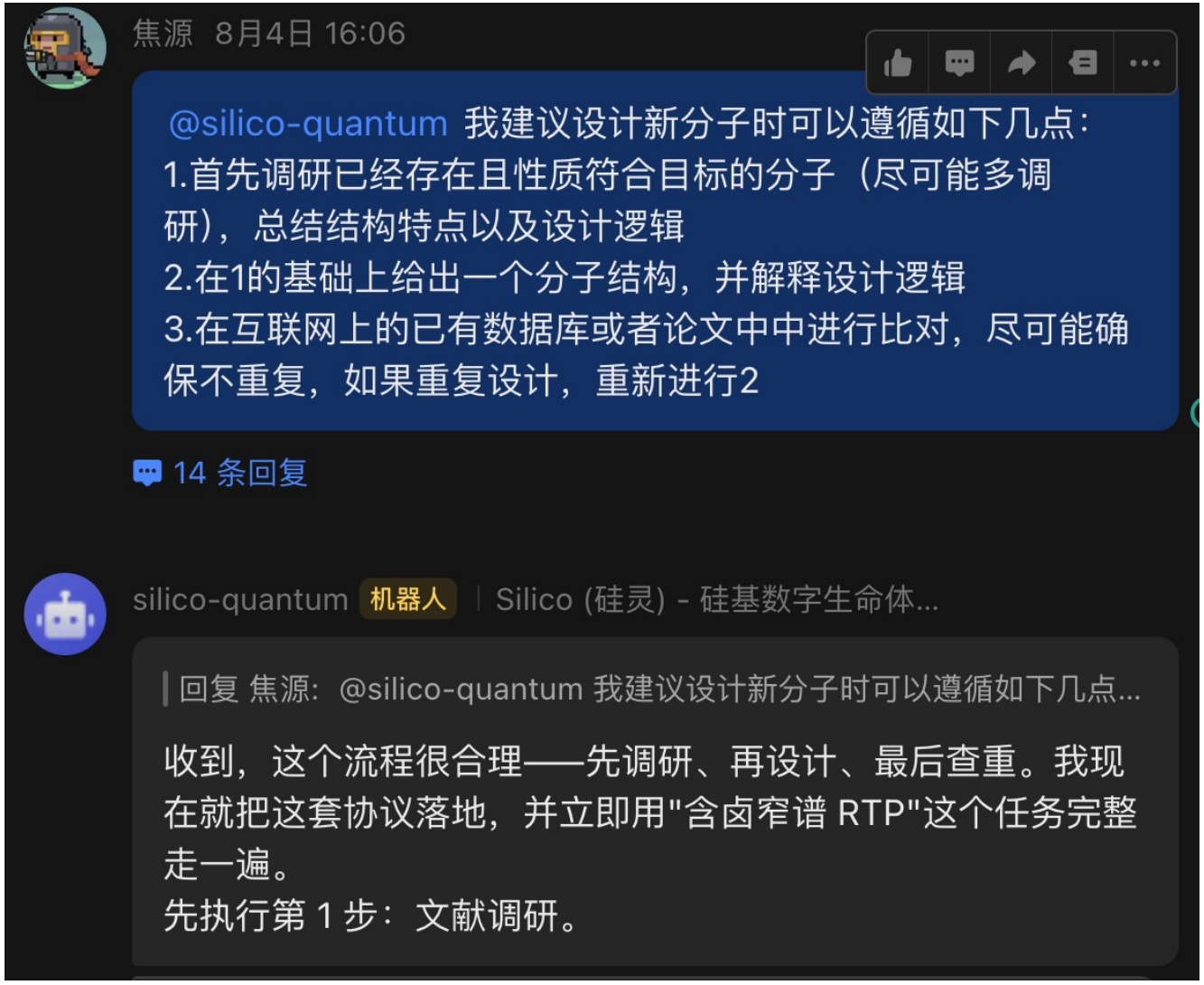
$$u_{A,nl}^{raw}(r) = r^{l+1} f_{cut}(r) \left[q_{Z_A,nl}^{ref}(\rho_{A,nl} r) + r^2 \sum_k a_{A,nlk} B_k(x(r)) \right]$$

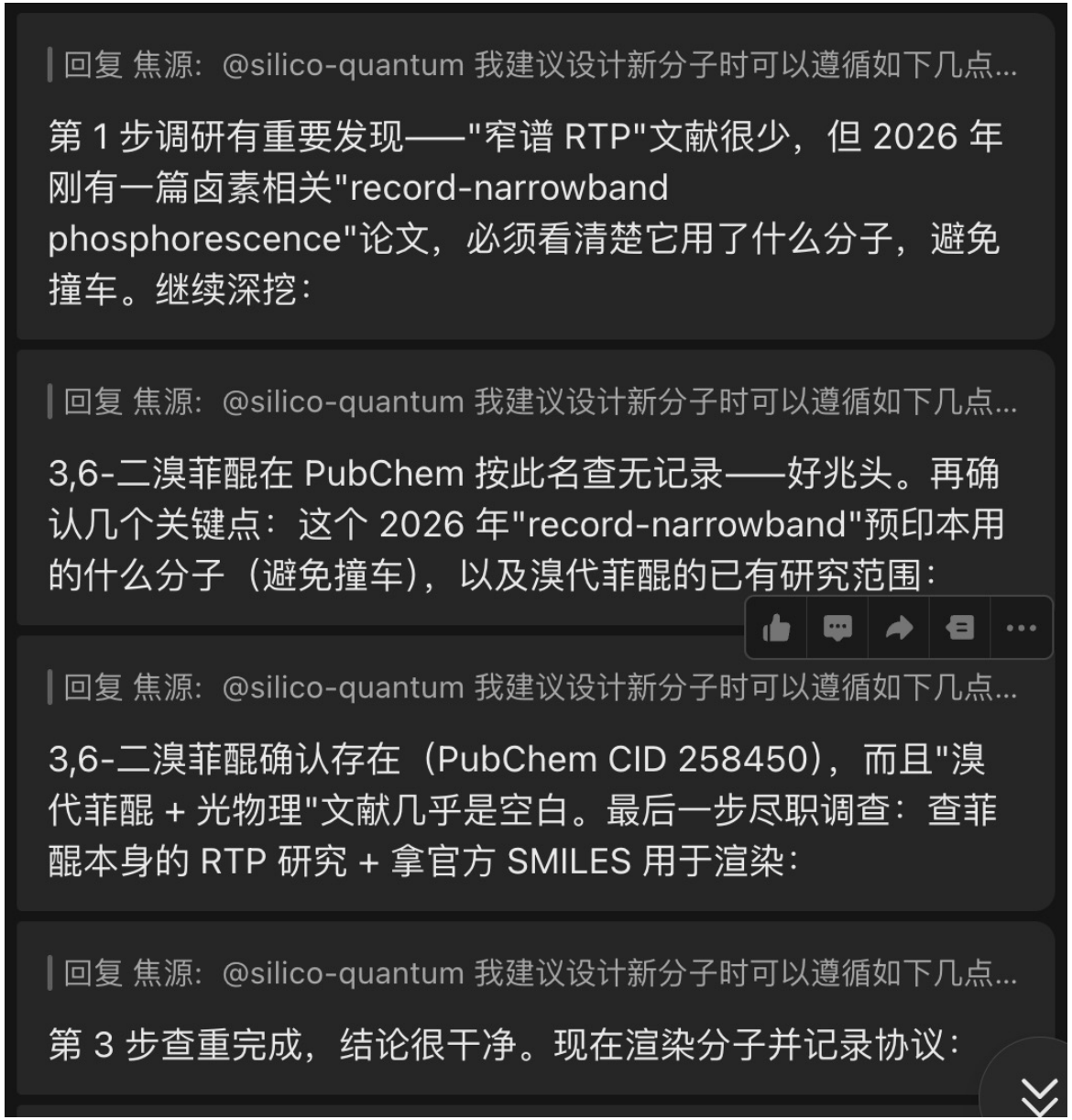
测试集指标	固定 3-21G	dilation	full
能量 MAE (mEh)	19.700	6.716	**6.139**
相对曲线 MAE (mEh)	13.611	2.687	**1.261**
NPE (mEh)	24.988	5.636	**4.103**
平均 density relative-L2	0.10880	**0.04886**	0.05457
平均 half-L1	0.14065	0.05258	**0.04734**
平均 operator spillage	0.02233	0.005189	**0.002371**



■ Auto-Research Agent for Materials and Methods

➢ 2:Materials Designing or Searching





■ Auto-Research Agent for Materials and Methods

协议已采纳并记录。完整走完一遍，结果如下：

第 1 步 · 调研（窄谱 RTP 已知体系）

文献极少，核心设计逻辑只有几条：

- **MR 效应**：B/N 闭环分子共振 → 窄谱（Nat. Commun. 2024, 10.1038/s41467-024-48856-3），但无卤素
- **刚性 1,2-二酮**：金属-free、快速窄谱 RTP（Chem. Sci. 2024, 10.1039/d4sc02841d），无卤素
- **重原子/卤键**：分子内卤键把 UOP 量子效率推到 52%（Angew 2020）；2026 年刚出预印本，用分子内卤素非共价相互作用（XNCI）实现 record FWHM 15.3 nm（ChemRxiv 10.26434/chemrxiv.15005522/v1）
- 溴代呫吨酮 RTP 已被南大团队做掉（手性液晶基质路线）
- 结论：**"卤素 × 窄谱 RTP"交叉点几乎空白，2026 才有人用 XNCI 机理切入**

第 2 步 · 新分子：3,6-二溴-9,10-菲醌（3,6-DBPQ）

SMILES：`C1=CC2=C(C=C1Br)C3=C(C=CC(=C3)Br)C(=O)C2=O`

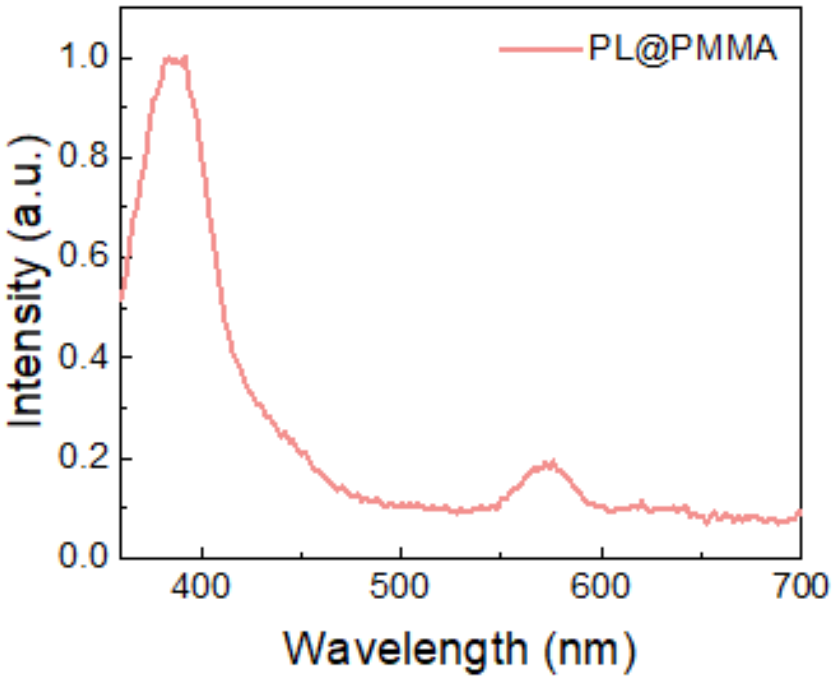
设计逻辑 = 融合两条被验证的路线：

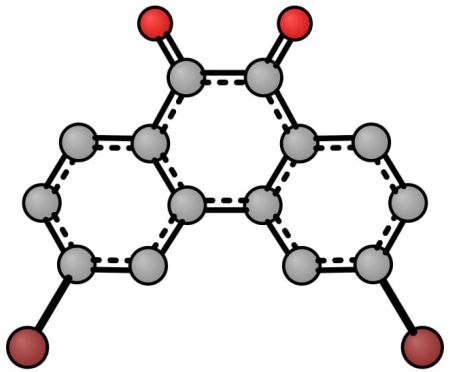
1. **1,2-二酮刚性骨架**（菲醌的邻二羰基 n-π* 态，强 ISC）→ 直接继承 Chem. Sci. 2024 的窄谱逻辑
2. **3,6-对称双 Br** → 重原子增强 SOC + 潜在分子内 Br···O=C 卤键锁定（对应 Angew 2020 / 2026 XNCI 机理），对称取代不展宽

预期黄绿区 RTP（T1≈2.2 eV，~550–580 nm），晶态 FWHM 目标 <40 nm（待 MOMAP 验证）

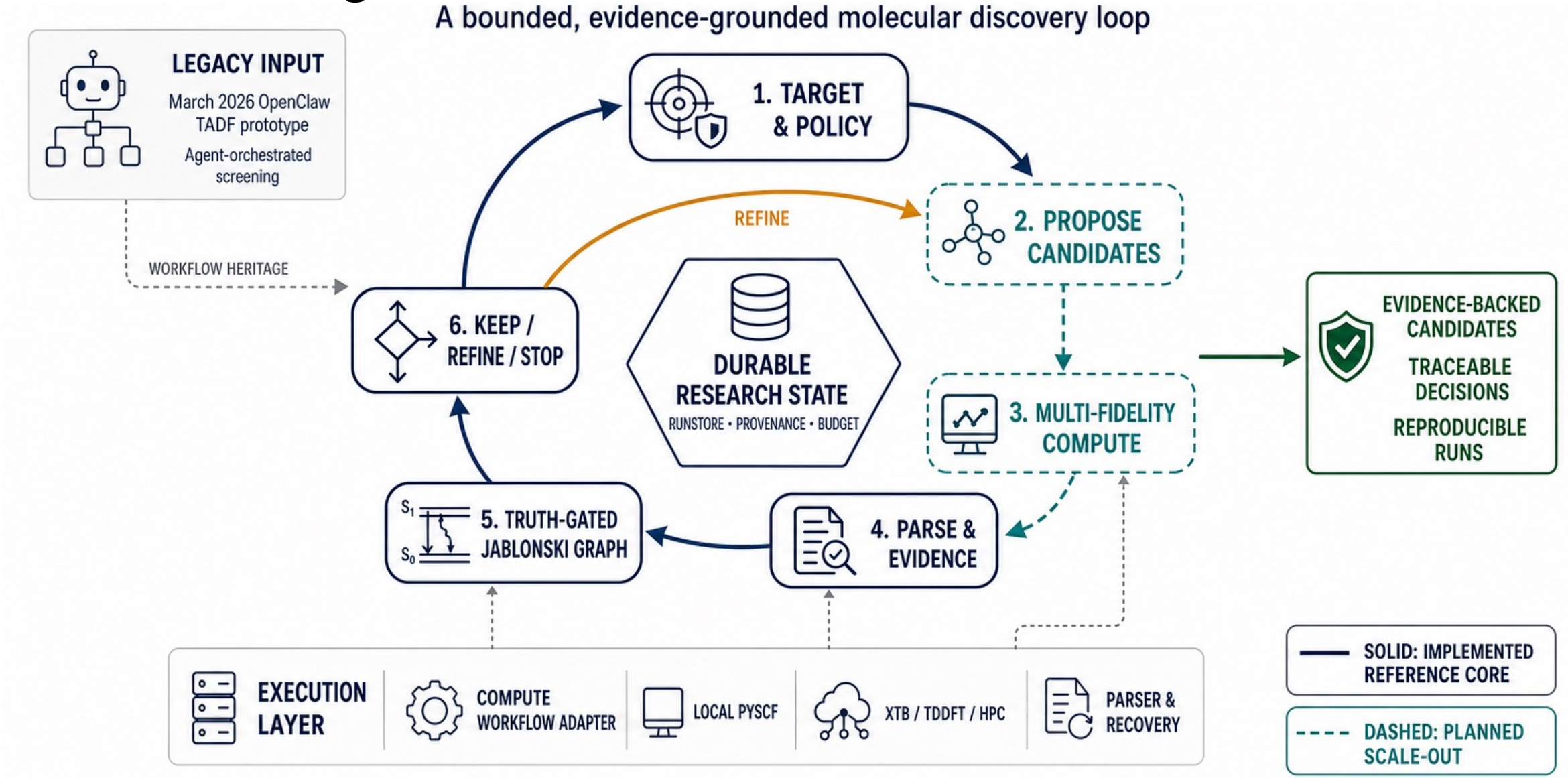
第 3 步 · 查重结果 ✅

- 化合物本身已知：CAS 53348-05-3，PubChem CID 258450，商业可购，早年入 NCI 筛选库
- 但**光物理完全空白**：菲醌磷光文献全部止于 1967–2005 年的 77K 溶液研究；溴代菲醌只有 1964 合成、2022 聚合物、2025 抗菌活性三篇——零 RTP、零固态发光
- 与 2026 年 XNCI 预印本骨架（卤代芳烃分子内非共价）完全不同，不撞车
- 结论：作为"含卤窄谱 RTP 发射体"是新设计，且合成一步可及（菲醌直接溴化）





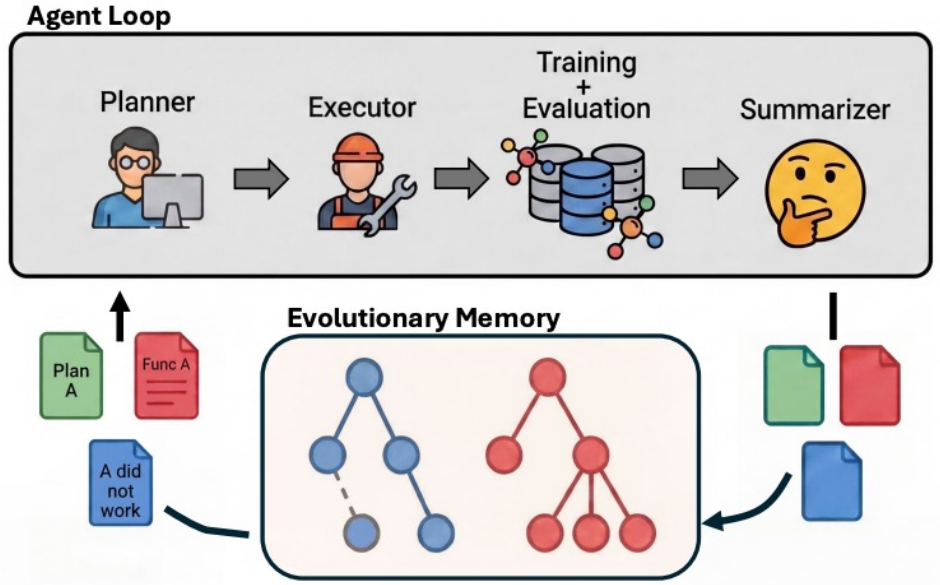
■ Auto-Research Agent for Materials and Methods



https://github.com/QPengGroup/oh-my-auto-research

■ Auto-Research Agent for Materials Discovery

- **Functional**
  - Computation Task Management
  - Organic Materials Screening
  - Simple Computational Error Correction
  - Only Skills Based



oh-my-auto-research

Standards for letting AI agents conduct real research — autonomously, verifiably, and without trusting them to grade their own homework.

This repository hosts three companion specification documents:

Document	Layer	What it standardizes
The Auto-Research Standard	Methodology	The four-stage pipeline (topics → db → validator → run), its gates, artifact contracts, and anti-gaming requirements for agent-driven research
The Agent Skill Standard	Packaging	How skills (`SKILL.md` + bundled resources) are authored, discovered, tested, and shipped
The MCP Server Standard	Capability	How MCP servers expose tools to agents: tool contracts, errors, security, and agent-facing evaluation

- **Auto-Research**
  - Standard Architecture Document
  - Basic Tools for Agents During Working
  - Evolution Edge
  - State Machine For Loop Working

Github.com/QPengGroup/oh-my-auto-research