

大语言模型驱动分子模拟智能体：从自动化到智能化

韩润一^{1,2}, 徐晓放^{1,2}, 黄剑兴³, 李路路¹, 叶茂¹, 刘中民¹

(¹ 中国科学院大连化学物理研究所, 辽宁 大连 116023; ² 中国科学院大学, 北京 100049; ³ 华为技术有限公司, 深圳 广东 518129)

摘要: 传统分子模拟方法在处理化工、材料等复杂体系时, 常受限于对专家经验的高度依赖、操作流程繁琐与结果可重复性差等瓶颈, 难以满足高通量筛选与工业应用的需求。现有的基于脚本的工作流自动化方案虽实现了流程标准化, 但其静态预设的执行逻辑无法应对模拟过程中出现的非收敛、力场适配偏差及采样效率低下等动态异常, 缺乏自主的科学决策能力。近年来, 大语言模型 (Large Language Model, LLM) 的引入为分子模拟提供了新的逻辑推理引擎与人机交互范式。通过构建集成自然语言指令、领域知识检索增强 (Retrieval-augmented Generation, RAG)、模拟流程自动编排及代码自适应生成的智能体, 实现了模拟任务从“静态脚本执行”到“动态自适应推演”的转变。本文综述了大语言模型赋能的分子模拟智能体最新进展, 重点阐述了其在晶体结构自动构建、力场参数智能推荐、反应路径搜索、自由能计算流程优化及多软件协同仿真中的关键技术路径, 分析了智能体在意图理解、复杂任务原子化拆解、全流程工具链编排及异常闭环修正等方面的应用实践, 并展望了未来在多模态实验数据融合、基于主动学习的力场自主进化及可信计算框架方面的发展方向, 旨在为计算化学与化工过程模拟的智能化升级提供理论参考与技术指引。

关键词: 分子模拟; 计算化学; 人工智能; 智能体; 大语言模型; 工作流; 计算机模拟

中图分类号: TQ 021.9

文献标志码: AA

文章编号: 0438-1157 (XXXX) XX-0001-19

Large language model-driven agents for molecular simulation: from automation to intelligence

HAN Runyi^{1,2}, XU Xiaofang^{1,2}, HUANG Jianxing³, LI Lulu¹, YE Mao¹, LIU Zhongmin¹

(¹ Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, Liaoning, China; ² University of Chinese Academy of Sciences, Beijing 100049, China; ³ Huawei Technologies Co., Ltd., Shenzhen 518129, Guangdong, China)

Abstract: Traditional molecular simulation methods, when applied to complex systems in chemical engineering and materials science, are often limited by bottlenecks such as a high reliance on expert experience, cumbersome operational procedures, and poor result reproducibility. These limitations hinder their ability to meet the demands of high-throughput screening and industrial applications. Although existing script-based workflow automation schemes have achieved process standardization, their statically preset execution logic fails to address dynamic anomalies during simulations—such as non-convergence, force field mismatches, and low sampling efficiency—thereby lacking autonomous scientific decision-making capabilities. In recent years, the integration of Large

收稿日期: 2026-04-29 修回日期: 2026-05-21

通信作者: 李路路(1992—), 男, 博士, 助理研究员, lilulu@dicp.ac.cn; 叶茂(1973—), 男, 博士, 研究员, maoye@dicp.ac.cn

第一作者: 韩润一(2002—), 女, 博士研究生, hanryi@dicp.ac.cn

基金项目: 国家自然科学基金项目(22503100); 中国科学院项目(XDA0490000)

引用本文: 韩润一, 徐晓放, 黄剑兴, 李路路, 叶茂, 刘中民. 大语言模型驱动分子模拟智能体: 从自动化到智能化[J]. 化工学报, XXXX, XX (XX): 1-19

Citation: HAN Runyi, XU Xiaofang, HUANG Jianxing, LI Lulu, YE Mao, LIU Zhongmin. Large language model-driven agents for molecular simulation: from automation to intelligence[J]. CIESC Journal, XXXX, XX(X): 1-19

Language Models (LLMs) has provided a novel logical reasoning engine and human-computer interaction paradigm for molecular simulation. By constructing agents that integrate natural language instructions, Retrieval-augmented generation (RAG) for domain knowledge, automated orchestration of simulation workflows, and adaptive code generation, a paradigm shift in simulation tasks from "static script execution" to "dynamic adaptive deduction" has been realized. This paper reviews the latest advancements in LLM-empowered molecular simulation agents. It highlights key technical pathways in the automatic construction of crystal structures, intelligent recommendation of force field parameters, reaction pathway searches, optimization of free energy calculation workflows, and collaborative simulations across multiple software platforms. Furthermore, it analyzes the practical applications of these agents in intent understanding, the atomic decomposition of complex tasks, full-process toolchain orchestration, and closed-loop anomaly correction. Finally, this review envisions future development directions, including multimodal experimental data fusion, the autonomous evolution of force fields based on active learning, and trusted computing frameworks, aiming to provide theoretical references and technical guidance for the intelligent upgrade of computational chemistry and chemical process simulations.

Keywords: molecular simulation; computational chemistry; artificial intelligence; agent; large language model; workflow; computer simulation

引 言

分子模拟作为连接微观动力学与宏观系统性质的桥梁,已成为化学工程、生物物理学与材料科学中不可或缺的核心工具^[1],它能够在原子尺度上揭示物质的构效关系、动态过程与反应机理,从而指导新药设计^[2-3]、新材料开发^[4]和催化剂设计^[5]。

然而,当前传统分子模拟流程在应对复杂真实体系的高通量筛选与多尺度预测时,仍受限于工具本身的高门槛与流程的强经验依赖性^[6]。尤其在催化反应网络解析、多相体系界面传质以及从电子结构到反应器尺度的跨尺度衔接等化工典型场景中,其建模与求解高度依赖经验判断与反复试错。从体系模型的构建、物理参数的选择,到计算结果解析及后续方案迭代优化,均需研究者具备深厚的计算化学功底与丰富的实践经验^[7]。这种强经验驱动的模式,不仅导致模拟流程繁琐、周期冗长,更使得研究结果的可重复性与鲁棒性受个体差异影响较大^[8],难以构建标准化、可复用、可迁移的自动化计算流程。

为缓解流程繁琐问题,近年来发展了基于Python脚本(如ASE(Atomic Simulation Environment,原子级模拟Python环境)^[9],Pymatgen(Python Materials Genomics,材料科学分析Python库)^[10])和工作流管理器^[11]的自动化方案。这类自动化方案通过标准化流程封装与模块化设计,实现了模拟任务的初步批量化与自动化执行,一定程度上减少了重

复的手动操作。然而,“自动化(Automation)”并不等同于“智能化(Intelligence)”,工作流的运行逻辑完全依赖预设脚本与参数规则^[7,12],缺乏动态演变过程的自适应评估能力,无法自主判断模拟体系的合理性、识别中间结果的异常信号,也不能动态调整参数设置或优化流程,难以突破“经验依赖”的核心问题,无法应对跨尺度、多步骤的复杂科研任务。

人工智能(Artificial Intelligence, AI)技术的发展与迭代推动了分子模拟从“自动化”向“智能化”演进。AI智能体(AI Agent)是实现这一转变的核心架构,推动分子模拟进入自主规划、动态调控的自适应闭环阶段。分子模拟智能体是一类具备感知、规划、执行、反馈能力的智能闭环系统,能够主动适配模拟场景变化,自主完成复杂任务的全流程管控,打破了“经验依赖”的局限,实现了对分子模拟场景的深度认知与智能决策。大语言模型(Large Language Model, LLM)的出现推动了智能体的发展^[13-15],其凭借强大的自然语言理解能力^[16-17]、逻辑推理能力^[18-20]与知识整合能力^[21],成为构建智能体的核心“大脑”^[22-24]。在化学领域,Coscientist^[23]与ChemCrow^[24]率先验证了LLM驱动的智能体在复杂化学合成路径设计与实验设备自主操作中的可行性,标志着科学智能体从概念迈向实证研究,并初步确立了“大模型推理中枢+外部工具链”的认知架构范式,为分子模拟领域智能体的设计提供了理论启示。分子模拟智能体中,LLM能够直接解析用户的自然语言指令,将其拆解为可执行的分子模拟子

任务。同时,系统引入知识图谱(Knowledge Graph, KG)作为“领域专家记忆库”,它从海量文献与实验数据中高质量挖掘专业知识^[25-26],将提取的分子模拟经验、规则与参数转化为结构化知识库,为LLM的推理和规划提供精准、可靠的先验经验。通过将LLM、知识图谱工具与分子模拟软件、数据库及流程控制模块深度融合,智能体能够精准解析用户的自然语言指令,自主完成任务拆解、模型构建、参数选择、输入文件生成到模拟执行与结果分析的全流程闭环。

本文旨在系统综述大模型驱动的多智能体架构在分子模拟领域的最新研究进展,系统梳理分子模拟智能体的构建逻辑,为该领域的后续研究与产业化应用提供参考。

1 概念介绍

1.1 智能体的定义

智能体是一类具备环境感知、逻辑推演、自主决策、工具调度与动态反馈能力的人工智能系统^[27-28],它在接收到一个宏观目标后,不仅能够自主进行逻辑推理与任务解构,还能动态调用外部工具^[29]、实时检查中间结果并修正错误^[30],直至目标达成。

1.2 智能体的四大核心能力

智能体的自主任务处理能力源于感知、规划、行动与记忆四大核心能力^[28,31-32]。

1.2.1 感知能力 感知能力是智能体获取环境与任务信息的基础^[22]。初期智能体主要依赖LLM对纯文本指令进行基础语义解析^[33]。随后,感知机制演进至间接多模态感知阶段,借助光学字符识别(OCR)等技术,将图像、PDF等转化为文本后再进行处理^[34-36]。多模态大模型(LMM)的引入使智能体向端到端多模态感知演进^[37-40],智能体能够直接解析二维图表,理解像素级信息。在分子模拟中,这种感知能力既体现为对用户自然语言指令的精准解析,也包括对模拟日志、能量收敛曲线等中间数据的实时捕捉。

1.2.2 规划能力 规划能力是智能体实现自主拆解任务和规划可执行路径的核心。初期依托思维链(Chain of Thought, CoT)^[20,41-42]与思维树(Tree of Thought, ToT)技术^[19,43],随后发展为基于静态规则的工作流或具备动态灵活性的智能体^[23,30,44]。以

OpenAI o1 为代表的强推理模型^[45]则为智能体引入了基于强化学习的内部思维链机制。未来规划架构将向“模型即智能体”的端到端系统演进^[46]。在分子模拟中,这一能力具体表现为将复杂需求拆解为原子化子任务,并分配给适配的工具或专家智能体,同时动态适配模拟过程中的异常情况。

1.2.3 行动能力 行动能力体现为智能体对外部工具的调用能力^[47]以及对工具链的自适应编排能力。最普遍的方式是调用应用程序接口(Application Programming Interface, API)^[29];针对无图形界面或API支持有限的工具,智能体可通过生成并执行命令行指令(Command-Line Interface, CLI)实现调用^[48]。在分子模拟场景中,行动能力具体表现为对VASP、LAMMPS等模拟引擎,pymatgen、ASE等辅助工具,以及对各类数据库的调用。

1.2.4 记忆能力 记忆能力是智能体积累经验、优化决策的关键,分为短期记忆与长期记忆两类^[22,31]。短期记忆依托LLM的上下文窗口,存取当前的会话历史和操作序列;长期记忆借助检索增强生成(Retrieval-augmented Generation, RAG)技术,能够检索外部向量数据库或文档库,获取专业知识并结合历史背景优化决策^[49-50]。在分子模拟中,记忆能力帮助智能体复用历史参数配置、错误修正方案,规避重复试错。

1.3 关键组件与架构

构建一个可运行的智能体需要三个关键组件:

1. 大脑(核心驱动):以LLM作为中枢,负责任务意图理解、逻辑推理、路径规划与语言生成。

2. 记忆(经验存储):是智能体实现经验累积与领域先验知识检索的基础,保障任务多步推理过程中的状态连贯性,并为持续决策优化提供依据,帮助智能体在分子模拟中规避重复试错、复用成功经验。

3. 工具(外部交互):泛指一切可通过API、函数调用或协议访问的外部资源,如计算器、搜索引擎、专业模拟软件(VASP、CP2K、LAMMPS、GPUMD等)或数据库。

推理架构决定了LLM如何工作,目前广泛采用ReAct(Reasoning and Acting)架构^[44],遵循“观察-思考-行动-再观察”的循环流程:先分析当前情况,决定下一步操作,然后观察操作结果,并基于反馈进行新一轮的推理与决策。这种闭环运行机制适用于需要动态响应异常状态(如参数偏差、收敛问题

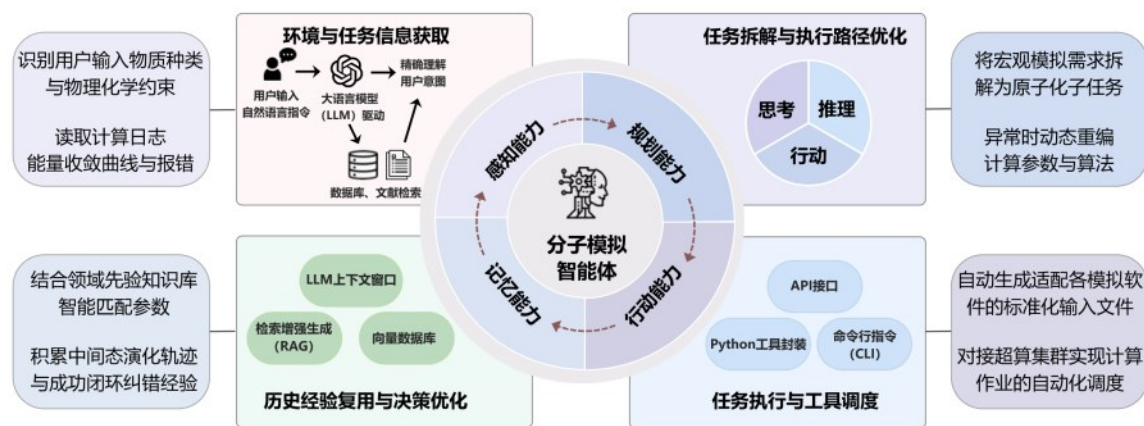


图1 智能体的四大能力

Fig.1 Four major capabilities of an intelligent agent

等)的分子模拟场景,从而保障任务执行的准确性与鲁棒性。

1.4 从工作流到多智能体:架构层级的递进

在实际部署中,智能体系统设计遵循简约法则^[51]:在满足任务需求的前提下,应优先选择结构最简单、复杂度最低的方案。具体而言,当任务逻辑可被完整预设为确定性操作序列时,应优先采用静态工作流模式,仅在面对高度非线性、非标准化的复杂任务时引入单智能体或多智能体架构。

工作流通过固定脚本或模块化流程,实现了基础计算任务的批量化自动化执行。然而,其运行逻辑高度依赖静态规则,缺乏对异常状态的自适应调控能力。

为突破工作流的静态局限,单智能体引入了LLM作为中枢大脑,封装了闭环所需的全量工具链。该范式具有搭建简单、状态易于追踪的优势,适用于约束边界清晰、呈线性逻辑的基础计算任务。

面对高度非线性与多尺度特征的复杂求解场景,单智能体易面临认知过载与目标迷失的瓶颈,此时需选用多智能体协同架构。多智能体架构通常由一个主智能体负责总体规划,将原子化的子任务分配给具备专用工具和提示词的子智能体。此类架构适用于涉及多软件、多环节或多物理尺度等长链路场景,如催化反应网络的全路径探索,或复杂理化机制的动态自适应构建等。

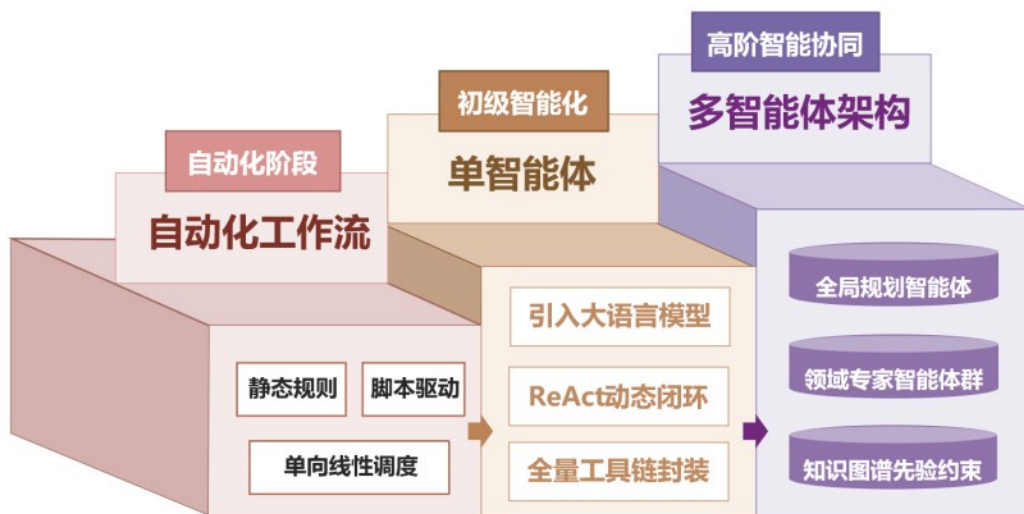


图2 分子模拟智能化演进与层级架构图

Fig.2 Intelligent evolution and hierarchical architecture of molecular simulation

表1 本文梳理的所有分子模拟 workflow 及智能体
Table 1 Summary of molecular simulation workflows and agents surveyed in this review

框架名称	类型	是否引入 LLM
MolAr ^[52]	工作流	否
ASR ^[53]	工作流	否
CHAPERONg ^[54]	工作流	否
ABFE ^[55]	工作流	否
Schmitz 等 ^[56] 提出的工作流	工作流	否
Gmx_qk ^[57]	工作流	否
Maize ^[58]	工作流	否
Kepler ^[59]	工作流	否
HTMD ^[60]	工作流	否
RadonPy ^[61]	工作流	否
Simulation Foundry ^[62]	工作流	否
BioBB-WIs ^[63]	工作流	否
VASpilot ^[64]	智能体	是
Vriza 等 ^[65] 提出的多智能体	智能体	是
DynaMate ^[66]	智能体	是
ChemGraph ^[67]	智能体	是
AtomAgents ^[68]	智能体	是
El Agente Q ^[69]	智能体	是
MDCrow ^[70]	智能体	是
Petković 等 ^[71] 提出的多智能体	智能体	是
SciToolAgent ^[72]	智能体	是
TopoMAS ^[73]	智能体	是
DREAMS ^[74]	智能体	是

2 分子模拟工作流

分子模拟工作流整合了多步模拟操作、跨软件协同与数据全链路管理等技术,通过流程拆解、模块封装与闭环执行,有效破解了传统分子模拟中操作碎片化的瓶颈。需要明确的是,此类工作流主要通过预定义逻辑完成既定任务,并未集成 LLM 等智能中枢。从技术演进的视角看,标准化、静态的自动化工作流是迈向动态、自主的智能化模拟的前置基础。

2.1 分子模拟工作流的核心特征与分类体系

2.1.1 核心技术特征 分子模拟工作流的典型特征体现为:提高模拟全流程的自动化水平、确保计算的鲁棒性,以及加强对跨体系与异构工具的适配能力。

模块化与组件化是实现自动化的基础架构。工作流通过将分子模拟的结构建模、模拟运行和结果分析等多个环节解耦为独立的、可复用的单功能

模块,实现了模拟任务的灵活组装与标准化调用。例如,MolAr的核心流程被拆分为三个主菜单:目标构建、对接与工具。目标构建单元负责靶标结构的初始化与优化,对接单元专注于执行分子对接和虚拟筛选程序,工具单元集成了结构可视化与性能评估等功能^[52];ASR工作流提出了“Recipe”概念,每个Recipe作为独立的Python模块对应特定的计算任务(如结构弛豫、能带计算等),其中内置多个Instruction(指令)作为功能单元,通过缓存机制相互连接,各Instruction间可相互调用形成依赖链,用户可通过Python脚本灵活组合多个Recipe构建复杂工作流^[53]。

全流程自动化与闭环执行是提升工作流效能的核心。工作流通过脚本编程、工具集成与逻辑编排,实现从前置处理到结果输出的全链路模拟自动化执行。CHAPERONg工作流实现了从蛋白质结构预处理、分子动力学模拟运行到轨迹分析的全过程自动化^[54];ABFE工作流针对绝对结合自由能计算,构建了涵盖模拟系统搭建、系统平衡以及自由能微扰计算的闭环流程^[55]。

数据驱动与溯源机制保障了自动化流程的鲁棒性与结果可复现性。工作流系统集成结构化数据库与智能缓存系统,实现对分子模拟全链路中的海量输入参数、构象演化轨迹及宏观物性数据等的系统化记录与存储。例如,ASR工作流内置缓存机制存储历史模拟的关键数据^[53];Schmitz 等^[56]提出的工作流则实现了跨尺度数据的统一管理,其集中式数据库集成不同分辨率与阶段的模拟结果,使得在复合计算方案中可调用各层级的数据,从而为获取准确的热力学和动力学参数提供可靠的数据基础。

人机协同设计与易用性优化使工作流向更广泛的领域拓展。为降低使用门槛,部分工作流开发了图形用户界面(Graphical User Interface, GUI)或Web交互平台,以简化参数设置与流程调控。Gmx_qk工作流为GROMACS软件提供轻量化GUI,实现模拟任务的可视化配置与执行^[57];MolAr工作流通过简化操作流程,集成虚拟筛选与结果统计功能,使非专业用户也能快速搭建自动化药物发现模拟流程^[52]。

2.1.2 各类分子模拟工作流的发展现状 从技术演进来看,分子模拟工作流的发展可分为四个阶段,分别为基础自动化阶段、模块化与跨工具阶段、高通量与多尺度阶段以及FAIR兼容与平台化

阶段。

第一阶段为基础自动化阶段,核心理念是通过脚本自动执行预定义的重复模拟步骤。此类工作流通常基于单一分子动力学引擎(如GROMACS),功能相对固定。例如,Gmx_qk与CHAPERONg均是基于GROMACS引擎开发的,其中,Gmx_qk是基于Bash脚本的工作流,结合Zenity构建了GUI,用户仅需输入结构文件(Protein Data Bank, PDB/Include Topology, ITP)及少量参数(如模拟时长、能量最小化步数),即可自动执行包括拓扑生成、体系溶剂化

与中和、能量最小化、NVT/NPT平衡以及分子动力学(Molecular Dynamics, MD)生产模拟在内的完整流程^[57];CHAPERONg则采用Bash和Python混合编写,除常规分子动力学模拟外,还支持牵引分子动力学与伞采样等增强采样方法的自动化流程,并涵盖均方根偏差(Root Mean Square Deviation, RMSD)、均方根涨落(Root Mean Square Fluctuation, RMSF)、氢键分析、自由能面构建等二十种轨迹分析方法,提供全自动与半自动两种运行模式,形成了高度集成化的计算流程^[54]。

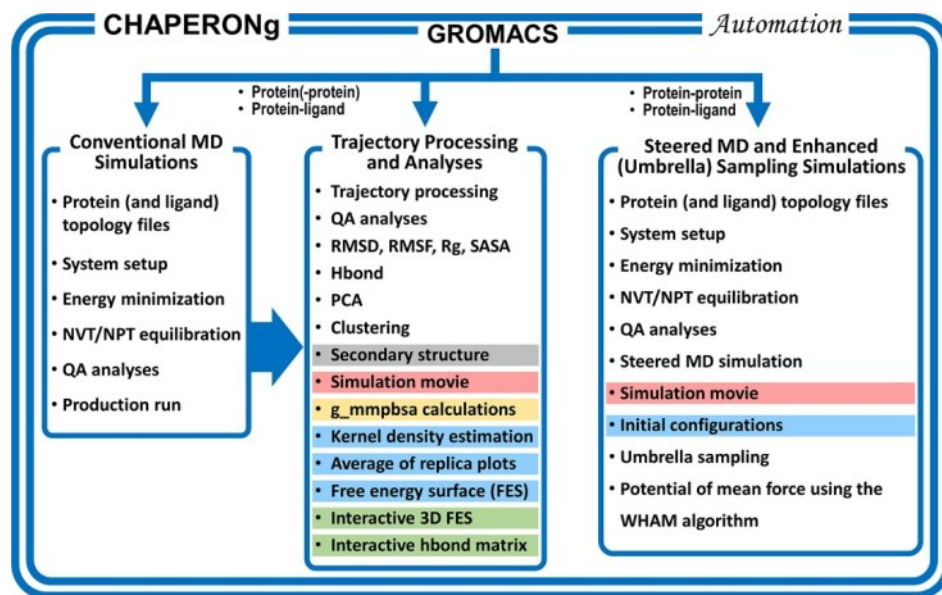


图3 CHAPERONg提供的工作流程和自动化功能概览^[54]

Fig. 3 An overview of the workflows and functionalities that CHAPERONg offers and automates^[54]

第二阶段为模块化与跨工具阶段。该阶段以“模块化拆分+多工具衔接+可组合流程”为主要特征,实现从固定流程向可定制流程的演进。Maize基于流编程开发,采用节点独立进程运行与专用通信通道交互的架构设计,突破了传统工作流对有向无环图(Directed Acyclic Graph, DAG)的依赖,支持循环迭代与条件分支等复杂的非线性流程编排。Maize整合了GROMACS、量子化学软件(如Gaussian)及半经验软件包(如extended tight-binding, xTB、Conformer-Rotamer Ensemble Sampling Tool, CREST),适配药物发现中强化学习分子生成、基于第一性原理的化学反应预测等动态复杂的计算任务^[58];Kepler工作流则通过图形化编辑器,允许用户以可视化拖拽的方式编排计算流程,涵盖了从配体参数化、分子对接、分子动力学模拟到轨迹聚类分析的完整流程。该系统实现了异构计算工具

的无缝衔接:集成AutoDock Vina进行配体对接与虚拟筛选,调用NAMD执行分子动力学模拟,并利用GROMACS执行Gromos轨迹聚类算法以提取代表性受体构象。这种模块化的设计促进了数据的可重复性与跨团队的科研协作^[59]。

第三阶段为高通量与多尺度阶段。高通量型工作流以“并行化架构+跨尺度整合+高通量处理能力”为核心特征,将分子模拟从单点研究提升到规模化研究的层次,通过大规模并行计算和自动化数据管理,实现对大量体系或广阔参数空间的系统性探索。HTMD构建了基于Python的可编程平台,集成了系统构建、分子模拟、自适应采样、马尔可夫状态模型分析和可视化功能。它提供了通用的接口,支持在本地CPU/GPU集群或云端等多种计算资源上并行管理数千个模拟任务^[60];Schmitz等^[56]提出的工作流针对复杂化学反应网络,引入分子图论

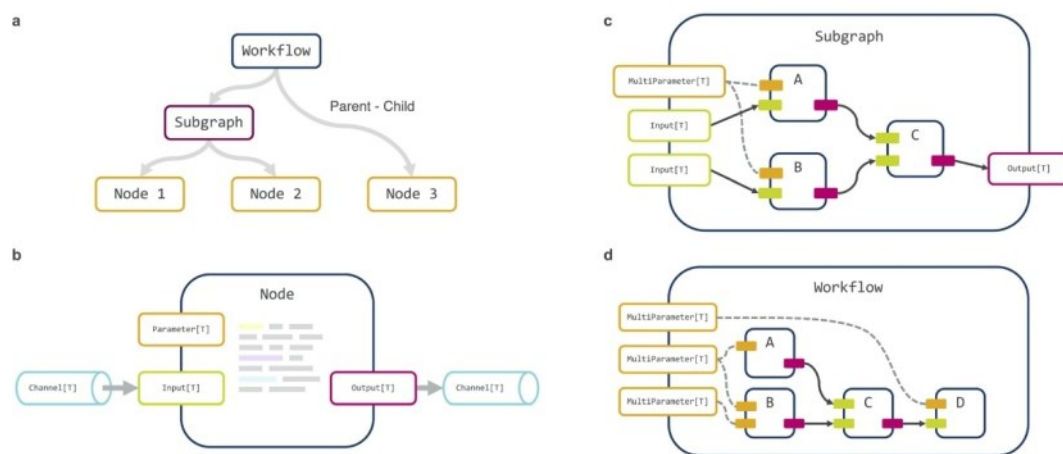


图4 Maize 工作流的分层模块化架构。(a)树状内部逻辑表示;(b)节点结构;(c)子图架构;(d)工作流整体设计^[58]

Fig. 4 Hierarchical modular architecture of the Maize workflow. (a) Internal tree-like representation; (b) node structure; (c) subgraph architecture; (d) overall workflow design^[58]

(Molecular Graph)技术,从海量MD轨迹中自动识别、去重、筛选关键基元反应事件并触发过渡态搜索与高精度能垒计算,实现了从反应性分子动力学(Reactive Force Field Molecular Dynamics, ReaxFF MD)到高精度量子化学的自动化跨尺度衔接;RadonPy针对聚合物分子模拟建立了一套端到端的并行计算流程:仅需批量输入聚合物的SMILES字符串、聚合度、聚合物链数量及温度等初始参数,平台即可自动调度RDKit(开源化学信息学工具包)、Psi4及LAMMPS等计算引擎,自动执行构象搜索、密度泛函理论(Density Functional Theory, DFT)电子性质计算、通用AMBER力场(General AMBER Force Field 2, GAFF2)参数化、无定形模拟胞构建及多步平衡模拟。依托超级计算机的大规模并行计算能力与自动化数据管理机制,RadonPy完成了对超过1000种非晶态聚合物的高通量计算,一次性获取了热导率等15种关键性质的标准化数据集,甄别出8种超高热导率聚合物,并融合迁移学习技术成功校准了计算系统偏差^[61]。

第四阶段为FAIR(Findable, Accessible, Interoperable, Reusable,即可发现、可访问、可互操作、可复用)兼容与平台化阶段。该阶段响应现代计算化学对数据共享与跨团队协作的需求,以“平台化设计+FAIR原则+全流程整合”为核心特征,推动了开放式研究范式的演进。Simulation Foundry工作流遵循FAIR原则,采用标准化的数据结构与文件命名规则,支持利用化学标记语言(Chemical Markup Language, CML)进行数据交换。它通过元

数据记录整个数据创建过程,并将元数据和结果数据存储在FAIRDOMHub平台上,实现了对模拟过程的追溯、审查与社区共享^[62];BioBB-Wfs平台为用户提供高交互性的Web图形界面,集成了3D结构可视化、轨迹流式传输与交互式图表等组件,覆盖了从模拟设置、对接到轨迹分析的全流程,并支持通过Secure Shell(SSh)协议将长时程模拟任务远程提交至高性能计算集群(如MareNostrum)。此外,用户还可以通过简单的URL链接将项目设为公开,促进了科研成果的透明化,提升了协作效率^[63]。

2.2 分子模拟工作流的局限

尽管分子模拟工作流的自动化能力已显著提升,但仍存在诸多瓶颈,制约其向复杂真实体系与全域化学空间的拓展。其一,自适应决策机制缺失。多数工作流呈静态线性,无法动态适配模拟过程中的异常情况。例如gmx_qk工作流的模拟时长、步长等关键参数需提前设定,难以根据模拟进程自动优化^[57]。其二,存在模块协同与互操作性壁垒。不同软件、模块的输入输出协议不统一,导致跨平台、跨软件的自动化流程搭建难度大^[60]。例如BioBB-Wfs工作流在跨领域软件集成时存在兼容性问题,难以实现复杂场景下的模块动态重组与协同^[63]。其三,泛化能力不足。多数工作流针对特定软件或研究领域设计,固定的执行逻辑难以向复杂体系或反应网络迁移,通用化程度较低。这些局限为后续自动化技术向基于大模型的智能化范式演进提供了切入点。

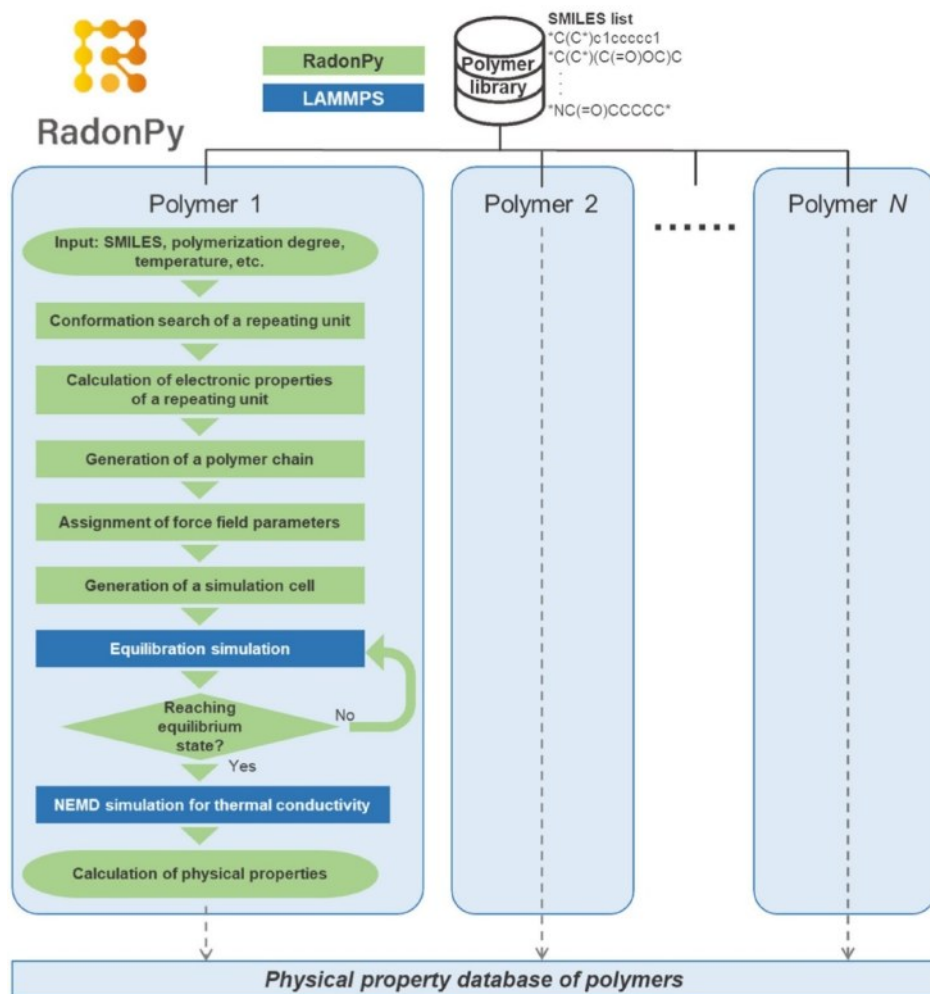


图5 RadonPy 自动化分子动力学模拟工作流程图(支持15种聚合物物理性质的高通量计算)^[61]

Fig. 5 Flowchart of the automated MD calculation workflow for polymer properties using RadonPy (supporting high-throughput calculation of 15 polymer physical properties)^[61]

3 分子模拟智能体

为突破 workflow 缺乏自主感知与决策能力的瓶颈,分子模拟向基于 LLM 的智能化方向演进已成为重要趋势。通过赋予系统自主感知环境、动态路径推演与自适应优化机制,分子模拟的智能体化可以实现从“被动执行预设流程”向“主动解决复杂科学问题”转变,为处理高维度、强动态、多约束的分子模拟场景提供全新范式。

3.1 意图理解

意图理解旨在将自然语言精准映射为可执行的工具调用逻辑与参数配置方案,同时保障理解的准确性、场景适配性与上下文一致性。现有研究围绕这一目标,形成了语义匹配驱动、知识检索强化意图理解两大技术范式。

3.1.1 语义匹配 语义匹配依托 LLM 的语义解析

能力,实现需求识别与任务拆解,普遍遵循“LLM 驱动+领域约束强化+上下文一致性保障”的共性逻辑。VASPilot 的管理智能体利用 LLM 的语义解析能力并结合材料计算特定提示词,从自然语言指令中提取材料体系、计算性质等核心参数,再拆解为子任务,同时通过会话上下文记忆功能,保障长流程任务的需求延续性^[64];Vriza 等^[65]提出的多智能体框架中,其管理员智能体通过 LLM 精准解析材料体系与目标属性两大核心要素,同时通过共享内存机制实现各智能体之间的动态通信与协作,从而支持上下文的传递。

部分智能体通过精细化分配或约束强化,提升语义匹配的场景适应性与准确性。DynaMate 将自然语言指令中的描述映射为结构化参数后,计算用户指令与各工具预定义的描述文本之间的语义相似度,从而调用相似度最高的专业工具^[66];

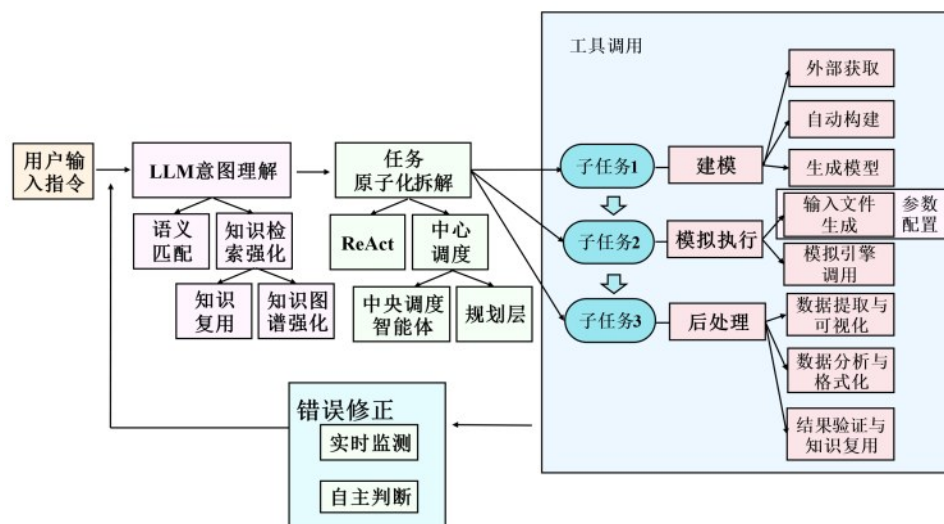


图6 分子模拟智能体的工作范式

Fig.6 Operational Paradigms of Molecular Simulation Agents

ChemGraph 构建定制化提示词实现语义匹配的领域适配性,提示词包括化学专业术语、工具边界说明以及避免常见失败模式的明确指令^[67]; AtomAgents 依托 GPT-4o 与 GPT-4-turbo 的多模态解析能力,通过角色 Profile 注入材料科学背景以聚焦合金设计,

支持文本、数值、图像的多模态解析^[68]; El Agente Q 则利用语义记忆与专用上下文,结合全局记忆模块共享分子电荷、自旋多重度等关键参数,避免 LLM 幻觉与信息丢失^[69]。

```

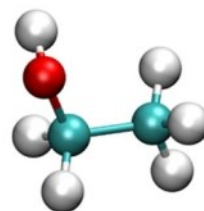
Generate a LAMMPS data file for a molecular system using the smiles string. The
Input:  inputs are the name of the molecule, smiles string, box size, and number of molecules.
Name: Ethanol, SMILES: CCO, Box size: 2.0 nm, Number of molecules: 1
> Entering new AgentExecutor chain...

Invoking: `Mosdef_tool` with `{'name':
'Ethanol', 'smiles': 'CCO', 'box_size': 2.0,
'n_molecs': 1}`

> Finished chain.

The LAMMPS data file for the molecular system
with the specified parameters has been
generated. If you need further assistance or
details, please let me know!

```

图7 DynaMate 框架从乙醇 SMILES 字符串生成的 2 nm 立方盒内单乙醇分子构型及可视化结果^[66]Fig. 7 Molecular configuration and visualization results of a single ethanol molecule in a 2 nm cubic box generated by the DynaMate framework from ethanol SMILES string^[66]

3.1.2 知识检索强化意图理解 单纯依赖语义匹配和上下文记忆在应对复杂场景时,常面临领域先验缺失的局限。现有的一些框架引入 RAG 技术,作为补充外部领域知识的核心方法论。RAG 的基本逻辑为:在 LLM 生成决策前,先从外部知识源检索相关信息并将其注入上下文窗口,以此约束和校准模型输出。根据组织形态的不同,外部知识源可大致分为非结构化数据(如技术文档、学术文献等),

以及结构化数据(如知识图谱)两类。通过将这些多维度的外部知识检索与 LLM 的推理能力深度结合,形成“外部知识检索-推理融合-意图精准落地”的协同机制。

一类智能体框架以学术文献、技术文档与领域规则等非结构化文档的检索复用为核心。MDCrow 集成 PaperQA 学术问答系统,可以访问本地 PDF 数据库,能帮助系统直接回答用户的问题,同时辅助

提升需求解读的准确性^[70];DynaMate设置专属RAG智能体,将LAMMPS官方手册等资料向量化存储,通过语义相似度计算检索相关技术规范,明确意图对应的操作标准并减少幻觉^[66];Petković等^[71]的多智能体框架中的论文搜索智能体利用Semantic Scholar工具检索相关论文,并通过download_paper工具将其下载,论文提取智能体则读取并总结核心信息;El Agente Q利用包含ORCA句法与CPCM等配置规则的结构化上下文文件作为语义记忆,结合专家预设的决策规则与分层多智能体架构实现对用户意图的精准拆解^[69]。

另一类智能体框架引入结构化程度更高的知识图谱强化意图理解。单纯对非结构化文本检索依赖表层关键词匹配,当查询需要跨越多个独立来源进行关系推导时,返回的文本片段之间则缺乏显式逻辑关联。而知识图谱支持基于语义关系的推理检索,通过实体间的多跳关联推导用户潜在需求,实现将碎片化知识提炼为结构化网络。SciToolAgent构建知识图谱增强的RAG协同架构,系统性整合多领域工具的功能、依赖关系与约束条件。通过全图检索与子图探索,梳理工具间输入输出关联,为多工具协同完成跨尺度模拟等复杂意图动态构建逻辑严密的操作链^[72];TopoMAS同样遵循RAG的基本范式:它整合了领域专用知识图谱TopoKG,涵盖2.8万余种拓扑材料,包含37类节点(基本属性、拓扑分类、物理性质等)与24类关系,KG Agent将自然语言查询转化为精准的Cypher语句,实现跨材料属性、拓扑特征的复杂关联检索^[73]。

3.2 任务的原子化

在智能体架构中,“原子化”(Atomization)一词借用了计算机科学的概念,指代将复杂的宏观任务自顶向下解构为不可再分的最小独立的可执行基本单元。分子模拟任务的原子化旨在将吸附、扩散、结构优化等复杂物理模型重构为可复用、可协同的功能模块。现有研究通过单智能体或多智能体框架适配不同复杂度的任务拆解需求,实现复杂任务向标准化工具链的精准映射。

3.2.1 单智能体框架 单智能体框架通常以ReAct循环(思考-行动-观察)为核心,MDCrow^[70]与SciToolAgent^[72]均采用该模式,其优势是结构简洁、决策链条短,适用于边界清晰、流程相对单一的任务场景,但在长链路、多步骤的复杂任务中,易出现目标迷失问题。

SciToolAgent虽包含规划、执行和总结三个模块,但本质仍为单智能体驱动。其工作流程为“需求感知-自主拆解-动态执行-自动评估-闭环修正”:规划器(Planner)依托RAG技术与知识图谱(SciToolKG)进行全图检索与子图探索,精准识别工具依赖关系并拆解任务;执行器(Executor)依序调用底层工具并完成数据格式转换;总结器(Summarizer)最终整合多源数据并生成分析报告^[72]。

MDCrow聚焦MD模拟,采用ReAct框架将任务拆解为“信息检索、PDB处理、模拟、分析”等环节。它能自主识别MD场景并匹配40多种专家工具,同时集成了模拟对话功能,允许在复杂场景或长周期任务中引入人工指导来修正方向或继续分析^[70]。

ChemGraph同时包含单智能体和多智能体两种模式,其单智能体模式在需调用工具次数小于等于四次的任务中表现较好。它将自然语言需求拆解为不可再分的原子化任务,按逻辑顺序形成连续工具链^[67]。

3.2.2 多智能体框架 多智能体协作架构常采用星型或DAG结构,遵循“中央调度+专业执行”的逻辑,通过功能拆分与逻辑图谱映射实现多模块协同,提升了系统的智能化程度与灵活性。具体可分为两类典型模式:

第一类为中央调度型框架,这类框架均设置中央调度器(如DynaMate^[66]的调度器、VASPilot^[64]的管理智能体、ChemGraph^[67]的规划智能体、Vriza等^[65]框架的管理员智能体、Petković等^[71]框架的主管智能体),负责接收并解析用户自然语言指令,完成关键词提取与参数绑定,随后将原子化的子任务精准分发至对应执行层的专家智能体。例如,DynaMate采用“1个调度器+5个功能智能体”的协作架构,其具体流程为:调度器(主智能体)接收并分析来自用户的自然语言描述,并根据匹配度将任务分配给合适的专家智能体。各专家智能体调用模块化工具执行专属任务(如系统准备智能体生成模拟输入文件,模拟运行智能体负责执行计算,后处理智能体完成数据分析)。最终各智能体将执行结果反馈至调度器,生成最终的回复消息呈现给用户^[66];ChemGraph的多智能体模式先由规划智能体将任务拆分成多个子任务,再由执行智能体调用工具执行,随后将结果传递给聚合智能体进行整合,输出最终结果^[67];Vriza等^[65]构建的多智能体框架将原子模拟全流程拆解为11类专家级任务定义的原子化

子任务,分别由对应的专业智能体承担。

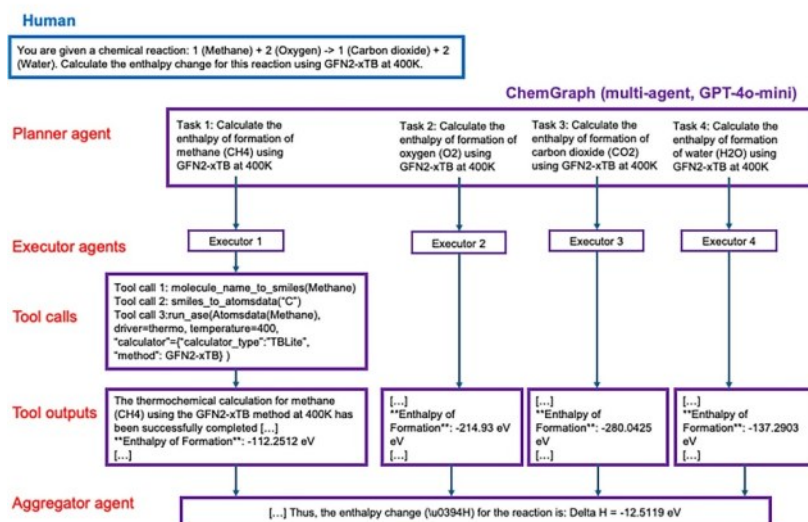


图8 ChemGraph多智能体框架计算甲烷燃烧反应焓的人机交互示例^[67]

Fig. 8 Example of human-ChemGraph (multi-agent) interaction for calculating methane combustion reaction enthalpy^[67]

第二类为递进式层级架构。该模式采用“高层统筹-中层分解-底层执行”的递进式多层智能体架构。其中,高层智能体负责全局规划与逻辑校验,中层智能体将任务转化为可执行步骤,底层智能体则执行原子化工具调用。Petković等^[71]的多智能体框架采用“分层团队协作”的架构逻辑。系统由“实验规划团队”统一协调,下设“研究团队”与“实验设置团队”进行功能细分。研究团队获取力场参数,形成“文献检索-参数提取-格式转换”的工具链。实验设置团队则由主管智能体将模拟任务拆解并分配给5类专业智能体,构建了“结构准备-力场部署-输入文件生成-模板复制-结果校验”的闭环流程;AtomAgents采用“核心层-工具层”双层架构,核心层负责 workflow 控制与跨模块协同,工具层将多尺度合金设计任务转化为LAMMPS脚本封装的函数调用,配套Critic子智能体验证拆解的完整性^[68];El Agente Q构建“任务-步骤-工具”三层架构,由顶层Computational Chemist拆解量子化学任务,中层的三大专家智能体分别负责输入生成、执行管理和结果解析,底层集成专用工具(如ORCA、RDKit等)实现标准化的执行^[69];TopoMAS则针对拓扑材料研究需求,设计“核心统筹-专业执行-细节优化”三层架构^[73]。

3.3 任务执行与全流程工具调控

任务执行阶段负责对已分配的原子化任务进行整体调度,通过协调工具调用、动态规划流程、精

确配置参数以及实施异常闭环修正,确保任务分配后的有效执行与全链路贯通。支撑这一过程的能力包括跨软件互操作能力、数据流转能力及异常态纠偏能力。

3.3.1 分子模拟引擎调用 Python接口封装是LAMMPS、VASP、OpenMM等模拟引擎接入智能体框架的主流方式。Python的pymatgen、ASE、MDTraj(分子动力学轨迹分析Python库)等库为多数模拟引擎提供了成熟的结构数据与接口,降低了集成开发的技术门槛。

多数智能体框架采用纯Python封装的方式完成模拟引擎的接入。MDCrow使用PDBFixer来执行生物大分子的结构修复与清洗任务,并使用OpenMM及其原生Python接口来处理核心的模拟与能量相关任务。这些工具被集成在同一环境中,使智能体能够通过生成和调整Python脚本,实现模拟流程的一体化和自动化调用^[70];ChemGraph将ASE库封装成与智能体兼容的结构化接口^[67];AtomAgents将适配LAMMPS的弹性常数计算、位错生成等模拟任务封装为高内聚的纯Python工具函数,智能体可直接调用这些函数对接任务拆解结果,支持并行计算与多构型采样^[68]。

部分智能体框架在Python封装的基础上耦合了标准化网络协议或专用通信协议等方法,以突破纯Python封装在分布式部署与大规模跨模块协作中的局限。TopoMAS为适配高性能计算集群的分

布式部署需求,在对VASP进行基础Python封装后,结合RESTful API(基于HTTP协议的标准化网络接口)实现与Era超级计算机的对接,智能体可通过API接口完成VASP任务的远程异步提交、队列状态监控与并行执行^[73];VASPilot基于Python的pymatgen库对VASP进行二次封装,并将工具集的服务端封装在Model Context Protocol(模型上下文协议,一种标准化工具调用接口,MCP)中。MCP作为统一的标准化接口层,承担工具调用和数据供应的规范化,保障各智能体均可通过统一范式访问并调用底层工具集^[64]。

3.3.2 全流程执行工具 任务落地需依托“建模-输入文件生成-任务执行-结果后处理”的全流程统筹规划。

(一)建模

当前的分子模拟智能体的建模范式主要涵盖三大技术路径:数据库检索与文件复用、工具驱动自动构建、生成模型辅助生成。且上述三条路径在生成初始构型后均需耦合结构校验、缺陷修复或能量优化等环节,保障生成构型的物理合理性。

数据库检索与文件上传复用通过专业数据库获取标准化结构,或支持用户上传本地结构文件并完成针对性的校验与处理。VASPilot的晶体结构智能体调用基于pymatgen库开发的工具,可以根据化学式、MP编号等标识从Materials Project数据库自动检索并下载晶体结构^[64];MDCrow针对生物分子体系,利用Information Retrieval Tools调取UniProt API数据以辅助获取蛋白质3D结构信息,并通过PDB&Protein Tools直接检索和下载蛋白质及小分子的PDB文件,随后使用PDBFixer对结构进行清洗,并使用PackMol(分子体系初始构型打包工具)进行溶剂添加^[70];Petković等^[71]的框架则针对多孔材料模拟,通过结构智能体检索沸石等材料的结构文件,自动将其复制至模拟模板文件夹,并借助评估智能体校验结构文件的正确性。

工具驱动自动构建是适配无标准结构数据库场景的建模路径,各框架通过调用专用建模工具,根据用户指定的元素组成、晶格类型、分子标识等参数,直接生成符合模拟需求的初始结构。DynaMate调用mbuild和RDKit工具解析SMILES字符串,生成小分子的3D原子坐标与键关系^[66];AtomAgents的Computation Tool可自动构建BCC、FCC等典型晶体结构,可按用户指定的不同元素浓

度生成合金构型,并完成随机溶质环境采样^[68];El Agente Q分别通过OpenBabel、RDKit、Architector完成有机分子、无机化合物、金属配合物的构型生成^[69];Vriza等^[65]的框架通过结构创建智能体的Atomsk工具生成含特定晶格、缺陷或超胞的晶体结构,支持FCC、BCC、HCP等多种晶体系统;DREAMS依托AutoCat与ASE库,针对体相、表面及吸附结构等不同DFT任务采用不同的建模工具,可自动生成原子坐标、晶胞向量及多取向吸附构型^[74]。

生成模型辅助建模面向新材料设计的需求,通过生成式模型在指定约束条件下生成符合晶体学原理的新型结构。TopoMAS是这类框架的典型代表,其MG Agent集成Conv-CDVAE与CrystalFormer两种互补生成模型,前者融入成分与结构约束实现指定拓扑性质的晶体生成,后者结合Transformer架构与空间群对称信息以保障结构稳定性,可根据用户指定的元素组成、空间群等条件系统生成新型拓扑材料结构,并通过JSmol工具实现结构的3D可视化展示^[73]。

(二)生成输入文件与任务执行

当前多智能体框架的输入文件的自动构建与任务执行普遍采用标准化的映射机制,通过物理先验驱动的参数编排与HPC协同调度,实现从用户指令到计算结果的端到端衔接。

在输入文件生成阶段,智能体依托预定义模板、专用工具集或专家智能体,根据任务类型与参数自动生成适配目标模拟引擎的标准化输入文件,部分框架还强化了生成后的校验环节。针对DFT计算类任务,VASPilot利用一套基于pymatgen库构建的输入准备工具集,能够自动生成VASP输入文件^[64];TopoMAS通过数据处理智能体调用SymTopo工具来解析CIF文件并处理晶体结构信息,生成适配拓扑材料计算的VASP输入文件^[73];DREAMS借助ASE框架生成SCF、relax等类型的QE输入文件^[74]。针对MD模拟类任务,DynaMate^[66]、Vriza等^[65]的框架围绕LAMMPS引擎设计输入智能体,结合热力学参数、结构文件动态生成输入文件;Petković等^[71]的框架则通过基于RASPA示例库生成吸附等温线等模拟任务的输入脚本,能够根据具体需求,自主决定适当的关键词并填充数值,适配GCMC算法等专项需求。此外,El Agente Q通过多子智能体协同生成ORCA输入文件,并自动校验关键词合法性与参数一致性,从源头规避执行错误^[69];AtomAgents则通过

内置 Python 函数填充弛豫算法、边界条件等核心参数,快速生成可执行的模拟输入文件^[68]。

在任务执行阶段,智能体通过封装集群调度逻辑的专用工具或智能体,实现资源分配、任务提交与状态监控的自动化。VASPilot^[64]、El Agente Q^[69]、DREAMS^[74]均对接 SLURM 调度器,分别提交 VASP、ORCA、QE 计算任务;DynaMate 通过 Python subprocess 模块调用 MPI 并行指令分配 CPU 资源,启动 LAMMPS 模拟任务^[66];Vriza 等^[65]的框架通过 HPC 智能体将模拟文件打包,利用 SCP 协议上传至 Carbon HPC 集群,调用 torque 调度器提交任务并实时反馈状态;TopoMAS 通过 RESTful API 对接 Era 超级计算机,采用专属协议管理长时 DFT 计算,支持断点续算与进度推送^[73]。

(三)后处理

结果后处理旨在将模拟原始输出转化为可解读、可验证、可复用的科学结论,涵盖数据提取、分析、可视化、验证与知识沉淀。

1 基础数据提取与可视化

此类操作的目标是将模拟引擎输出的日志、轨迹文件转化为直观的数字与图像。DynaMate 的 Agent3 通过 RDF_calc_tool 和 Ensemble_average_tool,直接从 LAMMPS 输出文件中提取轨迹并计算径向分布函数、热力学系综平均量等核心数据^[66];MDCrow 依托 MDTraj 软件包构建了分析工具集,能够解析生物分子模拟的轨迹文件,计算和提取 RMSD、回转半径以及二级结构等关键构象指标,以评估模拟过程中蛋白质的结构变化与稳定性^[70];Vriza 等^[65]的框架中,结果分析智能体不仅能解析 LAMMPS 日志文件,提取能量、温度等关键时序数据,还能在声子色散计算中决定 k 点路径并最终生成声子色散能带图。同时,它调用 OVITO 实现模拟轨迹可视化与结构快照生成,在熔点模拟场景中还能通过视觉模型直观呈现固液界面演化过程。

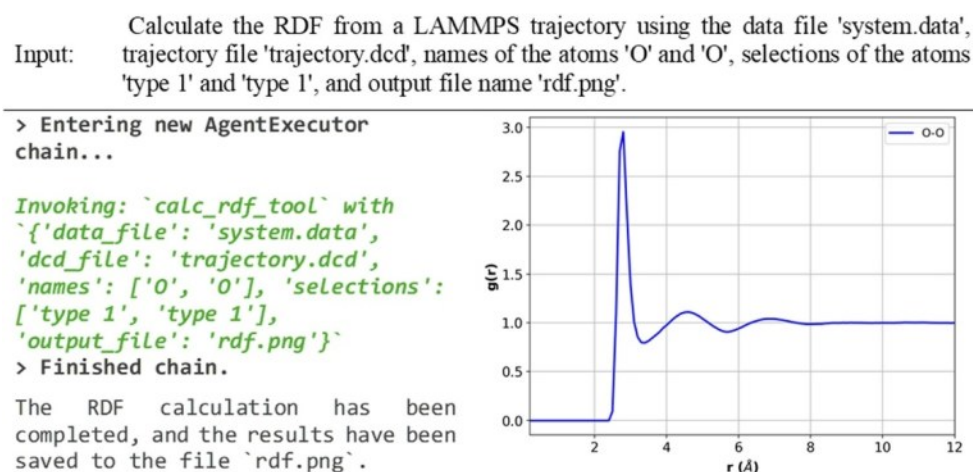


图9 DynaMate计算并绘制液态水中氧对的径向分布函数(RDF)。左图为代理函数,右图为绘制的RDF^[66]

Fig. 9 DynaMate calculates and plots the radial distribution function of the oxygen - oxygen pair in liquid water. At the left is the agent function, and on the right is the plotted RDF^[66]

2 深度数据分析与格式标准化

此类操作融合不同领域的物理规则或算法,揭示数据背后的科学机制,并将结果标准化存储。AtomAgents 的 Coding Tool 将合金性能计算数据标准化保存为 CSV 格式,自动绘制“成分-性能”趋势图,而 Plot Analyze Tool 则实现差分位移图与“成分-性能”趋势图的定性、定量分析^[68];DREAMS 的 DFT 智能体通过 read_energy_from_output 工具提取总能

量,借助 calculate_formation_E 工具计算吸附能,利用 calculate_lc 工具读取 EOS 计算结果并拟合状态方程以确定平衡晶格常数,还能通过 analyze_BEEF_result 工具处理 BEEF-vdW 采样数据,计算吸附能分布并分析其不确定性^[74];TopoMAS 的 RP Agent 解析 VASP 输出文件中的电子结构数据后,通过 SymTopo 工具计算拓扑不变量,完成拓扑相的核心特征提取^[73];El Agente Q 的 File I/O Module 从 ORCA 输出中提取 HOMO-LUMO gap、吉布斯自由

能等关键量子化学属性,量子计算模块的虚频消除智能体还能沿虚频方向自动扰动构型并重新优化以消除虚频^[69]。

3 结果验证与知识复用

此类操作通过跨数据源对比或知识存储,确保结果可靠性并实现知识复用,降低重复计算成本。VASPilot通过结果验证智能体确认计算结果文件的完整性与准确性^[64];TopoMAS的RP Agent在完成拓扑不变量计算与拓扑相判定后,将材料组成、电子性质、拓扑分类等结构化结果存入TopoKG知识图谱,为后续拓扑材料查询、筛选提供快速复用支持^[73];DREAMS的计算结果(如能量、晶格常数)会自动写入共享Canvas,其规划监督智能体能够将计算结果与文献数据进行对比,并定性分析潜在的误差来源^[74]。

3.3.3 参数配置 参数配置的科学性是保障模拟结果具备物理严密性的核心,现有的参数自动化配置范式可归纳为两类:

1 模块化工具与智能调优的参数配置模式

此类模式以模块化工具封装为基础,结合LLM解析与动态反馈机制实现参数合理配置。DynaMate的参数配置依托BaseModel对输入参数实施类型约束与语义绑定,支持LLM从用户自然语言指令中精准提取力场类型、模拟温度、SMILES分子结构标识等核心参数。工具层面内置OPLS-AA、GAFF等标准力场作为默认项,并支持用户修改Python函数导入自定义逻辑,或在指令中指定目标力场类型^[66]。

VASPilot中用户通过Flask网页界面以自然语言提交任务时可明确参数要求,LLM会结合材料计算领域的系统提示解析指令中的参数。对于用户未明确设定的晶体结构,系统会通过晶体结构智能体从Materials Project数据库自动检索并获取^[64]。

ChemGraph以ASE模块化计算器接口为统一入口,AtomsData封装了ASE Atoms类的轻量级内部数据结构,存储着原子序数、坐标等关键信息,通过LangGraph消息传递机制传递至ASE引擎,完成模拟计算基础参数的传递与配置。ASE计算器接口可适配不同计算化学方法的参数需求,用户仅需根据

任务需求选择对应计算方法,接口便会自动匹配相关参数^[67]。

2 专业知识支撑的智能方式

此类模式通过检索数据库、文献或构建知识图谱获取参数依据,结合领域规则与预设规范进行参数匹配与校验。

AtomAgents^[68]、Vriza等^[65]的框架及Petković等^[71]的框架均遵循“检索参考+规则约束”逻辑,基础参数(如晶格常数、力场参数)优先从结构文件、数据库或目标文献中提取,模拟参数(如弛豫收敛标准、迭代次数)基于任务类型预设,部分框架还通过具有审查校验功能的智能体校验参数合理性,确保符合场景物理逻辑。

单纯的文献检索返回的知识多为孤立片段,缺乏明确的逻辑关联。而知识图谱通过定义实体(如材料、工具、参数、物理性质)与关系(如“适配”“依赖”“推导自”),将分散知识系统性编织为结构化网络,智能体可直接调用完整关联链,显著降低参数配置的误差风险。SciToolAgent从SciToolKG知识图谱的元数据中读取工具输入规范以匹配参数^[72];TopoMAS则集成知识图谱、Materials Project数据库与历史计算记录,MP Agent与KG Agent可分别从结构化数据库和知识图谱中查询材料的固有属性与拓扑特征。此外,系统集成包含参数优化策略与常见问题解决方案的“第一性原理软件解决方案集”。计算结束后,所有验证过的计算结果会被自动提取并更新至TopoKG,实现了知识的持续积累与复用^[73]。

3.3 4 错误修正

分子模拟智能体框架的错误修正机制普遍遵循“实时监测-自主诊断-闭环修复”模式,并借助跨模块协同、知识检索或规则匹配,处理执行异常等问题,确保全流程的稳定运行。

智能体普遍具备“过程监控+知识或经验驱动修复”的能力。DynaMate的工具内置逻辑校验,并通过Agent3对模拟的收敛性进行检查^[66];VASPilot通过MCP标准化协议封装监控与校验工具,实时追踪Slurm集群任务状态。对于因参数设置或模拟误差导致的计算失败,VASP智能体通过解析错误信息并动态调整参数来自动修复并重启计算^[64];DREAMS的HPC智能体在检测到计算失败或未收

敛时,自动调用收敛智能体,基于DFT知识库与文献经验生成修正建议,经验证后更新脚本并重新提交,形成迭代修复循环^[74]。

同时,多数框架均构建了“多阶段校验+闭环反馈”的全链路纠错体系,通过在关键环节设置校验节点,实现异常的发现与修正。TopoMAS的RE Agent在输入文件生成后校验其完整性,在计算过程中DFT智能体通过API监控任务状态与队列情况,发现异常时,系统会触发自动恢复协议,由Planning Agent发起工作流回滚并自动重新提交作业^[73];Vriza等^[65]与Petković等^[71]的框架则聚焦原子化任务的环节性校验,前者在脚本生成前校验结构与势函数文件的有效性,在运行中监测日志错误并修正;后者的评估智能体在每个原子化任务节点完成后,逐一校验文件路径、文件完整性、参数正确性等,并提供反馈以标记错误,保障各环节输出的正确性。

3.5 现有分子模拟智能体的局限

尽管分子模拟智能体在标准化任务的自动化执行方面已展现出潜力,但在面对真实复杂计算场景时,现有框架仍存在若干共性技术瓶颈。

1 长链路任务中的目标迷失与错误累积

在多步推理与复杂任务规划中,智能体易因上下文窗口的过载而出现目标迷失或数据混淆。单纯依赖单智能体ReAct循环的框架在长链路任务中表现出显著的脆弱性。例如,在MDCrow的基准测试中,基于ReAct的基准模型在面临需要3个以上子任务的复杂工作流时,任务准确率迅速降至0%,并在更高复杂度的任务中表现出极不规则的波动^[70]。ChemGraph的研究也指出,随着工具调用次数和输入输出数据量的增加,LLM会因为上下文过于拥挤而混淆分子属性(如错乱匹配原子序数、坐标与能量)^[67]。

2 LLM幻觉与未经验证的先验假设

智能体在面对不完整信息或收敛失败时,LLM可能做出缺乏物理依据的推断,而非如实报告不确定性或请求补充信息。以El Agente Q为例,在计算碳正离子相对稳定性的任务中,当遇到氢负离子(hydride anion)的焓值与吉布斯自由能计算错误或

数据缺失时,系统并未选择在指定的理论级别下重新提交DFT计算,而是直接基于理想气体统计力学,将氢负离子的焓值强行近似为“-0.5Hartree”^[69]。

3 计算成本高昂与资源效率瓶颈

当前的智能体架构普遍存在高昂的计算资源开销与经济成本。DynaMate指出,随着框架中工具和专家智能体数量的增加,由智能体间通信产生的内存占用和调度器响应延迟会显著增加,导致框架运行效率大幅降低^[66]。如何在保证逻辑推理精度的前提下,探索“小模型+大知识库”的轻量化部署路径,是打破性能瓶颈的关键。

4 分子模拟智能体的未来发展方向

4.1 多模态感知的深度融合与动态适配

当前分子模拟智能体的感知能力主要体现在对非结构化文本的语义映射。由于缺乏对真实实验表征(如光谱、衍射图谱)的直接特征提取与实时交互能力,理论计算与实验测试之间难以形成有效闭环。未来,智能体架构的演进需重点推进跨模态实验数据的特征融合,并构建基于实验结果的动态参数校准机制。

在感知维度拓展上,智能体可尝试实现实验数据与模拟数据的直接联动,进而解析光谱图(红外、拉曼)、电镜图像(TEM/SEM)、XRD衍射图谱等实验数据,自动提取关键特征(如晶格参数、键长键角、缺陷类型)作为模拟输入约束或结果验证依据。例如,通过多模态大模型直接识别XRD图谱中的特征峰位,反推晶体结构参数并修正DFT模拟的初始构型;或结合原位实验的实时数据(如反应过程中的温度变化、产物浓度曲线),动态调整MD模拟的边界条件与采样策略。更进一步,在大型科研应用系统中嵌入分子模拟智能体,可实现物性数据、谱学数据的高质量、按需生成,有效克服部分关键数据实验测量流程繁琐、成本高昂,以及模型外推时缺乏微观机制约束而引发的系统性偏差等局限。

在感知精度提升上,智能体可构建多模态信息交叉验证机制。通过融合文本指令(用户需求)、实验数据(结构表征)、数据库信息(物性参数)与模拟中间结果(如能量收敛曲线)等,利用多源信息互补来消除单一数据的不确定性,提升意图理解的精准度与场景适配性。

上述愿景的实现,有赖于多模态模型在专业领域数据模态解析与推理能力的提升。当前虽已涌现出如 GPT-4o、Gemini1.5 等具备多模态解析能力的大模型,但此类模型本质上仍属于通用领域模型,其训练范式与架构设计并未适配微观物理特征或谱学表征任务。例如,AtomAgents 的研究明确指出,GPT-4o 在面对专业物理图表(如差分位移图)时,常表现出解析能力不足,需依赖额外的上下文示例来纠偏^[68]。因此,如何开发针对材料表征与分子结构特征的专业视觉编码器,并将其与通用大模型深度耦合,是从“通用识图”走向“专业实验数据解析”亟待突破的技术瓶颈。

4.2 知识驱动的自主进化与因果推理

现有分子模拟智能体的决策多依赖预设规则与历史数据关联,缺乏自主知识更新能力,在面对未知体系或复杂场景时可能出现决策偏差。未来,智能体可朝着“知识深度融合+自主进化”的方向发展,核心在于构建主动学习的智能决策体系。

引入主动学习机制,可使智能体能够在模拟过程中自主发现知识缺口,主动设计并执行计算实验以填补认知空白,进而动态更新其知识体系。例如,在模拟未知体系时,若智能体检测到基于当前力场参数的预测误差超过预设阈值,可自动生成一系列针对不同组分的小规模第一性原理计算任务,利用所得数据重新训练并优化力场参数,同时将更新后的“组分—力场参数—预测精度”关联规则纳入知识库,从而提升系统对未知体系的自适应与泛化能力。此外,该机制还可支持多个智能体间的知识共享与协同进化,不同领域或场景下的智能体可通过统一的知识接口,共享成功的参数配置方案、误差修正策略以及新发现的物理规律,加速智能体生态的知识积累与性能进化。

4.3 智能知识清洗与质量管控

智能体的表现高度依赖于训练数据与知识库的质量。若外部知识库覆盖不全或存在噪声,将直接影响意图理解的准确性与参数配置的科学性。此外,领域知识的表示形式(如知识图谱的结构化程度、文本描述的完整性)也会制约智能体在专业场景中的泛化能力。

为消解知识库中噪声,未来需建立“事前筛选—事中校验—事后纠错”的全流程质量管控体系。来源验证优先采信权威数据库与高影响力文献的数据,标记低可信度来源;逻辑验证通过物理化学规

则校验数据合理性;多源交叉对比不同数据库、文献、模拟工具的同类数据,采用统计方法(如中位数滤波、异常值检测)过滤偏离度超标的噪声数据。

4.4 可信计算与可解释性框架

分子模拟智能体正逐步向催化剂逆向设计、高性能材料定制等工业级应用场景延伸,而 LLM 固有的“黑盒”推理模式与潜在幻觉风险成为了其实用化的核心瓶颈。工业场景不仅要求输出正确结果,更要求结果的产生过程可追溯、可验证、可审计。因此,构建多智能体系统的可信计算与可解释性框架将是未来的重要发展方向。具体而言,提升系统的工业可信度需从以下三个维度展开。

1 决策链透明化

智能体在进行物理模型选择或异常纠错时,需提供可追溯的显式推理路径。未来多智能体系统应要求核心调度智能体在生成每一步操作的同时,附带其推理所依据的物理规则、所引用的知识图谱节点或文献出处,使人类专家能追溯并审查 AI 的决策逻辑。

2 构建结构化日志管理体系

传统智能体的交互多以非结构化的文本对话呈现,难以进行严密的工程审计。未来需引入标准化的过程记录协议,将智能体的任务拆解图、API 调用状态、参数演变历史以及模拟软件的收敛轨迹等,统一映射为结构化的日志文件或图数据库。这种全流程的数据留存不仅便于事后溯源,也为引入专家审查机制提供了直观的数据接口。

3 不确定性量化与边界控制

智能体在处理未知体系或外推化学空间时,易因先验约束匮乏而产生误判风险。因此,需在智能体执行链条中嵌入不确定性评估模块。当智能体对某一反应路径的判断或核心参数配置的置信度低于预设安全阈值时,应自动触发熔断保护程序,主动请求专家介入或降级回退至更为保守的计算方法,从而在源头规避高昂的算力消耗与物理试错成本。

5 总结

分子模拟正经历从人工操作到自动化 workflow、再到智能体驱动的范式演进。本文首先阐述了传统分子模拟流程存在高度依赖专家经验、操作流程繁琐且难以标准化的问题。随后,介绍了自动化 workflow 在提升操作效率与可重复性方面的作用,并指出其仍受限于静态流程与预设逻辑,无法自主适应模拟过程中出现的非收敛、参数适配偏差等动态异常问题。

在此基础上,本文重点探讨了 LLM 赋能的分子模拟智能体技术,阐述了其构型自动建模、参数自适应配置、输入文件自动生成、多软件联合调度的全链路技术路径。分子模拟智能体集成了基于 LLM 的自然语言指令语义映射、复杂任务的原子化拆解与规划、全流程工具链的编排以及异常状态的闭环修正等机制,同时借助知识检索增强技术实现领域先验知识的积累和复用,弥补了传统分子模拟方案的静态局限。

尽管现有智能体框架已在材料计算、药物发现等多个场景中展现出潜力,但仍面临多模态感知深度不足、知识更新机制不完善、错误修正依赖人工规则等挑战。未来研究应致力于推动多模态信息的端到端融合、发展知识驱动的自主进化机制、构建高质量知识库与噪声清洗体系,进一步提升智能体在复杂、动态模拟场景中的适应性与鲁棒性,最终实现分子模拟从“工具辅助”到“智能协同”的范式转变。

参考文献

- [1] DeFever R S, Targonski C, Hall S W, et al. A generalized deep learning approach for local structure identification in molecular simulations[J]. *Chemical Science*, 2019, **10**(32): 7503–7515.
- [2] Michel J. Current and emerging opportunities for molecular simulations in structure-based drug design[J]. *Physical Chemistry Chemical Physics*, 2014, **16**(10): 4465–4477.
- [3] De Vivo M, Masetti M, Bottegoni G, et al. Role of molecular dynamics and related methods in drug discovery[J]. *Journal of Medicinal Chemistry*, 2016, **59**(9): 4035–4061.
- [4] Vergadou N, Theodorou D N. Molecular modeling investigations of sorption and diffusion of small molecules in glassy polymers[J]. *Membranes*, 2019, **9**(8): 98.
- [5] Kee C W. Molecular understanding and practical in silico catalyst design in computational organocatalysis and phase transfer catalysis—Challenges and opportunities[J]. *Molecules*, 2023, **28**(4): 1715.
- [6] Liu H, Jin Y, Ding H J. MDBuilder: a PyMOL plugin for the preparation of molecular dynamics simulations[J]. *Briefings in Bioinformatics*, 2023, **24**(2): bbad057.
- [7] Thompson M W, Gilmer J B, Matsumoto R A, et al. Towards molecular simulations that are transparent, reproducible, usable by others, and extensible (TRUE)[J]. *Molecular Physics*, 2020, **118**(9/10): e1742938.
- [8] Knapp B, Ospina L, Deane C M. Avoiding false positive conclusions in molecular simulation: the importance of replicas[J]. *Journal of Chemical Theory and Computation*, 2018, **14**(12): 6127–6138.
- [9] Hjorth Larsen A, Jørgen Mortensen J, Blomqvist J, et al. The atomic simulation environment—a Python library for working with atoms[J]. *Journal of Physics: Condensed Matter*, 2017, **29**(27): 273002.
- [10] Horton M. Materials discovery screening with pymatgen[J]. *Nature Reviews Physics*, 2023, **5**(2): 76.
- [11] Birkenheuer G, Blunk D, Breuers S, et al. MoSGrid – a molecular simulation grid as a new tool in computational chemistry, biology and material science[J]. *Journal of Cheminformatics*, 2011, **3**(1): P14.
- [12] Di Tommaso P, Chatzou M, Floden E W, et al. Nextflow enables reproducible computational workflows[J]. *Nature Biotechnology*, 2017, **35**(4): 316–319.
- [13] Wei J, Tay Y, Bommasani R, et al. Emergent abilities of large language models[EB/OL]. 2022: arXiv: 2206.07682. <https://arxiv.org/abs/2206.07682>
- [14] OpenAI, Achiam J, Adler S, et al. GPT-4 technical report[EB/OL]. 2023: arXiv:2303.08774. <https://arxiv.org/abs/2303.08774>
- [15] Ouyang L, Wu J, Jiang X, et al. Training language models to follow instructions with human feedback[C]//Advances in Neural Information Processing Systems 35. 2022: 27730–27744.
- [16] Kumar P. Large language models (LLMs): survey, technical frameworks, and future challenges[J]. *Artificial Intelligence Review*, 2024, **57**(10): 260.
- [17] Wang X H, Jiang H Y, Yu Y, et al. Building intelligence identification system via large language model watermarking: a survey and beyond[J]. *Artificial Intelligence Review*, 2025, **58**(8): 249.
- [18] Xiao H G, Zhou F Z, Liu X Y, et al. A comprehensive survey of large language models and multimodal large language models in medicine[J]. *Information Fusion*, 2025, **117**: 102888.
- [19] Yao S Y, Yu D, Zhao J, et al. Tree of thoughts: deliberate problem solving with large language models[C]//Advances in Neural Information Processing Systems 36. New Orleans: Neural Information Processing Systems Foundation, 2023: 11809–11822.
- [20] Wei J, Wang X Z, Schuurmans D, et al. Chain-of-thought prompting elicits reasoning in large language models[C]//Advances in Neural Information Processing Systems 35. New Orleans: Neural Information Processing Systems Foundation, 2022: 24824–24837.
- [21] Taylor R, Kardas M, Cucurull G, et al. Galactica: a large language model for science[EB/OL]. 2022: arXiv:2211.09085. <https://arxiv.org/abs/2211.09085>
- [22] Wang L, Ma C, Feng X Y, et al. A survey on large language model based autonomous agents[J]. *Frontiers of Computer Science*, 2024, **18**(6): 186345.
- [23] Boiko D A, MacKnight R, Kline B, et al. Autonomous chemical

- research with large language models[J]. *Nature*, 2023, **624**(7992): 570–578.
- [24] M Bran A, Cox S, Schilter O, et al. Augmenting large language models with chemistry tools[J]. *Nature Machine Intelligence*, 2024, **6**(5): 525–535.
- [25] Venugopal V, Olivetti E. MatKG: an autonomously generated knowledge graph in Material Science[J]. *Scientific Data*, 2024, **11**(1): 217.
- [26] Statt M J, Rohr B A, Guevarra D, et al. The materials experiment knowledge graph[J]. *Digital Discovery*, 2023, **2**(4): 909–914.
- [27] Wooldridge M, Jennings N R. Intelligent agents: theory and practice[J]. *The Knowledge Engineering Review*, 1995, **10**(2): 115–152.
- [28] Xi Z H, Chen W X, Guo X, et al. The rise and potential of large language model based agents: a survey[J]. *Science China Information Sciences*, 2025, **68**(2): 121101.
- [29] Qin Y J, Liang S H, Ye Y N, et al. ToolLLM: Facilitating large language models to master 16000+ real-world APIs[C]// *International Conference on Learning Representations*. 2024: 9695–9717.
- [30] Shinn N, Cassano F, Gopinath A, et al. Reflexion: Language agents with verbal reinforcement learning[C]// *Advances in Neural Information Processing Systems 36*. New Orleans: Neural Information Processing Systems Foundation, 2023: 8634–8652.
- [31] Summers T R, Yao S Y, Narasimhan K, et al. Cognitive architectures for language agents[J]. *Transactions on Machine Learning Research*, 2023.
- [32] Durante Z, Huang Q, Wake N, et al. Agent AI: surveying the horizons of multimodal interaction[EB/OL]. 2024: arXiv: 2401.03568. <https://arxiv.org/abs/2401.03568>
- [33] Wang R Y, Jansen P, Côté M A, et al. ScienceWorld: is your agent smarter than a 5th grader? [C]// *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*. Abu Dhabi: Association for Computational Linguistics, 2022: 11279–11298.
- [34] Shen Y L, Song K T, Tan X, et al. HuggingGPT: solving AI tasks with ChatGPT and its friends in Hugging Face[C]// *Advances in Neural Information Processing Systems 36*. New Orleans: Neural Information Processing Systems Foundation, 2023: 38154–38180.
- [35] Yang Z Y, Li L J, Wang J F, et al. MM-REACT: prompting ChatGPT for multimodal reasoning and action[EB/OL]. 2023: arXiv: 2303.11381. <https://arxiv.org/abs/2303.11381>
- [36] Huang R J, Li M Z, Yang D C, et al. AudioGPT: understanding and generating speech, music, sound, and talking head[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, **38**(21): 23802–23804.
- [37] Liu H T, Li C Y, Wu Q Y, et al. Visual instruction tuning[C]// *Advances in Neural Information Processing Systems 36*. New Orleans: Neural Information Processing Systems Foundation, 2023: 34892–34916.
- [38] OpenAI. GPT-4o system card[EB/OL]. 2024: arXiv: 2410.21276. <https://arxiv.org/abs/2410.21276>
- [39] Gemini Team, Google. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context[EB/OL]. 2024: arXiv: 2403.05530. <https://arxiv.org/abs/2403.05530>
- [40] Wu S Q, Fei H, Qu L G, et al. NExT-GPT: Any-to-any multimodal LLM[C]// *Proceedings of the 41st International Conference on Machine Learning*. Vienna: PMLR, 2024.
- [41] Kojima T, Gu S S, Reid M, et al. Large language models are zero-shot reasoners[C]// *Advances in Neural Information Processing Systems 35*. New Orleans: Neural Information Processing Systems Foundation, 2022: 22199–22213.
- [42] Wang X Z, Wei J, Schuurmans D, et al. Self-consistency improves chain of thought reasoning in language models[EB/OL]. 2022: arXiv: 2203.11171. <https://arxiv.org/abs/2203.11171>
- [43] Zhou A, Yan K, Shlapentokh-Rothman M, et al. Language agent tree search unifies reasoning acting and planning in language models[EB/OL]. 2023: arXiv: 2310.04406. <https://arxiv.org/abs/2310.04406>
- [44] Yao S Y, Zhao J, Yu D, et al. ReAct: Synergizing reasoning and acting in language models[EB/OL]. 2022: arXiv: 2210.03629. <https://arxiv.org/abs/2210.03629>
- [45] OpenAI. OpenAI o1 system card[EB/OL]. 2024: arXiv: 2412.16720. <https://arxiv.org/abs/2412.16720>
- [46] Sang J T, Xiao J L, Han J R, et al. Beyond pipelines: A survey of the paradigm shift toward model-native agentic AI[EB/OL]. 2025: arXiv: 2510.16720. <https://arxiv.org/abs/2510.16720>
- [47] Schick T, Dwivedi-Yu J, Dessì R, et al. Toolformer: Language models can teach themselves to use tools[C]// *Advances in Neural Information Processing Systems 36*. New Orleans: Neural Information Processing Systems Foundation, 2023: 68539–68551.
- [48] Kang Y, Kim J. ChatMOF: an artificial intelligence system for predicting and generating metal-organic frameworks using large language models[J]. *Nature Communications*, 2024, **15**(1): 4705.
- [49] Gao Y F, Xiong Y, Gao X Y, et al. Retrieval-augmented generation for large language models: A survey[EB/OL]. 2023: arXiv: 2312.10997. <https://arxiv.org/abs/2312.10997>
- [50] Zhao A, Huang D, Xu Q, et al. ExpeL: LLM agents are experiential learners[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, **38**(17): 19632–19642.
- [51] Cai T F, Zhang N G, Li Y S, et al. Occam’s razor in AI-assisted complex diagnosis: a comparative effectiveness study of single large language models versus multi-agent systems in resource-constrained primary care settings[R/OL]. 2026. <https://ssrn.com/abstract=6148075>
- [52] Maia E H B, Medaglia L R, Da Silva A M, et al. Molecular architect: a user-friendly workflow for virtual screening[J]. *ACS Omega*, 2020, **5**(12): 6628–6640.
- [53] Gjerding M, Skovhus T, Rasmussen A, et al. Atomic simulation recipes: a python framework and library for automated workflows [J]. *Computational Materials Science*, 2021, **199**: 110731.
- [54] Yekeen A A, Durojaye O A, Idris M O, et al. CHAPERONg: a tool for automated GROMACS-based molecular dynamics simulations and trajectory analyses[J]. *Computational and Structural Biotechnology Journal*, 2023, **21**: 4849–4858.
- [55] Ries B, Alibay I, Anand N M, et al. Automated absolute binding free energy calculation workflow for drug discovery[J]. *Journal of Chemical Information and Modeling*, 2024, **64**(14): 5357–5364.
- [56] Schmitz G, Yönder Ö, Schnieder B, et al. An automatized workflow from molecular dynamic simulation to quantum chemical methods to identify elementary reactions and compute reaction constants[J]. *Journal of Computational Chemistry*, 2021, **42**(32): 2264–2282.
- [57] Singh H, Raja A, Prakash A, et al. Gmx_qk: an automated protein/

- protein – ligand complex simulation workflow bridged to MM/PBSA, based on gromacs and zenity-dependent GUI for beginners in MD simulation study[J]. *Journal of Chemical Information and Modeling*, 2023, **63**(9): 2603–2608.
- [58] Löhr T, Assante M, Dodds M, et al. Navigating the Maize: cyclic and conditional computational graphs for molecular simulation[J]. *Digital Discovery*, 2024, **3**(12): 2551–2559.
- [59] Jeong P U, Sørensen J, Vemu P L, et al. Progress towards automated Kepler scientific workflows for computer-aided drug discovery and molecular simulations[J]. *Procedia Computer Science*, 2014, **29**: 1745–1755.
- [60] Doerr S, Harvey M J, Noé F, et al. HTMD: high-throughput molecular dynamics for molecular discovery[J]. *Journal of Chemical Theory and Computation*, 2016, **12**(4): 1845–1852.
- [61] Hayashi Y, Shiomi J, Morikawa J, et al. RadonPy: automated physical property calculation using all-atom classical molecular dynamics simulations for polymer informatics[J]. *npj Computational Materials*, 2022, **8**(1): 222.
- [62] Gygli G, Pleiss J. Simulation foundry: automated and F.A.I.R. molecular modeling[J]. *Journal of Chemical Information and Modeling*, 2020, **60**(4): 1922–1927.
- [63] Bayarri G, Andrio P, Hospital A, et al. BioExcel Building Blocks Workflows (BioBB-Wfs), an integrated web-based platform for biomolecular simulations[J]. *Nucleic Acids Research*, 2022, **50** (W1): W99–W107.
- [64] Liu J X, Zhu T N, Ye C Y, et al. VASPilot: MCP-facilitated multi-agent intelligence for autonomous VASP simulations[J]. *Chinese Physics B*, 2025, **34**(11): 117106.
- [65] Vriza A, Kornu U, Koneru A, et al. Multi-agent AI framework for end-to-end atomistic simulations[J]. *Digital Discovery*, 2026, **5** (1): 440–452.
- [66] Mendible-Barreto O A, Díaz-Maldonado M, Carmona Esteva F J, et al. DynaMate: leveraging AI-agents for customized research workflows[J]. *Molecular Systems Design & Engineering*, 2025, **10** (7): 585–598.
- [67] Pham T D, Tanikanti A, Keçeli M. ChemGraph as an agentic framework for computational chemistry workflows[J]. *Communications Chemistry*, 2026, **9**(1): 33.
- [68] Ghafarollahi A, Buehler M J. AtomAgents: Alloy design and discovery through physics-aware multi-modal multi-agent artificial intelligence[EB/OL]. 2024: arXiv: 2407.10022. <https://arxiv.org/abs/2407.10022>
- [69] Zou Y H, Cheng A H, Aldossary A, et al. El Agente: an autonomous agent for quantum chemistry[J]. *Matter*, 2025, **8**(7): 102263.
- [70] Campbell Q, Cox S, Medina J, et al. MDCrow: Automating molecular dynamics workflows with large language models[J]. *Machine Learning: Science and Technology*, 2026, **7**(2): 025037.
- [71] Petković M, Menkovski V, Calero S. Towards fully automated molecular simulations: Multi-agent framework for simulation setup and force field extraction[EB/OL]. 2025: arXiv:2509.10210. <https://arxiv.org/abs/2509.10210>
- [72] Ding K Y, Yu J, Huang J J, et al. SciToolAgent: A knowledge-graph-driven scientific agent for multitool integration[J]. *Nature Computational Science*, 2025, **5**(10): 962–972.
- [73] Zhang B H, Li X, Xu H C, et al. TopoMAS: Large language model driven topological materials multiagent system[EB/OL]. 2025. <https://arxiv.org/abs/2507.04053>
- [74] Wang Z Q, Huang H S, Zhao H C, et al. DREAMS: Density functional theory based research engine for agentic materials simulation[EB/OL]. 2025: arXiv:2507.14267. <https://arxiv.org/abs/2507.14267>