

人工智能代理在药物研发中的应用进展

赵东海, 谢昌谕

浙江大学药学院, 浙江 杭州 310058

[摘要] 药物研发面临着高技术、高投入、高风险和长周期的严峻挑战,亟需能够系统性提升效率、精准预测并降低失败风险的颠覆性技术。人工智能代理作为一种由大语言模型驱动的新兴智能范式,正展现出重塑药物研发全流程的巨大潜力。其核心在于能够围绕复杂的科学目标进行自主推理、规划和工具调用,整合并协调不同研发环节,实现人工智能从“工具”到“主动合作者”的角色转变。一方面,人工智能代理通过知识整合及假设生成,能够发现未被充分探索的新靶点和新机制;另一方面,它能自动化地执行从分子设计、优化到合成规划等复杂任务,并通过与自动化实验平台深度结合,实现从虚拟设计到物理实验的闭环。在此基础上,人工智能代理正向更高层级的范式演进,构建综合性药物设计平台和发展通用型生物医学智能体。本文系统阐述了人工智能代理的核心框架及其在药物研发关键环节上的应用进展,并探讨了目前存在的局限及未来的发展方向,旨在为相关领域的研究提供参考。



[关键词] 人工智能代理; 药物研发; 大语言模型; 多智能体系统; 综述

[中图分类号] R91; TP18 **[文献标志码]** A

Application and progress of artificial intelligence agents in drug development

ZHAO Donghai, HSIEH Changyu (School of Pharmacy, Zhejiang University, Hangzhou 310058, China)

Corresponding author: HSIEH Changyu, E-mail: kimhsieh@zju.edu.cn, ORCID: 0000-0002-6242-4218

[Abstract] Drug discovery faces formidable challenges including high technology, high costs, substantial risks, and prolonged development timelines, necessitating disruptive technologies capable of systematically improving efficiency, enhancing predictive accuracy,

收稿日期(Received): 2025-09-08 修改返回日期(Revised): 2025-11-10 接受日期(Accepted): 2025-11-22 网络预发表日期(Online): 2026-01-01

基金项目(Funding): 国家自然科学基金(22373085); 浙江省自然科学基金(LHDMZ25H300001)

第一作者(First author): 赵东海, 硕士研究生, 主要从事加速药物设计的人工智能代理研究; E-mail: 22319109@zju.edu.cn; ORCID: 0009-0007-0315-2543

通信作者(Corresponding author): 谢昌谕, 教授, 博士生导师, 主要从事前沿计算技术在药物研发领域的应用研究; E-mail: kimhsieh@zju.edu.cn; ORCID: 0000-0002-6242-4218

and reducing failure rates. Artificial intelligence (AI) agent—an emerging intelligent paradigm powered by large language models—holds significant potential to transform the entire drug development pipeline. Their core capability lies in performing autonomous reasoning, planning, and tool utilization directed at complex scientific objectives, thereby integrating and orchestrating multiple research stages and transitioning AI from a mere “tool” to an “active collaborator”. Through knowledge integration and hypothesis generation, AI agents can identify underexplored therapeutic targets and novel mechanisms of action. In parallel, they can automate complex tasks such as molecular design, optimization, and synthesis planning, and further close the loop between virtual design and physical experimentation by interfacing with automated experimental platforms. Moreover, AI agents are evolving toward higher-level paradigms, including the development of integrated drug design platforms and general-purpose biomedical agents. This review systematically summarizes the core architectures of AI agents, highlights their applications across key stages of drug development, and discusses current limitations along with future directions, providing a reference for researches in related fields.

[**Key words**] Artificial intelligence agent; Drug development; Large language model; Multi-agent system; Review

[J Zhejiang Univ (Med Sci), 2026, 55(1): 4-15.]

[**缩略语**] 人工智能(artificial intelligence, AI);推理与行动协同(Reasoning and acting, ReAct);检索增强生成(retrieval-augmented generation, RAG);设计-合成-测试-分析(design-make-test-analyze, DMTA);组蛋白脱乙酰酶(histone deacetylase, HDAC);DNA 甲基化酶 1(DNA methyltransferase 1, DNMT1);含溴结构域蛋白 4(bromodomain-containing protein 4, BRD4);成壳型噬菌体诱导染色体岛(capsid-forming phage-inducible chromosomal island, cf-PICI);G 蛋白偶联受体 160(G protein-coupled receptor 160, GPR160);精氨酸酶 2(arginase 2, ARG2);严重急性呼吸综合征冠状病毒 2(severe acute respiratory syndrome coronavirus-2, SARS-CoV-2)

药物研发长期深陷于高技术、高投入、高风险和长周期的困境,传统依赖大规模实验试错的研究模式已愈发难以为继。在此背景下,寻求能系统性提升效率、精准预测并降低失败风险的新方法已成为制药行业的迫切需求^[1-3],而这直接推动了计算科学在药物研发中的应用范式不断升级^[4-5]。

随着深度学习技术的突破,AI在解决关键科学问题上展现出强大的能力。高精度预测模型可用于评估不同药物形式的生物活性和安全性^[6-8]。以变分自编码器、生成对抗网络、变换器以及扩散模型为代表的生成式 AI 技术,能够“从头设计”具有特定功能的候选药物,涵盖了从新颖小分子到

复杂蛋白质序列的广泛范围^[9-12],标志着 AI 在药物研发中的角色从“评估者”转变为“创造者”。然而,无论是预测还是生成模型,其应用都存在明显的“任务边界”,虽能够高效解决特定问题,却无法自主整合多步骤的复杂研发流程。人工智能代理(即智能体)是一种能感知环境、围绕目标进行自主决策、规划行动并与外部工具交互的智能系统^[13-14]。它与传统数据驱动 AI 模型的最根本区别在于,智能体能够将一个高层次的、模糊的目标(例如为某靶点寻找一种新型抑制剂)自主分解为一系列逻辑连贯且可执行的子任务,并按计划调用工具来完成这些任务^[15-16]。智能体能够通过自动化整个研究过程而非单个步骤来系统性地提

升研发效率,为解决药物研发的根本性困境提供了全新思路。

当前,大语言模型的能力飞跃正催化着智能体的范式革命:AI正从被动执行明确指令的工具,演变为能自主理解复杂目标、进行推理规划并执行多步任务的“主动合作者”,这一演变有望从根本上重塑药物研发的全貌^[14, 17]。鉴于智能体的快速发展和巨大潜力,本文系统综述智能体在药物设计领域的最新研究进展:首先剖析智能体的核心技术框架,随后介绍其在药物研发关键环节的应用,最后探讨当前面临的挑战并展望未来的发展方向,以期为相关领域的研究提供参考。

1 智能体的核心架构

智能体是能够自主运行的智能系统,它通过感知数字或物理环境、围绕科学目标进行推理,制订一系列行动计划,并利用一套外部工具来执行这些计划^[14]。为实现这一系列复杂的自主行为,其系统架构的重要特征在于不同模块间的协同工作。这种架构既可以表现为单个智能体内部精密配合的多个功能模块^[18],也可以扩展为由多个具备不同专业能力的独立智能体组成的多智能体系统^[19]。后者通过将宏观任务分解,让多个智能体协同工作,以解决单一智能体难以应对的复杂科学问题。无论是单一智能体还是多智能体系统,其核心架构主要由推理和规划、检索增强生成、记忆系统和工具使用四大模块构成(图1),它们共同支撑智能体的自主运行。

1.1 推理和规划

在智能体的架构中,大语言模型扮演着核心认知引擎的角色,负责推理、规划和反思等关键

任务^[20]。通过思维链^[21]等提示词工程,可以引导大语言模型生成中间推理步骤,以完成复杂的多步推理任务。在此基础上,ReAct框架进一步将推理与行动相结合,形成“推理—行动—思考”的决策闭环^[20]。为了提升完成任务的成功率,自我反思框架为智能体引入了更强的自我反思机制,使其能在ReAct循环中进行自我评估、修正和记忆,根据反馈优化决策^[22]。例如,新一代推理模型DeepSeek R1在训练过程中内化了这种推理和反思过程,即在生成最终答案之前生成思考过程^[23]。

大语言模型的推理和规划能力已应用于药物设计中,其典型范例是ReAct框架用于分子优化^[20](图2)。在优化一个高活性但成药性差的激酶抑制剂时,智能体首先推理出当前分子的核心缺陷(用户提供描述或调用工具计算),随后调用分子编辑器、对接模拟和计算化学工具,执行具体的化学修饰并预测新分子的各项属性。通过评估结果,智能体启动下一轮推理,针对性地逐步修正分子结构直至筛选出符合优化目标的候选分子。

1.2 检索增强生成

仅依靠大语言模型内部知识进行推理存在知识陈旧、事实性幻觉等固有局限,RAG是解决这些问题的关键技术,它通过结合外部知识库使大语言模型的输出更具事实准确性和时效性^[24-25]。RAG的核心流程分为三个步骤:首先,通过向量嵌入(将文本信息转化为计算机能理解和比较的数字坐标)等技术从外部知识库中检索相关事实;然后,将检索到的事实与用户的原始问题整合,构建一个信息更丰富的增强型提示;最后,大语言模型基于增强型提示生成一个有事实依据

且更可靠的回答^[24]。这一过程有效地将大语言模型的推理与外部事实相结合,大幅提升了回答的准确性。在药物研发中,基于RAG技术的智能体通过动态检索生物医学文献和专利中的最新证据,对此类多元证据进行整合及推理,生成可靠的靶点假设、药物—靶点相互作用预测及合成路径规划,支持从虚拟设计到实验验证的闭环研发流程^[26]。

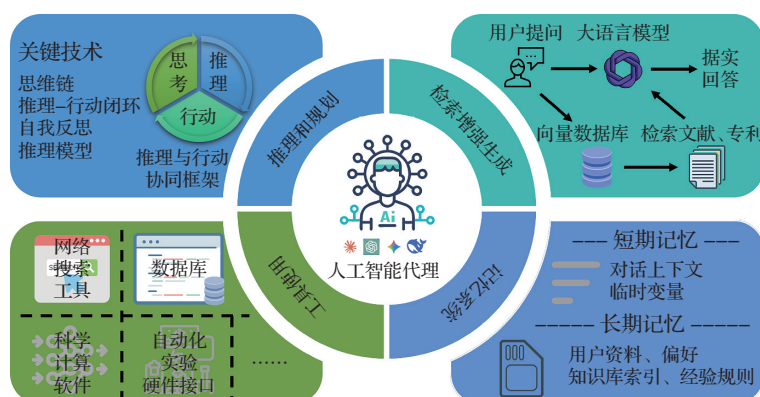
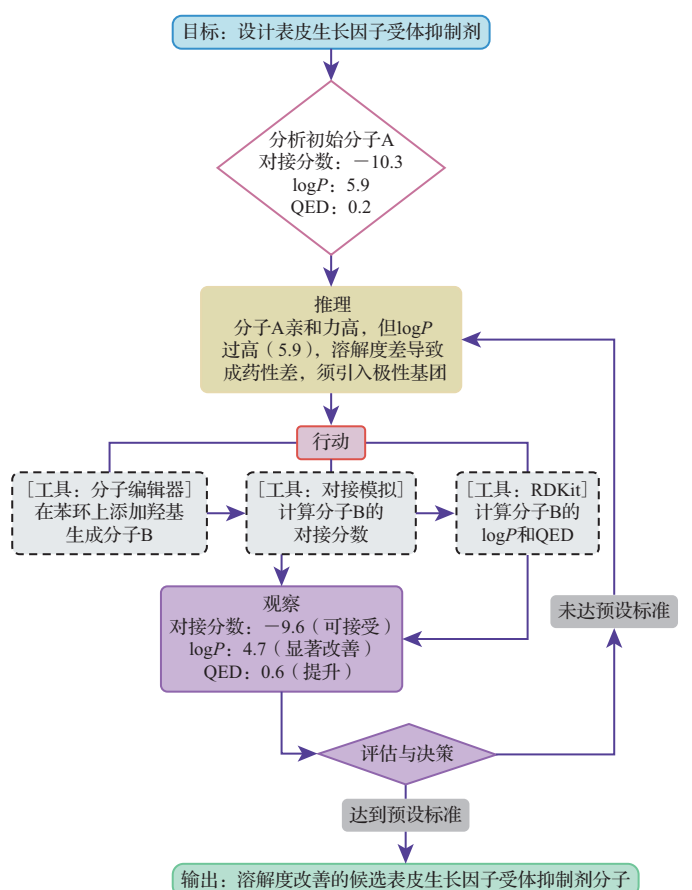


图1 人工智能代理的核心架构

Figure 1 Core architectures of artificial intelligence agent



logP: 疏水常数(辛醇-水分配系数的对数值); QED: 定量估计类药性;
RDKit: 化学信息学工具包; ReAct: 推理与行动协同。

图2 基于ReAct框架的药物分子设计优化决策流程图示例

Figure 2 Example of molecular optimization decision flowchart based on the ReAct framework

传统RAG技术的有效性高度依赖于外部知识库的质量,一个充斥着过时、偏倚甚至错误信息的知识库,会直接损害最终生成内容的可靠性。为了应对这一挑战,研究者引入了更为先进的RAG技术(如GraphRAG^[27])以增强鲁棒性。在知识源筛选和整合层面,一些研究工作开始采用混合检索策略,如BioLab结合使用了知识图谱(一种用于表示实体及其相互关系的结构化网络)等结构化数据源和非结构化文本数据库^[28]。知识图谱能够提供经过验证的、具有明确关系的事实,从而在检索阶段引入一层事实校验;同时,通过对多源信息进行来源评估和交叉验证,动态赋予其权重,从而在生成答案时优先采纳来自高权威性文献或数据库的内容。在处理信息偏倚和不一致性方面,智能体正从被动的“检索-生成”模式演变为主动的“反思-修正”模式^[29-30]。例如,智能体可以在生成初步答案后,启动一个自

我批判环节,主动检索能够证实或证伪其初步结论的补充证据。当检测到来自不同数据源的矛盾信息时,智能体能够标记这些冲突点,并依据预设的方式(如采纳最新数据、将冲突提交给人类专家裁决)处理,从而显著提升了在复杂信息环境中生成准确、可靠答案的能力^[29-30]。此外,为应对实时变化的信息环境,RAG技术进一步引入了动态知识库机制(如增量索引和数据流整合),实现知识库的持续更新和高效维护^[31-32]。随着智能体技术的发展和大型语言模型自身能力的提升,传统RAG技术逐渐演化为智能体RAG,其典型应用体现为各类深度研究的智能体框架,这些框架能自主完成信息搜集和整合,生成详实的调研报告^[33],在该模式下,智能体能够主动规划、迭代优化检索和生成流程,实现从静态检索到自主研究的跨越^[34]。

此外,RAG技术的应用已从知识检索扩展到工具检索。面对日益增多的可用工具和大型语言模型有限的上下文长度,智能体需要一种动态选择工具的机制。例如,SciToolAgent^[35]

和Biomni^[36]可通过RAG技术检索完成任务所需的工具,有效地解决了工具库规模超出模型上下文容量的难题。

1.3 记忆系统

RAG技术为智能体提供了动态访问外部知识库的能力,但真正自主的智能体不仅能即时查询事实,还具备记忆能力:将交互过程中的经验、结论和失败教训持续沉淀。因此,记忆模块应运而生,其可负责存储和检索关键信息,涵盖了用于追踪当前任务的短期记忆和积累经验知识的长期记忆,为智能体执行长期、复杂的任务提供支持。短期记忆通常依赖于大型语言模型的上下文窗口,但其容量和持久性受限于窗口大小,会随着交互结束而丢失^[37]。长期记忆则负责捕捉用户偏好、积累在执行不同任务的过程中获得的经验和知识,必须借助外部存储。其通常需要引入外部向量数据库,通过RAG技术进行内容的存

储和动态访问,从而为智能体提供持久的记忆支持^[38]。这种长期记忆对于需要持续迭代推理的复杂生物医学研究至关重要。例如,STELLA系统引入了模板库作为其长期记忆的载体。在成功揭示化疗耐药性的分子机制后,该系统能将验证过的推理 workflow 提炼为结构化模板并存入记忆库;当启动新一轮类似研究时,智能体便能主动检索并复用这些成功策略。这种基于外部存储的记忆机制有效突破了传统大语言模型上下文窗口的限制,确保了智能体在长期、复杂的研发任务中持续积累经验并避免重复试错^[39]。

1.4 工具使用

在获得知识和记忆的支持后,智能体还需进一步将认知层面的推理和记忆转化为可执行的外部行动。工具使用模块是其连接外部世界、执行物理操作的通道,它如同智能体的双手,负责将大语言模型生成的抽象行动指令精准解析为对特定软件、数据库或自动化设备应用程序编程接口的调用和代码执行的命令,从而将内部规划落实为外部世界的实际操作^[40]。这些工具的范围较为广泛,包括用于信息检索的网络搜索工具、执行复杂算法的科学计算软件、存储结构化数据的数据库,以及用于操控自动化实验设备的硬件接口等。工具调用高度依赖于大语言模型的格式化输出能力,需按要求将工具调用请求转化为 json 格式或其他格式进行输出。为确保不同工具之间无缝协作,智能体的工具调用模块采用标准化接口协议,最常用的是 OpenAI 公司的函数调用协议和 Anthropic 公司的模型上下文协议,后者不仅在接口协议层面提供了统一的工具描述、输入输出格式和调用约束,还在通信协议层面规范了智能体和工具之间的消息交换和调用流程,从而实现跨平台、多工具的高效协作^[41]。

2 智能体在药物研发中的前沿应用

智能体正逐步渗透到药物研发的全周期中,展现出重塑传统研发模式的巨大潜力。以下先探讨智能体如何赋能药物研发中的核心任务,包括知识整合及假设生成、先导化合物生成和优化、生物大分子设计以及自动化实验平台构建;在此基础上进一步探讨智能体向更高级范式的演进,即构建综合性药物设计平台和发展通用型生物医学智能体。

2.1 赋能特定研发任务

2.1.1 知识整合及假设生成 由谷歌、斯坦福大学等机构共同开发的 AI co-scientist 是一个具有代表性的多智能体系统,旨在模拟并辅助科研人员提出创新性科学假设^[19]。该系统通过“生成—反思—排名”的迭代探索框架,调用网络搜索等外部工具,并融合人类研究员的关键反馈,形成一个自我修正的闭环。其性能体现在 GPQA 科学问答基准^[42]上的准确率达到了 78.4%,超越人类专家的平均水平^[19]。其潜力已在多个实际研究场景中得到验证:在针对肝纤维化的研究中,该系统自主提出了三种新颖的表现遗传学靶点(HDAC、DNMT1 和 BRD4),随后的实验证实了 HDAC 和 BRD4 抑制剂具备明显的抗纤维化活性^[43];在针对 cf-PICI 跨物种基因转移这一此前未公开的生物学机制研究中,该系统独立推理出“囊膜—尾部劫持”核心机制假说,该假说精准预测了人类耗时多年才通过实验揭示的机制^[44-45]。上述研究结果有力证明了 AI co-scientist 系统具备绕过人类固有认知路径、直接接触及复杂生物学底层机制的能力。

在 AI co-scientist 系统的基础上,后续研究进一步扩展了智能体在科研工作流中的角色和能力。例如,Robin 系统通过整合多个功能专一的智能体(Crow 负责从海量文献中提取信息,Falcon 负责对候选假设进行科学评估,Finch 负责对实验数据进行编程分析和可视化),驱动了一个“假设—实验—分析—再假设”的半自主化科研闭环^[46]。为验证其核心筛选能力,该系统在研究干性年龄相关性黄斑变性的过程中生成了 30 个治疗候选假设,并通过大语言模型进行成对比较以生成排名,该排序结果与人类专家筛选出的前 10 位治疗候选高度吻合,且大语言模型裁判在重复评估相同的成对比较任务时展现出 88% 的自我决策一致性。而且,Finch 在后续的 RNA 测序分析中也表现出高可靠性,超过 50% 的独立分析流程均能识别出相同的差异表达基因。基于这种经过验证的可靠工作流,Robin 系统在实践中成功通过迭代循环发现并验证了利帕舒地尔(Ripasudil)作为治疗干性年龄相关性黄斑变性的新型候选药物^[46]。OriGene 系统则将智能体的角色进一步聚焦为“虚拟疾病生物学家”,引入了独特的自我进化能力,使其能从过往成功的解决方

案中学习并提炼新的思维模板,在“靶点发现”任务中实现了系统性能的持续自我提升^[47]。该系统在定制的靶点检索问答基准测试中,多项选择题的准确率达 60.08%,超过人类专家的平均水平。该系统已成功自主识别并指导研究人员实验验证了肝癌靶点(GPR160)和结直肠癌靶点(ARG2)^[47]。

2.1.2 先导化合物生成和优化 在确定药物靶点后,研发工作的核心便转向分子设计和优化,即在广阔化学空间中寻找兼具活性和成药性的分子。针对这一复杂的优化难题,现有 AI 的价值在于充当智能协调者,而非取代海量的计算工具。例如,ChatChemTS 系统旨在帮助用户通过自然语言来操作 ChemTSv2 这一分子设计框架^[48],自动化进行参数配置和结果分析等繁琐步骤,从而降低 ChemTSv2 的使用门槛^[49]。MultiMol 系统则通过数据驱动的执行智能体和文献指导的研究智能体协同工作,模拟化学专家的工作流程,专门用于指导复杂的多目标分子优化^[50]。该系统在六项多目标分子优化任务的评估中成功率高达 82.30%,远超当前最强基线方法(成功率为 27.50%),展现出卓越性能;在实际应用层面,该系统成功优化了黄嘌呤胺同系物的选择性,并改善了沙奎那韦的生物利用度^[50]。Fleming 系统则展示了多智能体框架在解决真实世界复杂药物研发问题中的强大潜力,该系统专为研发抗结核药物而设计,通过中央管理智能体协调多个专业的子智能体进行工作,这些专业子智能体分别负责分子设计、活性预测,以及成药性评估。通过自动化整合分子设计的复杂流程,该系统能高效地发现具有潜力的新型抗菌药物,经药物化学家评估,其筛选出的候选分子中,高达 82.6% 被认定为合理且值得进一步开发^[51]。

2.1.3 生物大分子设计 智能体的应用已扩展到抗体、多肽等结构更复杂、研发挑战更高的生物大分子药物设计领域。该领域代表性工作之一是 Swanson 等^[52]开发的“虚拟实验室”(一个模仿人类科研团队的多智能体协作系统)。该系统由首席研究员智能体领导,下设多学科科学家智能体(如免疫学家、机器学习专家等),并由科学评论员智能体提供反馈,同时接受人类研究员更高层次指导。该系统成功应用于针对 SARS-CoV-2 变体的纳米抗体设计中,其建立的迭代流程整合

了 ESM、AlphaFold-Multimer 和 Rosetta 等工具,最终,该系统在数天内设计出 92 种新型纳米抗体,经实验验证超过 90% 的抗体具有可表达性和可溶性,其中两种抗体成功展现了对 SARS-CoV-2 新型变体(如 JN.1 和 KP.3)的结合能力^[52]。另一项代表性工作是 BioLab^[28],一个旨在实现端到端自主生物学的多智能体系统。其架构基于“劳动分工”原则,由规划智能体(分解任务)、推理智能体(制订计划)、xBio-Tools 智能体(调用工具)、评论家智能体(质量控制)和记忆智能体(迭代学习)等协同工作。这种精密的协作架构使其在多个生物医学推理基准测试中表现持续优于 GPT-4o、Gemini-2.5 等前沿大语言模型^[28]。在针对程序性死亡受体 1 靶点的 T 细胞免疫疗法研究中,BioLab 自主完成了从靶点挖掘到抗体优化的全流程,经湿实验验证,其设计出的优化抗体(Pem-MOO-2)半抑制浓度值(0.01 nmol/L)显著优于帕博利珠单抗(0.027 nmol/L),有力证明了该系统在抗体设计上的可靠性和应用潜力^[28]。

2.1.4 自动化实验平台构建 在药物研发中,实验验证通常是成本最高、耗时最长的环节。传统实验依赖人工操作,效率有限且难以实现大规模迭代优化。随着自动化设备和云实验室的快速发展,智能体正逐渐与自动化实验平台深度结合,形成了能够“自主设计—执行—分析”的闭环实验体系,不仅加速了候选分子的合成,还提升了实验可重复性和可扩展性。Boiko 等^[53]开发的 Coscientist 系统是该领域的开创性工作之一。该系统由 GPT-4 大语言模型驱动,能自主调用互联网搜索、代码执行和硬件文档查阅等工具。其核心能力在于能查阅并理解复杂的技术文档,从而自主编写代码以精确控制本地液体处理机器人或远程云实验室,实现了从高层级化学研究目标到执行物理实验的自主化。在验证实验中,Coscientist 系统不仅在化合物合成和规划测试中展现了高度准确性(七种化合物中有四种化合物的合成路线被评为满分),还操控本地液体处理机器人自主完成了两种复杂的钯催化交叉偶联反应,成功合成了目标产物,展示该系统连接虚拟设计和物理实验的强大能力^[53]。为应对多步骤、多设备协同的复杂实验场景,后续研究在 Coscientist 等单智能体系统基础上构建了层级化的多智能体系统。例如,Song

等^[54]开发的 ChemAgents 系统通过引入一个顶层任务管理器来协调下属的文献阅读员、实验设计员等多个专家智能体。该系统在探索新型催化剂的研究中自主完成了对 100 种候选物的高通量合成和测试,并通过贝叶斯优化策略和预训练的机器学习模型成功筛选出一种催化剂,其性能显著优于随机合成的催化剂。

2.2 构建综合性和通用型系统

2.2.1 构建综合性药物设计平台 除了在药物设计的特定环节发挥作用,部分智能体已经发展为“端到端”的综合性药物设计平台。这类平台不再局限于单一任务的解决,而是致力于构建自主的研发闭环。它们不仅能高效模拟从靶点发现到候选药物评估的完整数字化研发流程,还能进一步覆盖湿实验验证环节,打通经典的 DMTA 循环,以最大限度地减少人工干预,加速临床药物研发进程。

一些平台致力于构建完整的“虚拟制药”流程。例如,PharmAgents 系统通过部署四个专业智能体分别负责靶点发现、先导化合物识别、先导化合物优化和临床前候选药物评估,实现了从输入疾病名称到输出临床前候选药物的全过程自动化^[55]。实验表明,该系统将药物设计成功率从 15.72% 提升至 37.94%,同时将毒性低估风险控制在 12% 以内,展现出通过经验积累实现自我演进的能力^[55]。具有类似设计理念的还有 AgentD 系统,它由中央大语言模型统一协调包括生物学数据检索、领域特定问答、分子生成及其属性预测、属性感知的分子优化和三维蛋白-配体复合物结构生成在内的独立功能模块,以支持药物发现的不同环节^[56]。在针对 Bcl-2 靶点的案例研究中,AgentD 系统首先利用检索模块获取靶点序列及已知抑制剂信息,并通过问答模块解析相关的药物协同机制;随后,系统结合生成模型构建初始分子库,利用预测和优化模块进行两轮迭代,成功将库中具有高类药性分子的数量从 34 个提升至 55 个,通过关键类药规则的分子也从 29 个增至 52 个;最后,系统进一步生成了候选分子的三维蛋白-配体复合物结构以辅助结合模式验证。这一全流程的自主执行有力证明了 AgentD 系统作为综合性平台在药物设计任务中的重要作用。

Tippy 系统是面向生产阶段的自动化综合性

药物设计平台^[13]。该系统围绕现代药物研发中经典的 DMTA 循环来构建其多智能体架构,通过五个专业智能体协作,与实验室信息管理系统和自动化设备深度集成,专注于加速 DMTA 周期。例如,Tippy 系统自主完成了从分子设计、合成规划、高效液相色谱法分析到数据解读的完整 DMTA 循环,最终成功合成了目标分子,经高效液相色谱法分析验证其纯度达到 95.3%,产率为 72%,展现了其作为生产级自动化 AI 系统高效执行完整 DMTA 循环的潜力^[13]。

2.2.2 发展通用型生物医学智能体 随着智能体在药物研发各个关键环节的深入应用,强调广覆盖性和跨场景适应性的通用型生物医学智能体开始成为全方位的“科研助手”。

实现通用性有两条主流路径。一是构建庞大而全面的工具生态。Biomni 系统^[36]通过统一平台集成了超过 150 个专业工具和近 60 个数据库,并引入了基于大语言模型的工具选择器和以代码为核心的通用行动接口,以应对复杂工作流的挑战。性能评估显示,该系统展现了强大的泛化能力,在八项具有挑战性的现实科研场景中,其性能比基础大语言模型提升了 402.3%;在数据库问答基准测试中,达到了与人类专家相当的 74.4% 准确率,并在序列问答基准测试中以 81.9% 的准确率超越了人类水平^[36]。二是引入知识图谱以优化工具组合。SciToolAgent 系统^[35]通过引入科学工具知识图谱(SciToolKG)来协同管理海量科学工具。在定量评估中,该系统在科学工具调用评估基准测试中的准确率高达 94%,在处理多工具复杂任务时比 ReAct 等基线方法高出 10% 至 20%^[35]。

STELLA 系统^[39]旨在实现智能体自我进化。其核心是两个动态进化机制:可进化的推理策略库(能从成功经验中提炼模板)和可动态扩展的工具集(能根据任务需求自主创建新工具),这种设计使其性能可随经验积累而持续提升。在多个生物医学基准测试中(如 LAB-Bench 中的文献问答基准测试),该系统达到了较高的准确率,更重要的是,其准确率能随经验积累而增加,在与人类的交互学习中实现了近乎翻倍的性能提升^[39]。

总之,智能体的应用已覆盖药物研发全周期,呈现出从单点任务突破向系统化、通用化平台演进的清晰趋势,并逐步实现了计算与实验的深度融合。代表性的智能体及其所用大语言模

型总结见附表1。

3 智能体未来发展方向

3.1 实现基础模型和智能体的协同进化

提升智能体在复杂药物研发中的成功率是其迈向产业化的核心瓶颈^[57]。当前改进智能体的方法多停留在基于上下文提示的调整^[39, 47],这难以从根本上改进大语言模型的核心决策能力。强化学习为解决这一挑战提供了最自然的范式^[58],即让智能体通过与环境的持续交互(试错),并从任务反馈的奖励信号中自主学习,从而自主优化决策策略^[59]。例如,Biomni-R0系统采用近端策略优化算法,提升了参数量较小的开源模型完成任务的成功率^[36]。未来需探索更先进的强化学习策略,如通过ARPO^[60]算法以优化长序列任务,或通过SPEAR^[61]算法以提升资源效率。将这类算法融入智能体,将使其决策模式从短期适应进化为具有原则性的自我优化,加速复杂药物研发的产业化进程。

3.2 构建自驱动实验室

在药物研发领域,智能体的另一重要发展方向是演进为具身智能体,即构建闭环DMTA循环的自驱动实验室^[62]。然而,通向真正自主、鲁棒的自驱动实验室之路仍需跨越两大根本性障碍:一是硬件生态的割裂。不同厂商的自动化设备如同“自动化孤岛”,这些设备互不兼容的专有接口阻碍了跨设备工作流的整合和复现^[62]。未来必须构建开放、标准化的实验室操作系统,例如推广PyLabRobot^[63]这类开源软件库,为实验室自动化设备提供与硬件无关的编程接口,并推动SiLA2等标准化通信协议的行业普及,实现“即插即用”^[64]。二是物理系统的不确定性。数字世界中完美的实验计划常因试剂批次差异、仪器性能不稳定等物理微扰而失效^[65]。未来的智能体必须从“开环执行者”升级为“闭环控制者”,这要求其深度集成多模态传感器进行实时状态感知,并借鉴AgentOps(一种针对智能体的观察、分析、优化和自动化的运营框架)等研究工作的理念,具备自主的异常检测、故障诊断和流程恢复能力^[66],例如,当检测到移液压力异常时,能自主推断原因并执行纠错程序。同时,须将贝叶斯优化等不确定性感知算法融入闭环决策,以应对含噪数据。例如,在面对生物活性结果接近临界值

时,能自主评估是投入资源重复验证还是转向探索新分子,从而实现更高效、鲁棒的自主探索^[67]。

3.3 构建安全、可控的药物研发智能体

智能体的强大自主性也带来了严峻的安全挑战,必须构建多层次的保障和监管框架。风险来自三个层面:使用者和开发者层面(如恶意设计生物毒素或操作失误^[68-69]);智能体自身层面(如因自身幻觉误导实验方向^[70]、设计出含毒性的分子^[71-72]、控制失误引发物理危险^[73]);以及社会和监管层面(如技术垄断、医疗不平等^[74]和事故问责机制不明^[75])。针对这些挑战,首先,必须加强对开发者(通过能力限制和审计追踪)和用户(通过身份验证和行为监控)的双重监管,辅以法律框架明确责任归属^[76]。其次,尽管已有系统(如ChemCrow^[18]、SciToolAgent^[35])尝试通过调用PubChem^[77]、UniProt^[78]等数据库来规避已知风险,但尚存在覆盖不全、无法捕捉实时风险等局限,未来应在工具层面内置多层防护,例如集成基于Reaxys^[79]知识图谱的反应安全过滤器,或引入实时毒性预测模块。同时,须通过针对性微调和红队测试提升大语言模型自身的风险识别能力^[76]。最后,积极推动技术开源共享及学术界与产业界的深度融合以缓解垄断,并建立清晰的责任归属和国际协同治理机制来明确问责^[80-81]。此外,为有效应对这些挑战,建立专门的安全评估基准至关重要。不同于现有的通用基准,未来的基准必须聚焦于评估智能体在药物研发真实场景中防范恶意问题、识别毒性分子等特有挑战的能力^[82-83]。

4 结 语

智能体作为一种新兴的智能范式,正为面临“三高一长”困境的药物研发领域带来一场深刻变革。它超越了单一的预测或生成模型,进化为能够整合知识、推理规划、调用工具来完成复杂科学任务的“主动合作者”。从自动化生成科学假设到闭环执行DMTA循环,智能体已在药物发现的多个关键环节展现出重塑研发流程的巨大潜力。

尽管前景广阔,但将智能体从学术探索全面推向产业化应用仍面临关键瓶颈。首先,智能体在长周期、高复杂性任务中的成功率和稳定性亟待通过强化学习等手段实现根本性突破。其次,

迈向真正自主的具身智能体(如自驱动实验室)必须解决硬件生态割裂和物理系统不确定性的双重难题。最后,其强大的自主性也带来了前所未有的安全性挑战,亟需构建覆盖技术、伦理和监管的多层次保障框架。

展望未来,一个由人类科学家和智能体协同共进的新时代正在开启。智能体与具身智能技术的融合,将加速自驱动实验室落地,让人类科学家的智慧从重复性劳动中解放出来,而能聚焦于顶层战略规划、关键决策和非共识的创造性研究。实现这一宏伟蓝图,需要计算机科学、化学、生物学及自动化等多领域研究者的通力协作,共同推动这一变革性技术安全、可控地走向成熟。

本文附加文件见电子版。



志谢 本研究得到国家自然科学基金(22373085)和浙江省自然科学基金(LHDMZ25H300001)支持。浙江大学药学院毛乐霄同学参与稿件修改

Acknowledgements This study was supported by National Natural Science Foundation of China (22373085) and Zhejiang Provincial Natural Science Foundation of China (LHDMZ25H300001). MAO Lexiao from the School of Pharmacy of Zhejiang University provided support in revising this manuscript

作者贡献 赵东海和谢昌谕参与论文选题和设计或参与资料获取、分析或解释,起草研究论文或修改重要智力性内容。所有作者均已阅读并认可最终稿件,并对数据的完整性和安全性负责。具体见电子版

Author Contributions ZHAO Donghai and HSIEH Changyu participated in brewing and designing experiments, or acquisition, analysis, or interpretation of data for the work; drafting the work, or revising it critically for important intellectual content. All authors have read and approved the final manuscript, and take responsibility for the integrity and security of the data. See the electronic version for details

数据可用性 本研究未生成任何新数据集,所有分析数据均已公开,并已在文中明确标引

Data Availability This study did not generate any new datasets, all data analyzed are publicly available, and have been properly cited

医学伦理 本研究不涉及人体或动物实验

Ethical Approval This study does not contain any studies with human participants or animals performed by any of the author

利益冲突 所有作者均声明不存在利益冲突

Conflict of Interests The authors declare that there is no conflict of interests

©The author(s) 2026. This is an open access article under the CC BY-NC-ND 4.0 license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

参考文献(References)

- [1] MALHEIRO V, SANTOS B, FIGUEIRAS A, et al. The potential of artificial intelligence in pharmaceutical innovation: from drug discovery to clinical trials[J]. **Pharmaceuticals (Basel)**, 2025, 18(6): 788.
- [2] LIU X, YI W, XI B, et al. Identification of drug-disease associations using a random walk with restart method and supervised learning[J]. **Comput Math Methods Med**, 2022, 2022: 7035634.
- [3] WANG Y, QI J, CHEN X. Accurate prediction of epigenetic multi-targets with graph neural network-based feature extraction[J]. **Int J Mol Sci**, 2022, 23(21): 13347.
- [4] SALAHSHOORI I, GOLRIZ M, NOBRE M A L, et al. Simulation-based approaches for drug delivery systems: navigating advancements, opportunities, and challenges [J]. **J Mol Liq**, 2024, 395: 123888.
- [5] TASSAKKA A C M A R, ISKANDAR I W, ALAM J F, et al. Docking studies and molecular dynamics simulations of potential inhibitors from the brown seaweed *Sargassum polycystum* (Phaeophyceae) against PLpro of SARS-CoV-2[J]. **BioTech (Basel)**, 2023, 12(2): 46.
- [6] CHEN L, TAN X, WANG D, et al. TransformerCPI: improving compound-protein interaction prediction by sequence-based deep learning with self-attention mechanism and label reversal experiments[J]. **Bioinformatics**, 2020, 36(16): 4406-4414.
- [7] CHEN L, FAN Z, CHANG J, et al. Sequence-based drug design as a concept in computational drug design [J]. **Nat Commun**, 2023, 14(1): 4217.
- [8] WU Z, JIANG D, WANG J, et al. Mining toxicity information from large amounts of toxicity data[J]. **J Med Chem**, 2021, 64(10): 6924-6936.
- [9] PRYKHODKO O, JOHANSSON S V, KOTSIAS P C, et al. A *de novo* molecular generation method using latent vector based generative adversarial network[J]. **J Cheminform**, 2019, 11(1): 74.
- [10] WANG J, HSIEH C Y, WANG M, et al. Multi-constraint molecular generation based on conditional transformer, knowledge distillation and reinforcement learning[J]. **Nat Mach Intell**, 2021, 3(10): 914-922.

- [11] MADANI A, KRAUSE B, GREENE E R, et al. Large language models generate functional protein sequences across diverse families[J]. **Nat Biotechnol**, 2023, 41 (8): 1099-1106.
- [12] WATSON J L, JUERGENS D, BENNETT N R, et al. De novo design of protein structure and function with RFDiffusion[J]. **Nature**, 2023, 620(7976): 1089-1100.
- [13] FEHLIS Y, CRAIN C, JENSEN A, et al. Accelerating drug discovery through agentic AI: a multi-agent approach to laboratory automation in the DMTA cycle [J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2507.09023.
- [14] KRISHNAN N. AI agents: evolution, architecture, and real-world applications[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2503.12687.
- [15] GAO S, FANG A, HUANG Y, et al. Empowering biomedical discovery with AI agents[J]. **Cell**, 2024, 187 (22): 6125-6151.
- [16] WU J, ZHU J, LIU Y, et al. Agentic reasoning: a streamlined framework for enhancing LLM reasoning with agentic tools[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2502.04644.
- [17] DORON G, GENWAY S, ROBERTS M, et al. Generative AI: driving productivity and scientific breakthroughs in pharmaceutical R&D[J]. **Drug Discov Today**, 2025, 30(1): 104272.
- [18] BRAN A M, COX S, SCHILTER O, et al. Augmenting large language models with chemistry tools[J]. **Nat Mach Intell**, 2024, 6(5): 525-535.
- [19] GOTTSWEIS J, WENG W H, DARYIN A, et al. Towards an AI co-scientist[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2502.18864.
- [20] YAO S, ZHAO J, YU D, et al. ReAct: synergizing reasoning and acting in language models[J]. **arXiv e-prints**, 2022. DOI: 10.48550/arXiv.2210.03629.
- [21] WEI J, WANG X, SCHUURMANS D, et al. Chain-of-thought prompting elicits reasoning in large language models[J]. **arXiv e-prints**, 2022. DOI: 10.48550/arXiv.2201.11903.
- [22] SHINN N, CASSANO F, BERMAN E, et al. Reflexion: language agents with verbal reinforcement learning[J]. **arXiv e-prints**, 2023. DOI: 10.48550/arXiv.2303.11366.
- [23] DEEPSEEK AI, GUO D, YANG D, et al. DeepSeek-R1: incentivizing reasoning capability in LLMs via reinforcement learning[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2501.12948.
- [24] GAO Y, XIONG Y, GAO X, et al. Retrieval-augmented generation for large language models: a survey[J]. **arXiv e-prints**, 2023. DOI: 10.48550/arXiv.2312.10997.
- [25] SHARMA C. Retrieval-augmented generation: a comprehensive survey of architectures, enhancements, and robustness frontiers[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2506.00054.
- [26] GORYANIN I, GORYANIN I, DEMIN O. Revolutionizing drug discovery: integrating artificial intelligence with quantitative systems pharmacology[J]. **Drug Discov Today**, 2025, 30(9): 104448.
- [27] EDGE D, TRINH H, CHENG N, et al. From local to global: a graph RAG approach to query-focused summarization[J]. **arXiv e-prints**, 2024. DOI: 10.48550/arXiv.2404.16130.
- [28] JIN R, GUO Y, QU Y, et al. BioLab: end-to-end autonomous life sciences research with multi-agents system integrating biological foundation models[J]. **bioRxiv**, 2025. DOI: 10.1101/2025.09.03.674085.
- [29] ASAI A, WU Z, WANG Y, et al. Self-RAG: learning to retrieve, generate, and critique through self-reflection [J]. **arXiv e-prints**, 2023. DOI: 10.48550/arXiv.2310.11511.
- [30] JEONG M, SOHN J, SUNG M, et al. Improving medical reasoning through retrieval and self-reflection with retrieval-augmented large language models[J]. **Bioinformatics**, 2024, 40(Suppl 1): i119-i129.
- [31] CHEN Q, TAO W, ZHU Z, et al. ComRAG: retrieval-augmented generation with dynamic vector stores for real-time community question answering in industry[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2506.21098.
- [32] SU W, TANG Y, AI Q, et al. DRAGIN: Dynamic retrieval augmented generation based on the information needs of large language models[J]. **arXiv e-prints**, 2024. DOI: 10.48550/arXiv.2403.10081.
- [33] XU R, PENG J. A comprehensive survey of deep research: systems, methodologies, and applications [J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2506.12594.
- [34] SINGH A, EHTESHAM A, KUMAR S, et al. Agentic retrieval-augmented generation: a survey on agentic RAG[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2501.09136.
- [35] DING K, YU J, HUANG J, et al. SciToolAgent: a knowledge-graph-driven scientific agent for multitool integration[J]. **Nat Comput Sci**, 2025, 5(10): 962-972.
- [36] HUANG K, ZHANG S, WANG H, et al. Biomni: a general-purpose biomedical AI agent[J]. **bioRxiv**, 2025. DOI: 10.1101/2025.05.30.656746.
- [37] KANG J, LAROCHE R, YUAN X, et al. Think before you act: decision transformers with working memory[J]. **arXiv e-prints**, 2023. DOI: 10.48550/arXiv.2305.16338.
- [38] PARK J S, O'BRIEN J C, CAI C J, et al. Generative agents: interactive simulacra of human behavior[J]. **arXiv e-prints**, 2023. DOI: 10.48550/arXiv.2304.03442.
- [39] JIN R, ZHANG Z, WANG M, et al. STELLA: Self-evolving LLM agent for biomedical research[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2507.02004.
- [40] MASTERMAN T, BESEN S, SAWTELL M, et al. The landscape of emerging AI agent architectures for reasoning, planning, and tool calling: a survey[J]. **arXiv e-prints**, 2024. DOI: 10.48550/arXiv.2404.11584.
- [41] PAN Q, XU D, YAO X J, et al. FROGENT: an end-to-end full-process drug design agent[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2508.10760.

- [42] REIN D, LI HOU B, COOPER STICKLAND A, et al. GPQA: a graduate-level google-proof Q&A benchmark [J]. **arXiv e-prints**, 2023. DOI: 10.48550/arXiv.2311.12022.
- [43] GUAN Y, CUI L, INCHAI J, et al. AI-assisted drug repurposing for human liver fibrosis[J/OL]. **Adv Sci (Weinh)**, 2025, 12(44): e08751.
- [44] PENADÉS J R, GOTTWEIS J, HE L, et al. AI mirrors experimental science to uncover a mechanism of gene transfer crucial to bacterial evolution[J]. **Cell**, 2025, 188(23): 6654-6665.e2.
- [45] HE L, PATKOWSKI J B, WANG J, et al. Chimeric infective particles expand species boundaries in phage-inducible chromosomal island mobilization[J]. **Cell**, 2025, 188(23): 6636-6653.e17.
- [46] ESSAM GHAREEB A, CHANG B, MITCHENER L, et al. Robin: a multi-agent system for automating scientific discovery[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2505.13400.
- [47] ZHANG Z, QIU Z, WU Y, et al. OriGene: a self-evolving virtual disease biologist automating therapeutic target discovery[J]. **bioRxiv**, 2025. DOI: 10.1101/2025.06.03.657658.
- [48] ISHIDA S, AASAWAT T, SUMITA M, et al. ChemTSv2: functional molecular design using *de novo* molecule generator[J/OL]. **WIREs Comput Mol Sci**, 2023, 13(6): e1680.
- [49] ISHIDA S, SATO T, HONMA T, et al. Large language models open new way of AI-assisted molecule design for chemists[J]. **J Cheminform**, 2025, 17(1): 36.
- [50] YU J, ZHENG Y, KOH H Y, et al. Collaborative expert LLMs guided multi-objective molecular optimization [J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2503.03503.
- [51] WEI Z, EKTEFAIE Y, ZHOU A, et al. Fleming: an AI agent for antibiotic discovery in mycobacterium tuberculosis[J]. **bioRxiv**, 2025. DOI: 10.1101/2025.04.01.646719.
- [52] SWANSON K, WU W, BULAONG N L, et al. The Virtual Lab of AI agents designs new SARS-CoV-2 nanobodies[J]. **Nature**, 2025, 646(8085): 716-723.
- [53] BOIKO D A, MACKNIGHT R, KLINE B, et al. Autonomous chemical research with large language models [J]. **Nature**, 2023, 624(7992): 570-578.
- [54] SONG T, LUO M, ZHANG X, et al. A multiagent-driven robotic AI chemist enabling autonomous chemical research on demand[J]. **J Am Chem Soc**, 2025, 147(15): 12534-12545.
- [55] GAO B, HUANG Y, LIU Y, et al. PharmAgents: building a virtual pharma with large language model agents [J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2503.22164.
- [56] OCK J, SHARMA M R, BADRINARAYANAN S, et al. Large language model agent for modular task execution in drug discovery[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2507.02925.
- [57] HARTUNG T. AI, agentic models and lab automation for scientific discovery—the beginning of scAInce[J]. **Front Artif Intell**, 2025, 8: 1649155.
- [58] HUANG C, FAN Z, WANG L, et al. Self-evolved reward learning for LLMs[J]. **arXiv e-prints**, 2024. DOI: 10.48550/arXiv.2411.00418.
- [59] WANG G, XIE Y, JIANG Y, et al. Voyager: an open-ended embodied agent with large language models[J]. **arXiv e-prints**, 2023. DOI: 10.48550/arXiv.2305.16291.
- [60] DONG G, MAO H, MA K, et al. Agentic reinforced policy optimization[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2507.19849.
- [61] QIN Y, TAN X, HE Z, et al. Learn the ropes, then trust the wins: self-imitation with progressive exploration for agentic reinforcement learning[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2509.22601.
- [62] COOPER A I, COURTNEY P, DARVISH K, et al. Accelerating discovery in natural science laboratories with AI and robotics: perspectives and challenges[J]. **Sci Robot**, 2025, 10(106): eadv7932.
- [63] WIERENGA R P, GOLAS S M, HO W, et al. PyLabRobot: an open-source, hardware-agnostic interface for liquid-handling robots and accessories[J]. **Device**, 2023, 1(4): 100111.
- [64] PORR M, LANGE F, MARQUARD D, et al. Implementing a digital infrastructure for the lab using a central laboratory server and the SiLA2 communication standard [J]. **Eng Life Sci**, 2021, 21(3-4): 208-219.
- [65] LUO Z, KASIRZADEH A, SHAH N B. The more you automate, the less you see: hidden pitfalls of AI scientist systems [J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2509.08713.
- [66] MOSHKOVICH D, ZELTYN S. Taming uncertainty via automation: observing, analyzing, and optimizing agentic AI systems[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2507.11277.
- [67] GUPTA R, HARTFORD J, LIU B. LLMs for bayesian optimization in scientific domains: are we there yet?[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2509.21403.
- [68] WANG M, ZHANG Z, BEDI A S, et al. A call for built-in biosecurity safeguards for generative AI tools[J]. **Nat Biotechnol**, 2025, 43(6): 845-847.
- [69] ALBER D A, YANG Z, ALYAKIN A, et al. Medical large language models are vulnerable to data-poisoning attacks[J]. **Nat Med**, 2025, 31(2): 618-626.
- [70] XU Z, JAIN S, KANKANHALLI M. Hallucination is inevitable: an innate limitation of large language models [J]. **arXiv e-prints**, 2024. DOI: 10.48550/arXiv.2401.11817.
- [71] HATTOH G, AYENSU J, OFORI NP, et al. Can large language models design biological weapons? evaluating moremi bio[J]. **arXiv e-prints**, 2025. DOI: 10.48550/arXiv.2505.17154.
- [72] TAN B, ZHANG J, XIAO C, et al. Progress of artificial intelligence in drug synthesis and prospect of its

- application in nitrification of energetic materials[J]. **Molecules**, 2023, 28(4): 1900.
- [73] YU S, RAN N, LIU J. Large-language models: the game-changers for materials science research[J]. **Artif Intell Chem**, 2024, 2(2): 100076.
- [74] MIRAKHORI F, NIAZI S K. Harnessing the AI/ML in drug and biological products discovery and development: the regulatory perspective[J]. **Pharmaceuticals (Basel)**, 2025, 18(1): 47.
- [75] FOTHERINGHAM K, SMITH H. Accidental injustice: healthcare AI legal responsibility must be prospectively planned prior to its adoption[J]. **Future Healthc J**, 2024, 11(3): 100181.
- [76] TANG X, JIN Q, ZHU K, et al. Risks of AI scientists: prioritizing safeguarding over autonomy[J]. **Nat Commun**, 2025, 16(1): 8317.
- [77] KIM S, CHEN J, CHENG T, et al. PubChem in 2021: new data content and improved web interfaces[J]. **Nucleic Acids Res**, 2021, 49(D1): D1388-D1395.
- [78] CONSORTIUM U. UniProt: the universal protein knowledgebase in 2023[J]. **Nucleic Acids Res**, 2023, 51(D1): D523-D531.
- [79] ZHANG C, ARUN A, LAPKIN A A. Completing and balancing database excerpted chemical reactions with a hybrid mechanistic-machine learning approach [J]. **ACS Omega**, 2024, 9(16): 18385-18399.
- [80] BERDAHL C T, BAKER L, MANN S, et al. Strategies to improve the impact of artificial intelligence on health equity: scoping review[J/OL]. **JMIR AI**, 2023, 2: e42936.
- [81] RATTI E, MORRISON M, JAKAB I. Ethical and social considerations of applying artificial intelligence in healthcare-a two-pronged scoping review[J]. **BMC Med Ethics**, 2025, 26(1): 68.
- [82] ZHANG Z, CUI S, LU Y, et al. Agent-SafetyBench: evaluating the safety of LLM agents[J]. **arXiv e-prints**, 2024. DOI: 10.48550/arXiv.2412.14470.
- [83] ZHANG H, HUANG J, MEI K, et al. Agent Security Bench (ASB): formalizing and benchmarking attacks and defenses in LLM-based agents[J]. **arXiv e-prints**, 2024. DOI: 10.48550/arXiv.2410.02644.

[本文编辑 沈 敏 沈 洁]

· 学术动态 ·

王青青和来利华教授团队揭示微蛋白 FXYD3 调控黏液层形成维护肠道屏障

2025 年 11 月 3 日,浙江大学王青青/来利华教授团队在《细胞报告》(*Cell Reports*)发表“Goblet cell-expressed microprotein FXYD3 determines gut homeostasis by maintaining mucus barrier integrity”论文(DOI:10.1016/j.celrep.2025.116502),揭示杯状细胞表达的微蛋白 FXYD3 通过与肌浆网钙 ATP 酶 2(SERCA2)互作调控细胞内钙稳态,参与黏液层形成和维持,从而保障肠道稳态。这一发现为深入理解肠道黏膜防御机制提供了新的视角。

FXYD3 是一种已知可调节钠钾 ATP 酶活性的微蛋白,研究人员既往揭示了 FXYD3 在乳腺癌和皮肤炎症中发挥调控功能。该研究发现 FXYD3 高表达于结肠杯状细胞,肠上皮特异性敲除 FXYD3 的小鼠出现结肠黏液层缺损、菌群失调及细菌入侵,并且对葡聚糖硫酸钠诱导的结肠炎及病原菌感染更为易感。FXYD3 通过与 SERCA2 相互作用,维持 SERCA2 活性,从而调控杯状细胞内质网钙离子稳态,这一过程促进了黏蛋白的分泌及其关键糖基化修饰唾液酸化的完成,最终保障黏液层的结构与功能完整性。临床研究显示,溃疡性结肠炎患者肠道 FXYD3 表达显著下调,其表达水平与杯状细胞相关转录因子 SPDEF 和 ATOH1 呈正相关,有趣的是肠道菌群代谢产物丙酸盐和丁酸盐可诱导并维持 FXYD3 的表达,提示了由菌群-宿主互作介导的屏障调控新途径。

该成果阐明了 FXYD3 通过钙稳态调控黏液屏障功能的新机制,为炎症性肠病等黏膜疾病的诊治提供了潜在靶点,也为理解菌群与宿主之间的双向调控提供了理论依据。

杨文娟、薛越、祝培朔博士研究生及浙江大学医学院附属邵逸夫医院病理科姜支农主任医师为论文第一作者,研究得到国家重点研发计划和自然科学基金等资助。