

特色专题

人工智能驱动药物研发进展

何欣恒, 高斯涵, 李俊睿, 徐华强*

摘要 药物研发作为现代医药产业的核心驱动力, 面临传统模式“高投入、长周期、低产出”的困境, 亟需突破以应对日益复杂的健康需求。人工智能(AI)技术的快速发展为药物研发带来了革命性变革, 其在蛋白质结构预测、蛋白质设计、抗体药物设计及小分子药物设计等领域的应用显著提升了研发效率与成功率。深入分析了 AI 在蛋白质结构预测中的突破及其在靶点发现、虚拟筛选等环节的应用潜力; 探讨了 AI 驱动蛋白质设计从结构预测到功能创新的闭环模式; 剖析了 AI 在抗体序列优化、亲和力成熟及新型抗体设计中的作用; 梳理了 AI 在小分子药物靶点识别、虚拟筛选及 ADMET 优化中的最新成果。指出 AI 应用中面临的数据质量、模型可解释性及实验验证等挑战, 并展望了多模态数据融合、动态行为预测及自动化平台的未来发展方向。通过全面剖析 AI 赋能药物研发的现状与问题, 旨在为加速新药创制、提升人类健康福祉提供科学视角与思考启示, 提供一个关于 AI 赋能药物研发领域科技问题的全面且深入的视角, 并激发对未来发展方向的思考, 以期促进 AI 技术在药物研发领域的更有效应用, 加速新药创制进程, 最终惠及人类健康。

关键词 人工智能; 药物研发; AlphaFold3; 蛋白质设计; 抗体设计; 小分子药物设计

药物研发是支撑现代医药产业发展的核心驱动力, 是应对人类疾病挑战的关键环节。然而, 传统的药物研发模式长期以来深陷“三十”困境——平均超过 10 年的研发周期和超过 10 亿美元的研发投入, 却伴随着不到 10% 的药物上市成功率^[1]。这种高投入、长周期、低产出的模式, 严重制约了新药的产出^[2]。因此, 如何突破传统药物研发的瓶颈, 加速新药创制, 已成为全球医药领域共同面临的重大挑战^[3]。

近年来, 人工智能(artificial intelligence, AI)技术的发展为破解上述困境带来了曙光。AI 凭借其卓越的数据处理能力、模式识别能力以及泛化预测能力, 在药物研发的各个环节都展现出了巨大的应用潜

力^[4]。正在改变着药物研发的范式^[5]。

本文深入分析了 AI 驱动药物研发关键领域的国内外发展现状、最新研究成果、现存问题和未来发展方向, 力求提供关于 AI 赋能药物研发领域科技问题的全面视角, 以期促进 AI 技术在药物研发领域的更有效应用(图 1)。

1 AI 驱动蛋白质结构预测: AlphaFold3 的突破与应用

蛋白质是生命活动的功能执行者。精确的蛋白质结构是解析疾病机制、设计靶向药物的基础。传统的蛋白质结构解析方法虽然可以获得高分辨率的蛋白质结构, 但存在耗时耗力、通量低等局限性。因此,

中国科学院上海药物研究所, 上海 201203

收稿日期: 2025-04-23; 修回日期: 2025-06-03

基金项目: 国家重点研发计划项目(2022YFC2703105); 国家自然科学基金项目(32130022, 82121005); 中国科学院战略性先导科技计划项目(XDB37030103); 上海市科技重大专项(2019SHZDZX02)

作者简介: 何欣恒, 博士研究生, 研究方向为计算生物学, 电子信箱: he-xinheng@foxmail.com; 徐华强(通信作者), 研究员, 研究方向为结构生物学与药物发展, 电子信箱: eric.xu@simm.ac.cn

引用格式: 何欣恒, 高斯涵, 李俊睿, 等. 人工智能驱动药物研发进展[J]. 科技导报, 2025, 43(12): 29-37; doi:10.3981/j.issn.1000-7857.2025.04.00082

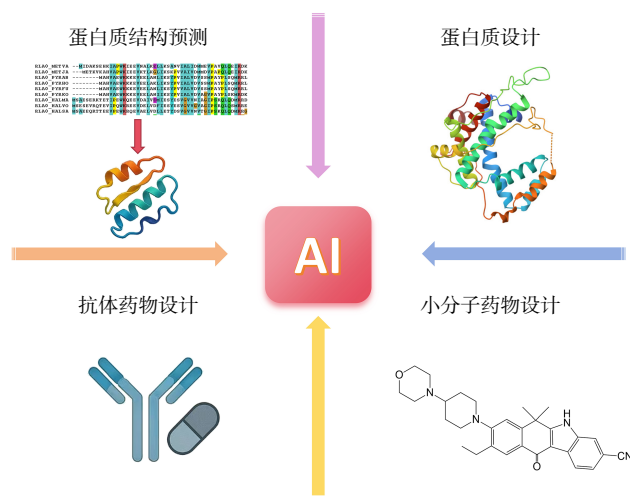


图1 AI技术应用于药物研发的重点领域

长期以来,研发人员迫切需求准确高效的蛋白质结构预测方法。AlphaFold3 等 AI 模型的出现,提高了蛋白质结构预测的精度和速度,被认为是加速药物发现的关键技术^[6]。

1.1 AI 在蛋白质结构预测中的应用

DeepMind 公司开发的 AlphaFold 系列模型是 AI 驱动蛋白质结构预测领域的里程碑式突破。AlphaFold2 的预测精度可以与实验方法相媲美。2024 年,DeepMind 发布了 AlphaFold3,进一步拓展了结构预测的能力^[7]。

AlphaFold3 不仅能够预测蛋白质的三维结构,还能预测包括脱氧核糖核酸(DNA)、核糖核酸(RNA)、配体以及修饰在内的生物分子复合物的结构。其采用了更先进的 Pairformer 注意力机制,能够基于成对表示高效地捕捉蛋白质序列中的长程依赖关系,并利用更全面的生物学数据库进行训练^[8]。

AlphaFold3 在多个评估指标上超越 AlphaFold2,尤其是在预测蛋白质-配体等复合物结构方面取得了 0 到 1 的突破。这为研究复杂的生物分子相互作用,及开发靶向特定复合物的药物提供了强有力的工具^[7]。

除了 AlphaFold 系列模型,蛋白质结构预测领域还涌现出多种高性能模型,包括 RoseTTAFold^[9]和 Protenix 等。RoseTTAFold 是 David Baker 实验室推出的深度学习蛋白质结构预测框架,核心创新在于同时更新序列、配对信息及三维坐标的“三轨”并行注意力网络。Protenix 是由字节跳动公司开发的 AlphaFold3

开源实现版本,旨在提高模型的可训练性和可扩展性。该模型在预测蛋白质-配体和蛋白质-核酸相互作用方面表现出色,且开源了训练代码,便于复现。

1.2 结构预测工具在药物研发中的最新研究成果

结构预测工具的突破性进展为药物研发带来了前所未有的机遇。本文以 AlphaFold3 为例,分析 AI 在药物研发中的最新研究成果。

首先,在靶点发现与验证方面,AlphaFold3 能够帮助研究人员快速获取与疾病相关的蛋白质结构模型。有了这些预测结构,研究人员可以初步理解蛋白质的功能机制,识别潜在的药物结合位点,从而加速靶点发现和验证的过程,然而这一过程通常还需要结构生物学的方法进一步验证^[10]。

在虚拟筛选与先导化合物发现方面,AlphaFold3 同样展现了巨大价值。传统的虚拟筛选通常依赖已知的蛋白质结构,现在,研究人员可以利用 AlphaFold3 预测的结构来进行筛选。通过结合分子对接等方法,可以筛选出潜在的先导化合物,提升发现效率。

抗体药物设计是另一个受益匪浅的领域。AlphaFold3 能预测抗体-抗原复合物结构,让研究人员得以深入理解抗体与抗原的结合机制,并据此初步指导抗体的序列优化、亲和力成熟以及人源化改造。

此外,AlphaFold3 在酶工程与蛋白质药物设计中也可以应用。它不仅能预测天然蛋白质的结构,还能用于分析工程化改造后的蛋白质结构。这为研究人员提供了强有力的工具,来指导酶的定向进化或新型蛋白质药物的设计^[11]。

1.3 AlphaFold3 的局限性与挑战

虽然 AlphaFold3 在许多情况下表现优异,但其准确性仍受限于训练数据的质量和数量。对于结构复杂的蛋白质,其预测精度可能无法保证^[12]。此外,AlphaFold3 目前在预测蛋白质的动态行为和构象变化方面能力有限^[8]。

计算资源需求高是另一个显著挑战。AlphaFold3 的训练和预测过程需要消耗大量算力,这种高门槛可能阻碍其在不同领域的普及。

结构预测结果的验证也很重要。AlphaFold3 给出的结构只是理论预测,依然需要通过实验方法加以确认。作为药物研发的第一步,预测结果只是起点。

2 AI 驱动蛋白质设计:从结构预测到功能创新

在蛋白质结构预测的基础上, AI 技术进一步推动了蛋白质设计的快速发展。蛋白质设计旨在通过计算方法创造具有特定功能的非天然蛋白质。传统蛋白质设计依赖于定向进化或理性设计,效率较低且成功率有限^[13]。而 AI 技术通过深度学习和生成模型,

能够快速生成具有目标功能的蛋白质序列。在 AI 技术推动下,蛋白质设计正从传统设计和定向进化迈向高效、智能化的新阶段^[10]。

2.1 AI 在蛋白质设计中的应用

AI 在蛋白质设计中的核心优势在于能够结合序列生成、结构预测和功能优化,显著提升效率和成功率,图 2 描述 AI 在蛋白质设计中的常见应用场景。

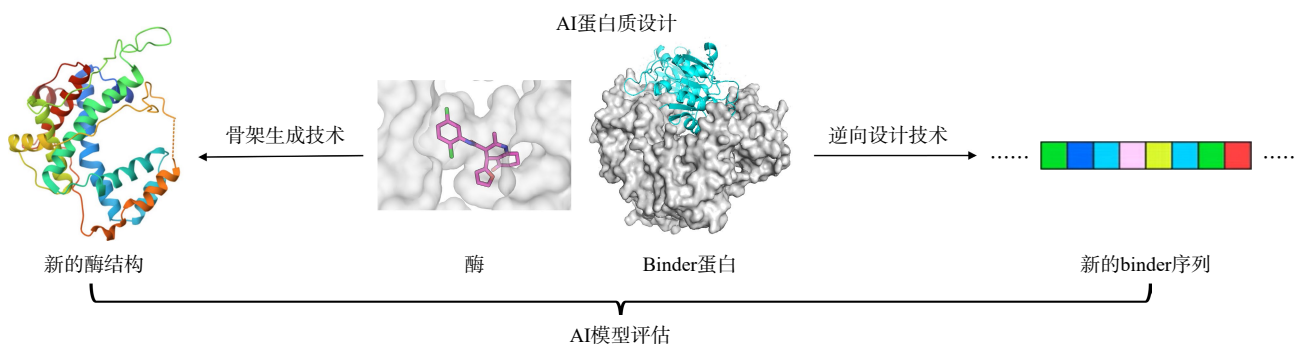


图 2 AI 在蛋白质设计中的应用场景

首先,在序列生成与优化方面, AI 通过逆向设计技术,可以从目标结构出发生成相应序列,使用生成模型也可以快速设计具有目标功能的蛋白质序列^[14]。

其次,在结构导向的设计方面, AI 利用蛋白质结构预测的突破进一步推动了设计能力。例如,基于深度学习模型 ProteinMPNN 能够学习蛋白质序列与功能之间的映射关系,生成具有特定催化活性或结合能力的序列。这些序列随后可通过结构预测工具进行折叠验证,确保其三维结构与预期功能相匹配^[15]。

在功能预测与验证方面, AI 模型能够评估蛋白质的潜在功能,或优化已有蛋白质的功能。例如, AI 可以预测酶的催化效率、蛋白质的结合亲和力或热稳定性,从而筛选出最优候选分子。这种能力在酶工程和治疗性蛋白质开发中尤为重要^[16]。

2.2 AI 驱动蛋白质设计的最新研究成果

近年来, AI 在蛋白质设计领域取得了显著进展。David Baker 团队开发的 RFdiffusion 模型已成功设计出多种新型蛋白质,包括高效催化酶和具有特定结合能力的蛋白质分子^[17]。

在工业应用方面, AI 驱动的蛋白质设计平台正逐步成熟。例如, Ginkgo Bioworks 等公司开发了基于 AI 的酶设计工具 Ginkgo Enzyme Services,通过优化

结构显著提升了工业酶的催化效率。

2.3 AI 驱动蛋白质设计面临的挑战

尽管 AI 在蛋白质设计领域展现出巨大潜力,但其发展仍面临若干关键挑战。

首先,数据依赖性为主要瓶颈。AI 模型的性能高度依赖于高质量的训练数据。然而,当前的实验数据在多样性上依然不足。训练数据往往局限于天然蛋白质的已知结构和功能。数据稀缺可能导致模型在面对全新设计任务时的泛化能力不足,限制了其在创新性蛋白质设计中的表现。

其次,功能预测的准确性仍需提升。虽然 AI 工具如 AlphaFold3 在蛋白质结构预测上取得了突破,但在预测蛋白质功能方面仍存在局限。蛋白质功能的实现往往涉及复杂的动态过程和环境依赖性,而当前 AI 模型主要基于静态结构预测,难以准确捕捉这些动态特性。这种预测误差可能导致候选分子在实验验证中失败,增加研发成本。

此外,实验成本的高昂是另一大制约因素。AI 驱动的蛋白质设计通常需要生成数千个以上候选序列,设计的候选分子需要通过湿实验验证,而这些实验步骤不仅耗时耗力,还涉及昂贵的试剂和设备成本。这使 AI 驱动蛋白质设计的广泛应用受到一定限制。

3 AI 赋能抗体药物设计:加速研发进程,提升药物疗效

抗体药物在生物制药领域中占据重要地位。其凭借靶向性强、特异性高等优势,在肿瘤、自身免疫疾病和感染性疾病等领域的治疗中展现了显著效果^[18]。但传统的抗体药物研发依赖于杂交瘤技术、噬菌体展示技术等,面临着周期长、成本高、成功率低的难题。AI 技术在抗体序列优化、人源化改造、亲和力成熟和新型抗体设计等方面的应用,有望加速抗体药物的研发进程,提高抗体药物的疗效和安全性。

3.1 AI 在抗体药物设计中的应用

AI 技术在抗体药物设计的各个环节都展现出强大潜力。首先,在抗体序列优化方面,AI 能够提升抗体疗效并降低免疫原性。例如,深度学习模型能够分析抗体互补决定区(complementarity-determining region, CDR)序列与抗原表位的相互作用,优化 CDR 设计以增强抗体结合能力。

抗体亲和力成熟是另一个 AI 发挥作用的领域。传统实验方法效率较低,且难以穷尽所有突变组合。而 AI 技术可以通过机器学习模型,根据抗体序列和亲和力数据预测最佳突变位点和组合,指导定向进化实验,从而提升亲和力成熟的效率和成功率^[19]。

在新型抗体设计方面,生成对抗网络通过生成器和判别器的对抗训练学习真实数据分布,生成器可从随机噪声中采样并输出新的抗体序列,从而实现抗体

设计的多样性。该技术可以生成全新的抗体序列,并结合预测模型筛选出最具潜力的候选分子^[20]。

近年来兴起的扩散模型(diffusion models)通过逐步去噪生成蛋白质三维结构,已被用于精确设计抗体骨架和优化 CDR 区结构,从而提高抗体与抗原结合的空间匹配度。这类模型不仅能生成符合物理化学约束的新结构,还能在结构空间中探索新颖结构,从而拓展抗体设计空间。

图神经网络(graph neural networks, GNN)能够将抗体和抗原表示为图,通过学习残基之间的空间与化学相互作用,实现亲和力预测与界面识别。例如, GearBind 是一种可预训练的几何图神经网络^[21],专为抗体亲和力成熟设计,显著提升了对抗体-抗原相互作用的建模能力。

语言模型(language models)也已成为抗体序列设计的重要工具。例如, Hie 等^[22]使用蛋白语言模型,无需结构或抗原信息即可高效指导抗体亲和力成熟,在仅 2 轮实验中将抗体亲和力提升最高达 160 倍。

抗体可开发性预测也是 AI 的重要应用方向。抗体药物的表达量、稳定性、聚集倾向等特性直接影响研发成功率。AI 模型通过学习序列与可开发性间的关系,能在早期阶段预测这些特性,排除问题分子。

为便于系统比较不同 AI 模型在抗体药物设计中的应用特征、技术机制与代表性成果,表 1 总结了扩散模型、图神经网络、生成模型与语言模型的核心原理、主要应用环节、优势。

表 1 AI 模型在抗体药物设计中的应用对比

模型类型	技术原理简述	主要应用环节	优势
扩散模型	通过逐步添加/去除噪声学习分布,实现蛋白三维结构的高精度生成	抗体骨架设计、CDR 结构建模	可精确建模抗原表位几何约束,探索新结构构象
图神经网络	将分子结构建模为图,学习节点(残基)和边(相互作用)之间的拓扑关系	抗体-抗原互作预测、亲和力预测	能处理抗体-抗原结构复杂性,支持界面残基识别与突变筛选
生成对抗网络	编码抗体序列到潜在空间或训练对抗网络,生成结构合理的新抗体序列	抗体序列生成、亲和力多样性探索	可合成高质量且多样的抗体序列,辅助构建抗体库
语言模型	使用蛋白序列作为输入,基于自注意力机制学习序列和功能信息	序列优化、指导亲和力成熟	适用于大规模数据、快速评估突变影响

3.2 AI 驱动抗体药物设计的最新研究成果

近年来, AI 在抗体药物设计领域的进展令人瞩目。例如 Absci、Adimab 和 Ginkgo Bioworks 等公司开发的平台,利用深度学习优化抗体,并在实验中得到验证,极大提升了开发效率。

到验证,极大提升了开发效率。

其中, Mason 等^[23]开发的一种端到端序列优化方法,通过深度神经网络模型在 10⁸ 规模的虚拟抗体库中筛选抗原特异性抗体,随机抽取的 30 个突变体中

全部表现出对 HER2 的有效结合,显示出 AI 在高通量抗体筛选中的强大能力。

在抗体亲和力成熟方面,研究人员开发了多种 AI 驱动技术。基于扩散模型或语言模型的方法能够高效探索突变空间,显著提高亲和力成熟的成功率。这些技术为传统实验方法提供了强有力的补充^[24]。

AI 辅助的新型抗体发现平台也层出不穷。基于生成模型或图神经网络的设计平台能够生成具有新颖结构和功能的抗体分子,为抗体药物研发注入更多可能性。如 David Baker 团队推出的基于结构的抗体生成工具 RFantibody,该方法使用抗体数据微调后的蛋白生成模型进行从头抗体结构、序列的设计和筛选^[25]。

在抗体可开发性预测方面,AI 工具同样取得了突破。Liang 等^[26]开发的 IsAb2.0 平台结合深度学习与多目标序列优化算法,成功实现了对纳米抗体 J3 的人源化重设计。通过引入少量优化突变,获得了热稳定性更高、表达水平更优且保留功能活性的 HuJ3 抗体变体,证明了 AI 在人源化路径规划中的实用性。

3.3 AI 驱动抗体药物设计面临的挑战

尽管 AI 在抗体药物设计中表现亮眼,但仍面临一些难题。数据质量和数据量是首要挑战。AI 模型依赖大量高质量数据,虽然抗体序列的数据量较大,但其亲和力数据相对稀缺,且质量不一。未来需要建立更大规模、更优质的抗体数据库,同时开发数据增强和挖掘技术。

模型可解释性也是一个关键问题。在抗体药物设计中,理解预测和生成背后的生物学机制至关重要,开发更具可解释性的 AI 抗体药物设计模型,并结合生物学知识进行解释,是亟待解决的方向^[27]。

实验验证的瓶颈同样不可忽视。AI 的预测结果最终需通过实验确认,而抗体表达、纯化、亲和力测定等步骤耗时耗力。开发高效、高通量的实验验证方法,并将 AI 与自动化实验平台整合,将是提升抗体设计效率的关键^[28]。

4 AI 助力小分子药物设计:加速发现、优化性质、提高效率

小分子药物是药物研发领域的核心力量,占据了药物市场的绝大部分份额。传统的小分子药物研发

的方式效率较低,研发周期长。AI 技术在小分子药物的靶点识别、虚拟筛选、先导化合物优化,以及药物的吸收(absorption)、分布(distribution)、代谢(metabolism)、排泄(excretion)和毒性(toxicity)——即 ADMET 性质预测等方面,正发挥着日益重要的作用。AI 的应用有望提升研发效率,并降低研发成本^[4]。

4.1 AI 在小分子药物设计中的应用

AI 在小分子药物设计的每个环节都展现出一定潜力。在靶点识别与验证方面,AI 能够分析海量生物学数据,包括基因组、转录组、蛋白质组和代谢组数据,找出已知产生药效的小分子的潜在靶点,并评估其可药性^[29]。

虚拟筛选与先导化合物发现是 AI 的另一大用武之地。传统虚拟筛选方法,如分子对接和药效团筛选效率低且假阳性率高。AI 模型通过学习化合物的结构和活性数据,可以构建更精准的预测模型,加速先导化合物发现^[30];生成模型则能基于口袋设计全新化合物,拓宽候选分子来源^[31]。

先导化合物优化同样受益于 AI。先导化合物往往存在活性不足、选择性差或 ADMET 性质不佳等问题。AI 模型能够预测这些特性,指导结构修饰,提升成药性^[32]。

小分子药物的合成路线设计也因 AI 而更高效。AI 可以预测反应产率和选择性,优化路线。例如,基于强化学习的模型可以自动探索合成路径,找到成本最低、效率最高的方案^[33]。

4.2 AI 驱动小分子药物设计的最新研究成果

近年,AI 在小分子药物设计领域成果斐然。基于深度学习虚拟筛选平台,如 Atomwise 和 Exscientia 可以预测化合物活性,已在多个项目中加速了先导化合物发现。例如,Exscientia 与 Sumitomo Dainippon Pharma 公司合作开发的 DSP-1181 用于强迫症治疗,已进入临床试验,仅用 12 个月,远快于传统的 5 年时间。

AI 驱动的先导化合物优化技术也不断突破。基于生成模型和强化学习的方法能够高效探索化合物结构空间,显著提升成功率。例如,Insilico Medicine 公司的 GENTRL 模型利用生成性张量强化学习在 21 d 内设计出有效的 DDR1 激酶抑制剂,总耗时仅 46 d,远超传统方法^[34]。

AI 驱动的药物性质预测工具同样快速发展。基于机器学习的 ADMET 预测模型和深度学习的毒性预测模型,帮助研究人员在早期评估药物潜力,提高成药性。

4.3 AI 驱动小分子药物设计面临的挑战

尽管进展显著, AI 在小分子药物设计中仍面临挑战。数据质量和偏差是首要问题。小分子药物的活性与性质数据常存在分布不均、标注错误等问题,影响模型性能。未来需建立更高质量的数据库,并改进数据清洗与增强方法^[35]。

模型的泛化能力和鲁棒性也有待提升。AI 模型在训练数据上表现良好,但在新数据集上可能失灵,数据集的偏差可能导致模型高性能表现实际上反映了数据偏差,而非物理原理的成功泛化。针对不同靶点和疾病,需开发更具通用性和稳定性的模型。

模型可解释性不足限制了应用。AI 预测结果常

难以解释,而在药物设计中,理解化学和生物学机制至关重要。开发可解释性更强的模型并提供决策支持工具将是关键。Jiménez-Luna 等^[36]指出,深度学习模型的“黑箱”性质限制了其在药物发现中的应用,亟需开发可解释性更强的模型并提供决策支持工具。

化学空间探索的局限性也不容忽视。AI 基于已知数据生成化合物,可能局限于现有化学空间。未来需结合化学合成和高通量筛选,尤其在竞争激烈的领域如激酶抑制剂发现中探索更广阔的未知领域^[37]。

5 AI 驱动药物研发的未来展望

未来, AI 将在药物研发中扮演更加核心的角色,改变传统研发范式,加速新药发现,降低成本,提升成功率,最终造福人类健康,表 2 总结了目前 AI 药物研发的挑战和进展。

表 2 AI 驱动药物研发进展与挑战总结

领域	关键进展	主要应用场景	现存挑战	未来发展方向
蛋白质结构预测	AlphaFold3实现生物分子复合物结构预测 预测精度接近实验方法 突破性扩展至配体,核酸相互作用预测	靶点发现与验证 虚拟筛选与分子对接 抗体-抗原复合物解析 酶工程与药物重定位	动态行为预测能力不足 计算资源需求高 实验验证成本高	动态构象与相互作用预测 算法优化降低算力需求 多模态数据整合与平台开发
蛋白质设计	RFdiffusion实现功能导向的从头设计 生成模型辅助结构改造 工业酶催化效率提升	新型治疗性蛋白质开发 酶从头设计 蛋白稳定性优化 合成生物学元件设计	非天然蛋白数据稀缺 功能预测准确性不足 实验验证成本高昂	多模态数据融合 动态行为预测模型开发 自动化设计-验证闭环平台
抗体药物设计	RFantibody实现结构导向的抗体生成 AI亲和力优化 AI辅助可开发性预测	抗体人源化优化 亲和力成熟加速 双特异性抗体设计 稳定性与聚集倾向预测	免疫原性预测模型不足 动态表位预测困难 实验通量限制验证效率	多组学数据整合 可解释AI指导工程化改造 自动化高通量验证平台开发
小分子药物设计	AI虚拟筛选与分子生成工具 AI化合物优化平台 ADMET预测模型	靶点识别与验证 全新分子骨架生成 合成路线智能规划 毒性预测	化学空间探索局限性 活性-选择性平衡困难 多靶点协同作用预测不足	知识图谱增强型生成模型 动态药效团建技术 自动化合成-测试闭环系统

5.1 蛋白质结构预测工具的未来发展方向

蛋白质结构预测工具有以下方面可以进一步推进。首先是提高预测精度,可以结合生物学领域内的知识与大语言模型,进一步提升其对复杂蛋白质和生物分子复合物的预测能力^[8]。

降低计算资源需求并提升计算效率也是重要方向。通过优化算法和模型,未来的 AlphaFold3 有望在

更普通的计算设备上运行,从而惠及更多研究者。

开发用户友好的工具和平台是另一个趋势。可以推出在线的结构预测平台、可视化工具以及与其他药物研发软件的集成接口,让非专业用户也能轻松上手,推动技术普及。

5.2 AI 驱动蛋白质设计的未来发展方向

AI 驱动蛋白质设计将在以下关键发展方向产生

突破,进一步赋能药物研发和生物技术创新。

首先,多模态数据整合是提升设计能力的重要途径。未来的AI模型需要超越单一的序列或结构数据,整合动态变化和细胞层面的多维度信息,以更全面地理解蛋白质的行为。

其次,动态行为预测将成为研究的重点方向。蛋白质的功能往往依赖于其在细胞环境中的动态构象变化。未来的AI工具可能通过结合分子动力学模拟或构象生成模型,预测蛋白质的实时相互作用。

此外,自动化设计平台的开发将显著推动技术普及应用。通过构建集成的全流程平台,AI可形成“设计—表达—测试—优化”的闭环系统,高效完成任务。

最后,与抗体药物设计的协同创新是未来发展的重要趋势。例如,AI可以通过设计新型蛋白框架,为抗体药物提供更多样化的结构选择,蛋白质设计还可用于优化抗体—抗原相互作用界面,通过调整局部结构增强结合亲和力。

5.3 AI驱动抗体药物设计的未来发展方向

展望未来,AI驱动抗体药物设计有以下发展方向值得关注。可解释AI与因果推断是一个重要方向。未来的研究将致力于开发更透明的模型,并结合因果推断方法,揭示抗体序列与功能之间的深层关系,为设计提供更可靠的理论支持^[27]。AI驱动的自动化抗体研发平台也将成为趋势。通过实现设计、合成、测试和优化的全流程自动化,研发速度将大幅提升。这种智能化平台的构建将改变抗体药物的开发模式^[19]。

5.4 AI驱动小分子药物设计的未来发展方向

未来,AI在小分子药物设计中的发展将更加多元化。通过整合化合物结构、性质、活性及靶点疾病信息,研究者将构建更全面的网络药理学模型,从而提升设计精度^[38]。

发展基于知识的AI模型也是可行的方向。将化学和生物学知识融入模型,如规则系统或知识图谱,能提高可解释性和决策支持能力^[39]。

AI驱动的自动化药物合成平台将进一步加速研发。通过集成路线设计、反应优化和自动化合成,整个流程将实现智能化,大幅缩短周期。例如,IBM公司的RXN for Chemistry平台已经在利用AI预测化学反应和合成路线,而Chemputer项目已经开发了通用化学机器人来自动化合成过程^[40]。

个性化小分子药物设计是未来的亮点。结合基因组学、蛋白质组学和代谢组学数据,AI也有可能为不同患者定制药物,提供更精准的治疗方案^[41]。

5.5 AI时代下的药物研发展望

在AI辅助中,药物研发将更依赖多学科交叉融合,AI技术将与生物学、化学、药学、医学和工程学深度结合,构建智能化研发平台。平台不仅能整合基因组、临床、影像等多源数据,通过AI进行分析、建模和优化,有望结合自动化实验技术实现高通量验证,推动从靶点发现到药物优化的全流程高效化。

随着个性化医疗技术的进步,AI将助力精准药物研发,通过分析患者个体化的基因组和临床数据,预测疾病风险和药物反应,设计个性化治疗方案,实现更精准的治疗效果(图3)。此外,自动化和智能化将成为药物研发的主流趋势,高通量筛选、机器人化学合成与AI设计平台的协同工作,将使研发全流程实现自动化,大幅缩短研发周期,让药物研发最终进入一个全新阶段。

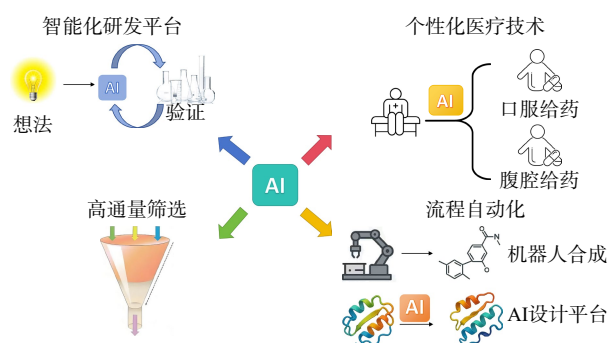


图3 AI在未来药物设计中可能扮演的角色

参考文献(References)

- [1] Hughes J, Rees S, Kalindjian S, et al. Principles of early drug discovery[J]. British Journal of Pharmacology, 2011, 162(6): 1239–1249.
- [2] Rake B, Gulbrandsen M, McKelvey M, et al. Rethinking medical innovation: Organizing R&D, responding to crisis, delivering health services[J]. Innovation, 2025, 27(1): 1–20.
- [3] Lu M K, Yin J Y, Zhu Q, et al. Artificial intelligence in pharmaceutical sciences[J]. Engineering, 2023, 27: 37–69.
- [4] Rehman A U, Li M Y, Wu B J, et al. Role of artificial intelligence in revolutionizing drug discovery[J]. Fundamental Research, 2025, 5(3): 1273–1287.
- [5] Fu C, Chen Q C. The future of pharmaceuticals: Artificial

- intelligence in drug discovery and development[J]. *Journal of Pharmaceutical Analysis*, 2025, 1(7): 101248.
- [6] Roy R, Al-Hashimi H M. AlphaFold3 takes a step toward decoding molecular behavior and biological computation[J]. *Nature Structural & Molecular Biology*, 2024, 31(7): 997–1000.
- [7] Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold3[J]. *Nature*, 2024, 630(8016): 493–500.
- [8] He X H, Li J R, Shen S Y, et al. AlphaFold3 versus experimental structures: Assessment of the accuracy in ligand-bound G protein-coupled receptors[J]. *Acta Pharmacologica Sinica*, 2025, 46(4): 1111–1122.
- [9] Baek M, DiMaio F, Anishchenko I, et al. Accurate prediction of protein structures and interactions using a three-track neural network[J]. *Science*, 2021, 373(6557): 871–876.
- [10] Desai D, Kantliwala S V, Vybhavi J, et al. Review of AlphaFold3: Transformative advances in drug design and therapeutics[J]. *Cureus*, 2024, 16(7): e63646.
- [11] Chen E, Pan E, Zhang S G. Structure bioinformatics of six human integral transmembrane enzymes and their AlphaFold3 predicted water-soluble QTY analogs: Insights into FACE1 and STEA4 binding mechanisms[J]. *Pharmaceutical Research*, 2025, 42(2): 291–305.
- [12] Wee J, Wei G W. Evaluation of AlphaFold3's protein-protein complexes for predicting binding free energy changes upon mutation[J]. *Journal of Chemical Information and Modeling*, 2024, 64(16): 6676–6683.
- [13] Cobb R E, Chao R, Zhao H M. Directed evolution: Past, present, and future[J]. *AIChE Journal*, 2013, 59(5): 1432–1440.
- [14] Strokach A, Kim P M. Deep generative modeling for protein design[J]. *Current Opinion in Structural Biology*, 2022, 72: 226–236.
- [15] Dauparas J, Anishchenko I, Bennett N, et al. Robust deep learning-based protein sequence design using Protein-MPNN[J]. *Science*, 2022, 378(6615): 49–56.
- [16] Yu T H, Cui H Y, Li J C, et al. Enzyme function prediction using contrastive learning[J]. *Science*, 2023, 379(6639): 1358–1363.
- [17] 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.
- [18] Su X, Li J, Xu X, et al. Strategies to enhance the therapeutic efficacy of anti-PD-1 antibody, anti-PD-L1 antibody and anti-CTLA-4 antibody in cancer therapy[J]. *Journal of Translational Medicine*, 2024, 22(1): 751.
- [19] Cheng J, Liang T J, Xie X Q, et al. A new era of antibody discovery: An in-depth review of AI-driven approaches[J]. *Drug Discovery Today*, 2024, 29(6): 103984.
- [20] Zhao W B, Luo X W, Tong F, et al. Improving antibody optimization ability of generative adversarial network through large language model[J]. *Computational and Structural Biotechnology Journal*, 2023, 21: 5839–5850.
- [21] Cai H Y, Zhang Z B, Wang M K, et al. Pretrainable geometric graph neural network for antibody affinity maturation[J]. *Nature Communications*, 2024, 15(1): 7785.
- [22] Hie B L, Shanker V R, Xu D, et al. Efficient evolution of human antibodies from general protein language models[J]. *Nature Biotechnology*, 2024, 42(2): 275–283.
- [23] Mason D M, Friedensohn S, Weber C R, et al. Optimization of therapeutic antibodies by predicting antigen specificity from antibody sequence *via* deep learning[J]. *Nature Biomedical Engineering*, 2021, 5(6): 600–612.
- [24] Saka K, Kakuzaki T, Metsugi S, et al. Antibody design using LSTM based deep generative model from phage display library for affinity maturation[J]. *Scientific Reports*, 2021, 11(1): 5852.
- [25] Bennett N R, Watson J L, Ragotte R J, et al. Atomically accurate *de novo* design of antibodies with RFDiffusion[J]. *bioRxiv*, 2025: 2024.03.14.585103.
- [26] Liang T J, Sun Z Y, Hines M G, et al. AI-based IsAb2.0 for antibody design[J]. *Briefings in Bioinformatics*, 2024, 25(5): bbae445.
- [27] Zheng J Y, Wang Y, Liang Q Y, et al. The application of machine learning on antibody discovery and optimization[J]. *Molecules*, 2024, 29(24): 5923.
- [28] Jiang X F, Luo S Z, Liao K B, et al. Artificial intelligence and automation to power the future of chemistry[J]. *Cell Reports Physical Science*, 2024, 5(7): 102049.
- [29] You Y J, Lai X, Pan Y, et al. Artificial intelligence in cancer target identification and drug discovery[J]. *Signal Transduction and Targeted Therapy*, 2022, 7(1): 156.
- [30] Zhang X J, Zhang O, Shen C, et al. Efficient and accurate large library ligand docking with KarmaDock[J]. *Nature Computational Science*, 2023, 3(9): 789–804.
- [31] Lin H T, Huang Y F, Zhang O, et al. DiffBP: Generative diffusion of 3D molecules for target protein binding[J]. *Chemical Science*, 2025, 16(3): 1417–1431.
- [32] Xiong G L, Wu Z X, Yi J C, et al. ADMETlab 2.0: An integrated online platform for accurate and comprehensive predictions of ADMET properties[J]. *Nucleic Acids Research*, 2021, 49(W1): 5–14.
- [33] Baker F N, Chen Z Q, Adu-Ampratwum D, et al. RLSynC:

- Offline-online reinforcement learning for synthon completion[J]. *Journal of Chemical Information and Modeling*, 2024, 64(17): 6723–6735.
- [34] Zhavoronkov A, Ivanenkov Y A, Aliper A, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors[J]. *Nature Biotechnology*, 2019, 37(9): 1038–1040.
- [35] Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development[J]. *Nature Reviews Drug Discovery*, 2019, 18(6): 463–477.
- [36] Jiménez-Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence[J]. *Nature Machine Intelligence*, 2020, 2(10): 573–584.
- [37] Li X Y, Xu Y Q, Yao H Q, et al. Chemical space exploration based on recurrent neural networks: Applications in discovering kinase inhibitors[J]. *Journal of Cheminformatics*, 2020, 12(1): 42.
- [38] Zhang P, Zhang D F, Zhou W A, et al. Network pharmacology: Towards the artificial intelligence-based precision traditional Chinese medicine[J]. *Briefings in Bioinformatics*, 2023, 25(1): bbad518.
- [39] Himmelstein D S, Lizee A, Hessler C, et al. Systematic integration of biomedical knowledge prioritizes drugs for repurposing[J]. *eLife*, 2017, 6: e26726.
- [40] Steiner S, Wolf J, Glatzel S, et al. Organic synthesis in a modular robotic system driven by a chemical programming language[J]. *Science*, 2019, 363(6423): eaav2211.
- [41] Taherdoost H, Ghofrani A. AI's role in revolutionizing personalized medicine by reshaping pharmacogenomics and drug therapy[J]. *Intelligent Pharmacy*, 2024, 2(5): 643–650.

Progress of artificial intelligence-driven drug discovery

HE Xinheng, GAO Sihan, LI Junrui, XU Huaqiang*

Shanghai Institute of Materia Medica, Chinese Academy of Science, Shanghai 201203, China

Abstract Drug discovery, as the core driving force of the modern pharmaceutical industry, faces the difficulty of the traditional model's "high investment, long cycle, and low output," urgently requiring breakthroughs to address increasingly complex health demands. The rapid development of artificial intelligence (AI) technology has brought revolutionary changes to drug discovery, significantly enhancing efficiency and success rates in areas such as protein structure prediction, protein design, antibody drug design, and small molecule drug development. This article focuses on the domestic and international progress of AI in these key domains, providing an in-depth analysis of AI breakthrough in protein structure prediction and its potential applications in target discovery and virtual screening. It explores the closed-loop model of AI-driven protein design, from structure prediction to functional innovation, and examines AI's role in antibody sequence optimization, affinity maturation, and novel antibody design. Additionally, it reviews the latest achievements of AI in small molecule drug target identification, virtual screening, and ADMET optimization. The article also highlights challenges in AI applications, including data quality, model interpretability, and experimental validation, while envisioning future directions such as multimodal data integration, dynamic behavior prediction, and automated platforms. By comprehensively analyzing the current state and challenges of AI-enabled drug discovery, this article aims to offer scientific perspectives and insights to accelerate new drug creation and enhance human health and well-being. It seeks to provide readers with a thorough and insightful view of technological issues in AI-empowered drug discovery, stimulate thinking about future directions, and promote the more effective application of AI technologies in this field, ultimately benefiting human health through an accelerated drug development process.

Keywords artificial intelligence; drug development; AlphaFold3; protein design; antibody design; small molecule drug design ●



(责任编辑 徐丽娇)