从物理智能到分子设计：跨尺度机制学习推动药物发现

作者：writer 来源：科学网

本文原地址：https://www.iikx.com/news/progress/42330.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

上海交通大学医学院张健等在MedScience发表Perspective文章《从物理智能到分子设计：跨尺度机制学习推动药物发现》(From physics-based intelligence to molecular design: linking mechanistic learning across scales)。

文章围绕物理信息生成模型、分子基础模型、蛋白构象动力学和变构药物发现，讨论了人工智能药物设计正在发生的一项重要变化：分子生成不能只依赖已有数据中的统计规律，还需要回到能量、结构、构象变化和分子识别机制本身。

生成式模型正在改变药物分子设计方式。过去，候选分子的发现往往依赖已有化合物库筛选，或围绕已知分子骨架进行修饰。如今，模型可以在更广阔的化学空间中生成新的分子结构，并结合靶点口袋形状、药效团特征和蛋白—配体相互作用进行优化。这为早期药物发现提供了新的工具，也让传统实验难以充分覆盖的分子空间进入设计范围。

不过，药物分子并不是只要结构看起来合理就能发挥作用。真实生物体系中的分子识别同时受到三维构型、电子分布、局部相互作用和能量稳定性的影响。如果模型主要学习已有分子数据库中的相关性，而缺少物理规律和机制信息的约束，生成结果可能在化学形式上成立，却难以稳定结合靶点，也未必能够产生预期的功能调控效果。这也是文章讨论的核心：分子设计正在从统计相关走向物理机制。所谓物理信息，并不是在生成分子之后再做一次简单筛选，而是将势能、电子密度、分子对称性、局部相互作用和构象约束等因素纳入模型学习过程，使分子生成从一开始就处在更接近真实物理环境的条件下。

在小分子生成方面，相关方法已经从单纯学习配体结构，逐渐转向利用靶点口袋环境指导分子构建。LiGANN、FLAG、DeepICL、PocketFlow、DiffSMol和PhoreGen等模型分别引入结合口袋形状、片段模式、药效团关系和蛋白—配体相互作用信息，使候选分子的生成不再只依赖配体统计特征，而是更多受到靶点局部化学环境的引导。ED2Mol进一步以电子密度作为分子生成的基础表征，从蛋白口袋中推断配体密度图，并在此基础上逐步构建分子结构。由于电子密度能够反映潜在能量面和局部相互作用环境，这类方法有助于提高生成分子的物理合理性，并增强模型对未见靶点口袋的泛化能力。

分子基础模型和机器学习力场则从更大尺度推动了这一变化。传统分子力场具有清晰的物理基础，但依赖预设函数形式，在面对复杂化学空间和大尺度生物体系时往往需要大量参数调整。机器学习力场和分子基础模型试图从大规模分子与生物复合物数据中学习相互作用规律，并在同一表

示空间中整合结构、距离、取向、对称性和能量信息。这样一来，模型不只是预测某个分子是否可能结合靶点，也可以进一步描述不同构象状态的稳定性和能量差异。

蛋白质功能的复杂性，使这一问题更加突出。许多生物过程并不由单一静态结构决定，而是依赖不同构象状态之间的转换。AlphaFold等方法显著提高了蛋白静态结构预测能力，但单一结构很难解释远端位点如何影响活性中心，也难以描述变构调控中的能量传递和构象群体变化。因此，AI在结构生物学中的作用正在从预测结构扩展到理解构象集合。AI2BMD、BioEmu、AlphaFlow等方法分别借助机器学习力场、扩散生成框架和流匹配策略，提高了蛋白构象采样和能量景观表征的效率，为研究变构通信、隐蔽口袋形成和功能状态转换提供了新的工具。

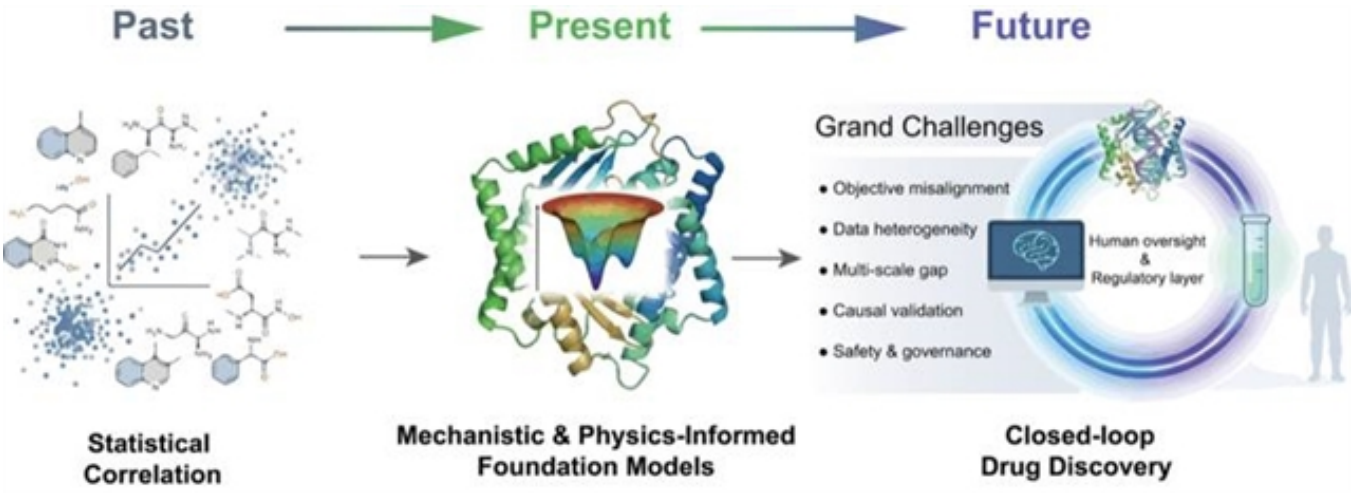


图1.通向闭环自主药物发现的概念路线图。分子设计正在从基于统计相关性的模型，走向机制与物理信息驱动的基础模型，并进一步指向闭环自主药物发现系统。该过程仍面临优化目标错配、数据异质性、多尺度整合不足、因果机制验证有限，以及安全治理和人工监督等挑战。

变构药物发现是检验机制学习能力的重要场景。与直接作用于保守正构位点的药物不同，变构药物通过结合远端调控区域改变蛋白构象分布，从而影响功能输出。这一策略有望提高选择性，也可能帮助克服耐药问题。但变构位点常常具有瞬时性和构象依赖性，有些口袋只在特定状态下形成，单靠静态结构、序列保守性或局部结合得分，很难判断一个位点是否真正具有调控意义。因此，变构药物设计的关键并不只是发现一个远端口袋，而是判断该口袋是否能够改变构象群体、影响自由能分布，并最终改变蛋白功能。对接分数或结合亲和力可以提供参考，但它们并不必然等同于变构调控效果。未来的变构药物发现需要把构象集合生成、通信通路推断、状态选择性分子设计和实验验证放在同一框架中考虑。

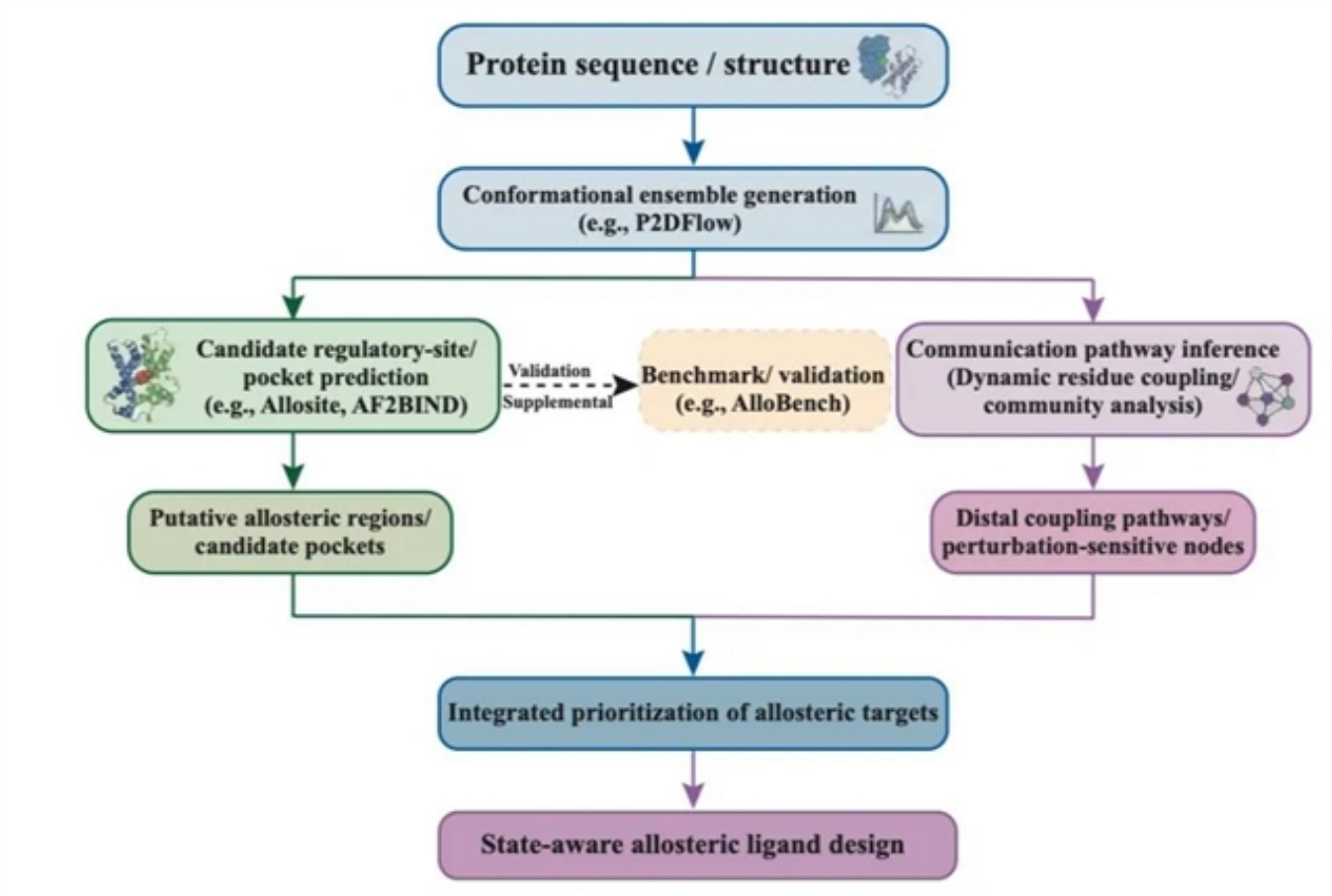


图2.AI驱动变构药物发现的概念路线图。AI驱动的变构药物发现需要整合构象集合生成、变构通信通路推断和状态选择性分子设计，并在热力学框架下结合实验观测，对模型预测进行验证和校正。

当前AI药物设计仍面临一些基础性问题。许多模型优化的是对接分数、重构精度或局部结合亲和力，但药物设计真正需要面向的是构象状态、自由能差异和生物功能的改变。与此同时，序列、结构、模拟轨迹和实验读数来自不同尺度，数据类型差异很大，从小分子生成到蛋白构象建模，再到残基通信分析和药理功能评价之间，也尚未形成稳定的多尺度连接。

药物设计最终仍要回到分子行为本身：候选分子能否稳定存在，蛋白构象是否发生转换，远端位点的通信能否改变功能状态，都需要接受物理规律和实验结果的检验。随着物理约束、构象动力学、变构通信和实验反馈不断进入模型体系，AI在药物发现中的角色也不再停留于筛选和生成候选分子，而是逐步延伸到解释并调控蛋白质内部的能量网络，为更可解释、可验证、可控的药物设计提供基础。（来源：EngineeringJournals微信公众号）

相关论文信息：https://journal.hep.com.cn/fmd/EN/10.1007/s11684-026-1263-6

特别声明：本文转载仅仅是出于传播信息的需要，并不意味着代表本网站观点或证实其内容的真实性；如其他媒体、网站或个人从本网站转载使用，须保留本网站注明的“来源”，并自负版权等法律责任；作者如果不希望被转载或者联系转载稿费等事宜，请与我们接洽。
作者：张健等 来源：《医学科学》

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发