

人工智能与天然药物 动态监测快报

总第 011 期 2025-5-30



中国科学院昆明植物研究所
云南省植物学会
中国科学院成都文献情报中心

目 录

政策法规	3
上海市人民代表大会常务委员会修改《上海市浦东新区促进张江生物医药产业 创新高地建设规定》	3
研究前沿	4
Science: AI 驱动的动态蛋白质从头设计.....	4
HighPlay: 基于强化学习和蛋白质结构预测的环肽序列设计	5
新型反应描述语言用于编码化学反应中的分子编辑操作	6
大规模转录组挖掘实现千金藤酸支架的大环多样化与生物活性增效	7
标签特征化: 基于网络同质性系统注释人源 GPCR 药物-靶点相互作用	8
PoseidonQ: 面向药物发现的免费机器学习平台, 实现高效可迁移 QSAR 模型的 开发、解析与验证	9
神农 Alpha: 一个人工智能 (AI) 驱动的共享和协作平台, 用于智能整理、 获取和翻译 NMM 知识	10
近期会议与活动	11

政策法规

上海市人民代表大会常务委员会修改《上海市浦东新区促进张江生物医药产业创新高地建设规定》

2025-04-29

上海市第十六届人民代表大会常务委员会第二十一次会议通过，对《上海市浦东新区促进张江生物医药产业创新高地建设规定》的第二次修正，修订后的规定自2025年4月30日起施行。

1. 增加一条，作为第十三条：对正在开展临床试验用于治疗严重危及生命且尚无有效治疗手段的疾病的药物，经医学观察可能使患者获益，并且符合伦理要求的，经审查、告知患者并取得知情同意后，可以在开展该药物临床试验的医疗机构内用于其他病情相同且无法参加药物临床试验的患者。

2. 增加一条，作为第二十条：浦东新区药品监管部门根据市药品监管部门的委托，依法实施第二类、第三类医疗器械生产许可，并按照职责对相关企业的生产活动进行监督检查。市药品监管部门应当加强指导和监督。

3. 将原第二十条改为第二十二條，增加一款，作为第三款：本市支持符合条件的医疗机构就浦东新区生物医药企业相关技术和产品开展临床研究和临床试验，优化管理流程，提高临床研究和临床试验质效。

4. 增加一条，作为第三十六条：浦东新区生物医药相关行业组织依照法律、法规和章程，发挥协调和自律作用，及时反映行业诉求，为会员提供信息咨询、宣传培训、市场拓展、权益保护、纠纷处理等方面的服务。

信息来源：上观新闻！

<https://sghexport.shobserver.com/html/baijiahao/2025/04/30/1558518.html>

研究前沿

Science: AI 驱动的动态蛋白质从头设计

2025-5-23

这项突破性研究开创了利用深度学习技术设计动态蛋白质的新范式。科学家们巧妙地将 AlphaFold2 的预测能力与分子动力学模拟相结合，成功构建出能像天然信号蛋白那样发生精确构象变化的合成蛋白质。

研究团队首先锁定了一个静态的钙离子(Ca²⁺)结合蛋白作为起点，通过计算机模拟筛选出潜在的替代构象。开发的关键技术在于：通过“硅基突变扫描”(in silico mutational scan)识别决定各构象状态的最小残基集，再借助深度学习指导序列优化，最终获得能在不同构象间动态转换的蛋白质设计。

这些设计不仅通过了核磁共振(NMR)和分子动力学模拟的双重验证，更展现出令人惊喜的调控特性：仅需改变单个氨基酸位点，就能显著改变构象平衡；通过引入钙离子结合位点，实现了用常见第二信使 Ca²⁺来调控蛋白质构象的巧妙设计。

特别值得注意的是，这些合成蛋白质展现出与天然信号蛋白类似的全 osteric 调控机制——远端突变可以精细调节钙离子结合位点的亲和力。深度学习预测、物理模拟和实验数据三者高度吻合，使得研究人员能够解析并重新编程构象特异的原子相互作用网络。

这项成果首次实现了对蛋白质内部结构域运动(intradomain motion)的从头设计。这种受自然启发的设计策略，为构建具有复杂信号整合能力的人工蛋白质系统开辟了新途径，未来不仅可用于设计全新的生物分子器件，还能改造天然蛋白质的功能，更是为构建未来生物计算、智能药物递送，乃至全新生命系统奠定了坚实的基础。

信息来源: <https://www.ebiotrade.com/newsf/2025-5/20250523073905371.htm>

DOI: 10.1126/science.adr7094

HighPlay: 基于强化学习和蛋白质结构预测的环肽序列设计

2025-5-23

药环肽结构的多样性和良好的生物相容性使其成为潜在的治疗药物。现有的环肽设计方法,无论是传统的还是新兴的人工智能辅助的,都依赖于大量的实验,面临着分子多样性有限、成本高、耗时等挑战。在本研究中,我们提出将强化学习(MCTS)与 HighFold 结构预测模型相结合的 HighPlay,仅基于目标蛋白序列信息设计环肽序列,实现环肽序列和结合位点的协同优化,并在不需要预定义目标信息的情况下动态探索序列空间。将该模型应用于三种不同靶点的环肽序列设计,通过分子动力学模拟对其进行筛选和验证,显示出良好的结合亲和力。具体来说,为 TEAD4 靶点设计的环肽序列在进一步的实验验证中显示出微摩尔水平的亲和力。

信息来源: Huitian Lin et.al. HighPlay: Cyclic Peptide Sequence Design Based on Reinforcement Learning and Protein Structure Prediction [J]. Journal of Medicinal Chemistry, 2025 DOI: 10.1021/acs.jmedchem.5c00896

新型反应描述语言用于编码化学反应中的分子编辑操作

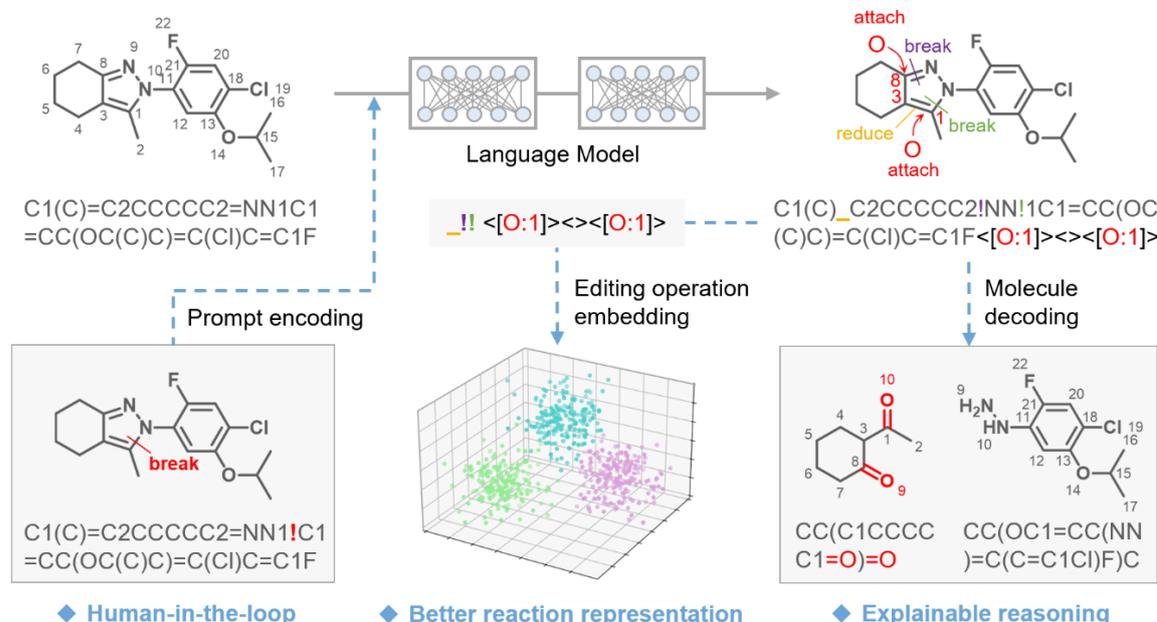
2025-05-13

中国科学院上海药物研究所开发出新型反应描述语言 ReactSeq。该语言基于逆合成分析思想，不仅定义产物结构，更明确了将其转化为反应物分子所需的分子编辑操作（MEOs），包括化学键断裂/变更、原子电荷改变、离去基团（LGs）连接等。基于 ReactSeq 构建的逆合成语言模型并非逐符号生成反应物，而是通过对产物分子实施 MEOs 实现转化，确保预测反应物与产物间的精确原子映射，显著提升模型可解释性。

使用 ReactSeq 时，普通 Transformer 模型即可实现当前最先进的逆合成预测性能。ReactSeq 特有的 MEOs 显式标记能编码人类指令：专家提示词可大幅提升模型表现，甚至引导其探索新反应。这些标记还有利于反应表征提取——相比聚合整个 ReactSeq 嵌入向量，聚焦 MEOs 标记嵌入可获得更本质、更可靠的反应表征。

研究团队结合该策略与自监督学习，开发出通用可靠的反应表征方法，在反应分类、相似反应检索、反应产率预测及实验步骤推荐等多项任务中验证了其有效性。

Our proposed method (SMILES to ReactSeq)

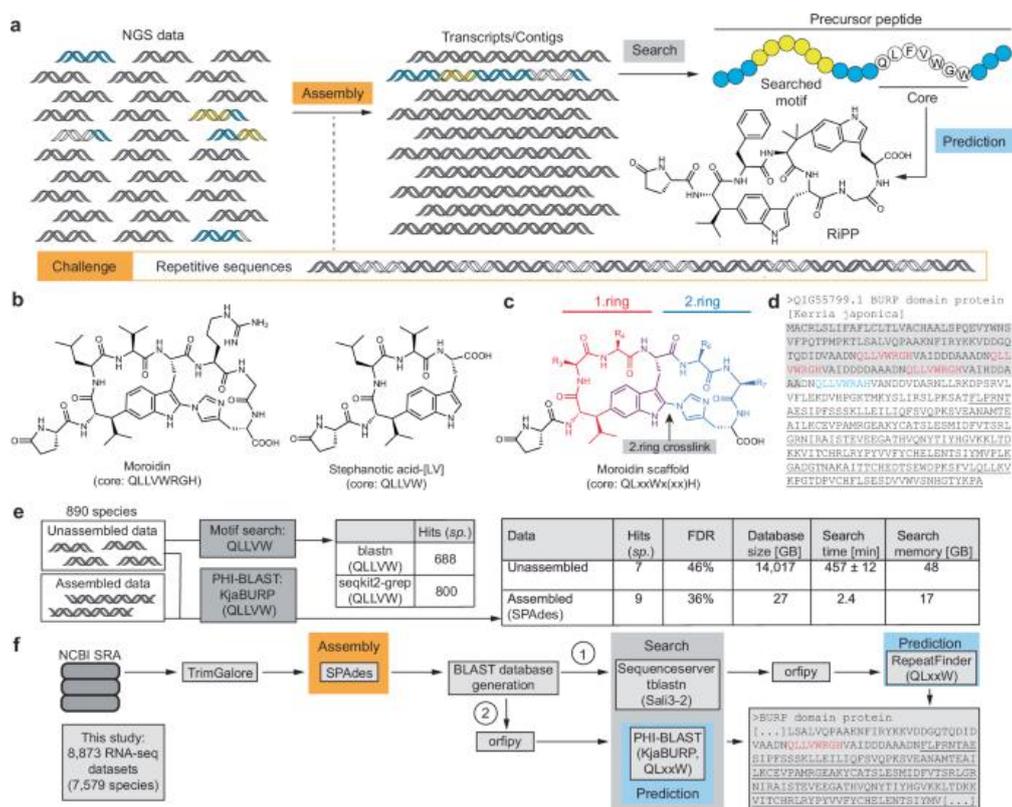


信息来源: Jiacheng Xiong, et.al. Bridging chemistry and artificial intelligence by a reaction description language[J] Nature Machine Intelligence volume 7, pages782–793 (2025)
10.1038/s42256-025-01032-8

大规模转录组挖掘实现千金藤酸支架的大环多样化与生物活性增效

2025-05-06

NCBI 序列读段存档库收录近万种植物的 RNA-seq 数据集，但因数据库规模庞大，未组装数据难以检索。本研究通过优化 RNA-seq 组装流程，将绝大部分公共 RNA-seq 数据转化为可检索的生物合成基因发现数据库。以植物核糖体合成翻译后修饰肽（RiPPs）——铜依赖性肽环化酶合成的吗啉类化合物为对象，测试了转录组挖掘流程的多样性应用。吗啉类化合物具有保守的千金藤酸双环支架，其 C 末端大环化结构对非小细胞肺腺癌细胞具有细胞毒性。研究发现多种第二环结构多样化的吗啉类似物，包括交叉连接与非交叉连接残基修饰类型，其中繁缕来源的吗啉类似物对肺腺癌细胞的毒性强于原型化合物。本研究中将千金藤酸型肽的分布扩展至禾本科、兰花蕉科、唇形科、石竹科及大戟科植物，同时证明大规模转录组挖掘能为真核生物 RiPP 先导结构的化学生物学探索提供更丰富的医药化学资源库。



信息来源: Xiaofeng Wang et al., Large-scale transcriptome mining enables macrocyclic diversification and improved bioactivity of the stephanotic acid scaffold [J]. Nature Communications

2025 DOI: 10.1038/s41467-025-59428-4

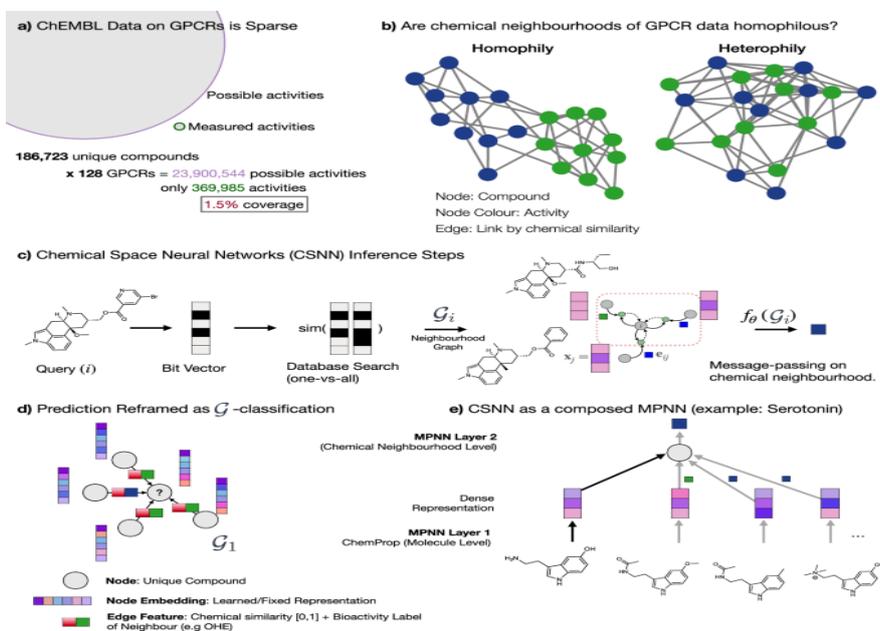
标签特征化：基于网络同质性系统注释人源 GPCR 药物-靶点相互作用

2025-5-3

人类 G 蛋白偶联受体 (hGPCR) 是约 34% FDA 批准药物的作用靶点, 但已知药物-靶点相互作用 (DTI) 覆盖率仅占潜在相互作用的 1.5%, 其余 98.5% 处于未探索的"黑暗"状态, 导致药物脱靶效应难以预测, 成为药物安全性的重大隐患。

区别于传统关注分布外探索 (药物发现) 的研究范式, 研究者提出名为"化学空间神经网络"的邻域预测模型。该创新方法利用网络同质性原理, 采用免训练的图神经网络架构, 并将标签转化为特征输入。研究表明, 化学空间神经网络的预测准确性与网络同质性呈现强相关性——通过推理阶段整合标记数据, 标签特征化策略显著提升了机器学习模型的分布内预测精度。

研究者在高通量酵母生物传感系统 (涵盖 3773 组药物-靶点相互作用、539 种化合物、7 种人源 G 蛋白偶联受体) 中验证了该模型的突破性进展: 不仅发现了 FDA 批准药物的新型靶向作用, 更拓展了构建可靠预测模型以指导实验验证的普适方法论。



信息来源: Frederik G. Hansson et.al., Labels as a feature: Network homophily for systematically annotating human GPCR drug-target interactions [J]. Nature Communications 2025
DOI:10.1038/s41467-025-59418-6

PoseidonQ: 面向药物发现的免费机器学习平台, 实现高效可迁移 QSAR 模型的开发、解析与验证

2025-04-09

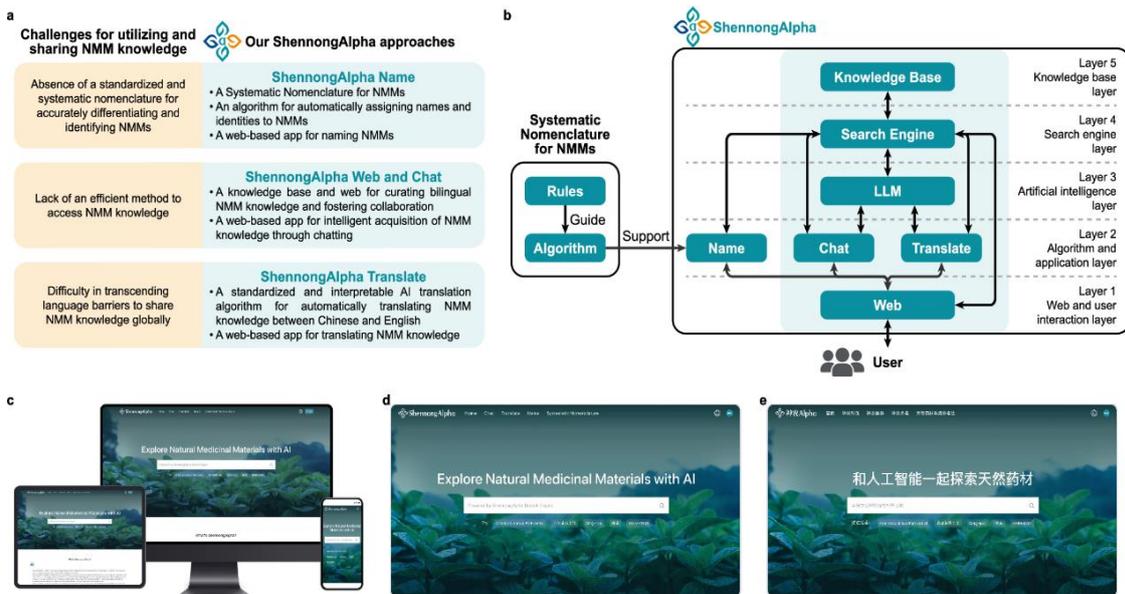
强大的机器学习算法的出现以及大量药理学数据的可用性为 QSAR 提供了新的燃料, 为获得高度预测模型提供了前所未有的选择, 以协助新的生物活性化合物的基本原理设计, 筛选和优先考虑大型分子文库, 以及将新药重新用于新的临床用途。研究者推出了 PoseidonQ (Personal Optimization Software for Efficient Implementation and Derivation of Online QSAR 的首字母缩写), 这是一个用户友好的软件解决方案, 旨在简化用于药物设计和发现的 QSAR 模型的推导。PoseidonQ 结合了 22 种机器学习算法, 17 种类型的分子指纹和 208 个 RDKit 分子描述符, 并能够快速推导回归和分类模型, 以及计算和易于解释的适用性域。重要的是, 该平台自动链接到 ChEMBL 数据库的最新版本, 从而提供了对大量精心整理的生物活性数据的简化访问, 用户还可以根据可定制的过滤设置收集高质量的实验数据。PoseidonQ 通过与 Streamlit Cloud 和 GitHub 的无缝集成, 促进了训练有素的 QSAR 模型作为基于 web 的应用程序的部署, 使用户能够毫不费力地访问、共享、改进和集成模型。PoseidonQ 适用于 Windows 和 Linux (ubuntu 22.04 发行版) 操作系统, 可以从 <https://github.com/Muzatheking12/PoseidonQ> 免费下载。

信息来源: Muzammil Kabier et al. PoseidonQ: A Free Machine Learning Platform for the Development, Analysis, and Validation of Efficient and Portable QSAR Models for Drug Discovery [J]. Journal of Chemical Information and Modeling 2025 DOI: 10.1021/acs.jcim.4c02372.

神农 Alpha: 一个人工智能 (AI) 驱动的共享和协作平台, 用于智能整理、获取和翻译 NMM 知识

2025-04-01

天然药材 (NMM) 在全球临床应用方面有着悠久的历史, 并拥有丰富的记录和知识。尽管 NMM 是药物发现和临床应用的主要来源, 但 NMM 知识的利用和共享面临着关键挑战, 包括关键信息的标准化描述、高效的管理和获取以及语言障碍。为了解决这些问题, 科研人员开发了 ShennongAlpha, 一个人工智能 (AI) 驱动的共享和协作平台, 用于智能知识管理、获取和翻译。对于标准化知识策展, 该平台引入了系统命名法, 以实现 NMM 的准确区分和识别。超过 14,000 名中国 NMM 及其知识已被纳入该平台。此外, 该平台还开创了基于聊天的知识获取、标准化机器翻译和协作知识更新。



信息来源: Zijie Yang et.al. ShennongAlpha: an AI-driven sharing and collaboration platform for intelligent curation, acquisition, and translation of natural medicinal material knowledge [J]. Cell Discovery 2025 DOI: 10.1038/s41421-025-00776-2

近期会议与活动

时间	题目	地点	相关链接
14-15 Jul 2025	15th World Glycobiology Congress	Berlin, Germany	https://www.clocate.com/glycobiology-world-congress/48957/
18-20 Jul 2025	The 5th Int'l Conference on Medicinal Chemistry and Drug Discovery(MCDD 2025)	Kunming, China	https://www.academicx.org/MCDD/2025/
2025年7月 18日-20日	中国化学会第十二届天然产物全合成-青年学术研讨会	贵阳, 中国	https://www.chemsoc.org.cn/meeting/home/m610
2025年7月 28日-30日	中国化学会第七届资源化学学术研讨会	昆明, 中国	https://www.chemsoc.org.cn/meeting/home/m609
2025年8月 22日-24日	中国化学会第十五届全国天然有机化学学术会议	沈阳, 中国	https://www.chemsoc.org.cn/meeting/home/m632
2025年9月 23日-25日	2025 亚洲药物设计大会	杭州, 中国	https://ahedd.aconf.org/

人工智能与天然药物动态监测快报

编写委员会

主 编：庄会富 陆颖

执行主编：杜宁

编 委：张于 颜秉超 胡坤 卿立燕 杨帅

编 辑：杜宁 田雅娟 黄蓉 史继强

工作联系人：杜 宁 电话：0871-65223176 邮箱：duning@mail.kib.ac.cn

田雅娟 电话：028-81258707 邮箱：tyj@clas.ac.cn