
中国科大发展出蛋白质主链去噪扩散概率模型

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/29843.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！

中国科大发展出蛋白质主链去噪扩散概率模型。

中国科学技术大学教授刘海燕与陈泉团队开发了不依赖于预训练结构预测网络的蛋白质主链去噪扩散概率模型SCUBA-D (SCUBA-diffusion)。SCUBA-D可自动从头设计主链结构或指定功能位点生成主链结构。大量实验结果验证了SCUBA-D的设计成功率和设计精度。10月9日，相关研究成果以*De novo protein design with a denoising diffusion network independent of pretrained structure prediction models*为题，在线发表在《自然-方法》(Nature Methods) 上。

刘海燕和陈泉团队致力于发展数据驱动的蛋白质设计方法。前期，该团队建立并实验验证了利用神经网络能量函数从头设计主链结构的SCUBA模型。而此次研究报道的SCUBA-D模型是基于深度学习的主链设计算法的迭代升级。SCUBA-D能够基于不同输入执行多类蛋白质结构设计任务。在模型设计上，通过在扩散模型训练中引入对抗损失，避免生成模型产生物理上不可行的结构，实现了高成功率的主链结构设计。由于没有使用已有结构预测网络作为预训练降噪网络，SCUBA-D可在设计中避免对已知天然结构的过度偏好，并能够发现已有模型在可设计蛋白质结构空间中的盲区。

该研究对SCUBA-D在多类蛋白质从头设计任务中的应用进行了实验验证。针对单体结构从头设计任务，研究对70条设计序列进行实验表征发现，近80%的序列(53条)可溶表达，实验解析的16个高分辨晶体结构与目标结构高度一致。在小分子结合蛋白设计任务中，研究对非经典血红素降解酶进行了保留结合位点的主链结构重设计，对设计的12条序列进行了实验验证。其中，5条序列具有与血红素的结合能力，3条序列与血红素的亲和力与天然蛋白相当或高于天然蛋白。在结合蛋白设计任务中的30个人工设计的Ras结合蛋白中，14个设计蛋白与Ras有相互作用，其中3个设计蛋白与Ras的结合亲和力与天然蛋白相当，且复合物晶体结构进一步验证了设计的精确度。

研究工作得到科学技术部、国家自然科学基金委员会、中国科学院等的支持。

[论文链接](#)

通过SCUBA-D模型进行蛋白质主链设计的原理。(a) SCUBA-D可基于噪声或者用户定义的结构草图设计新主链结构，也可以基于给定的含功能位点的局部结构设计新的完整主链；(b) 包含对抗损失函数的SCUBA-D模型架构。

研究团队单位：中国科学技术大学

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](https://www.iikx.com)转发