

DOI: 10.13376/j.cblls/2025082

文章编号: 1004-0374(2025)07-0818-10

· 蛋白质设施的主要技术系统与特色 ·



秦文明, 理学博士, 正高级工程师。中国科学院关键技术人才, 上海市东方英才拔尖人才。2003 年获得山东大学学士学位, 2009 年毕业于中国科学院上海生命科学研究院并获得博士学位, 美国阿贡国家实验室访问学者。现就职于中国科学院上海高等研究院国家蛋白质科学研究(上海)设施, 完整地参与了蛋白质设施的设计、建设和运行。研究领域为同步辐射线站方法学研究和基于晶体结构的药物发现等。作为项目负责人承担了上海市东方英才拔尖人才、中国科学院关键技术人才、中国科学院重大科技基础设施维修改造项目、战略性先导科技专项 B 类项目子课题、国家自然科学基金等多项课题。

X射线晶体学在生物大分子研究中的应用与展望

肖清洁, 吴婷婷, 顾亦君, 秦文明*

(中国科学院上海高等研究院国家蛋白质科学研究(上海)设施, 上海 201210)

摘要: X 射线晶体学作为解析生物大分子高分辨率三维结构的重要技术, 在蛋白质功能研究与药物研发中发挥重要作用。国家蛋白质(上海)设施晶体学线站自开放运行以来, 通过在数据采集能力、自动换样系统、探测器配置及数据处理流程等关键环节不断优化, 全面提升了线站的整体性能, 提高了结构解析的效率。依托高亮度同步辐射光源与成熟的自动化技术平台, 晶体学线站助力用户在药物设计、病毒机制解析、基因编辑、植物免疫等多个研究领域取得了一系列具有国际影响力的成果, 已助力发表相关论文 2 000 余篇, 累计在 PDB 数据库中发布高质量生物大分子三维结构近 5 000 个。未来, 晶体学线站将持续推进技术创新, 强化与冷冻电镜、核磁共振、小角散射和人工智能结构预测等方法的互补应用, 拓展从静态结构解析到动态机制研究的能力, 为生命科学研究和精准药物开发提供坚实的技术支撑。

关键词: X 射线晶体学; 同步辐射; 蛋白质结构解析; 高通量数据采集; 自动化晶体学平台; 结构指导药物发现

中图分类号: Q6-33 文献标志码: A

Application and prospects of X-ray crystallography in biomacromolecules

XIAO Qing-Jie, WU Ting-Ting, GU Yi-Jun, QIN Wen-Ming*

(National Facility for Protein Science in Shanghai, Shanghai Advanced Research Institute,
Chinese Academy of Sciences, Shanghai 201210, China)

Abstract: X-ray crystallography is one of the most important techniques for studying the detailed 3D structures of biomacromolecules. It plays a key role in understanding how proteins work and in supporting drug development. Since the crystallography beamline at the National Facility for Protein Science (Shanghai) became operational, many improvements have been made in areas such as data collection, automated sample handling, detector setup,

收稿日期: 2025-05-30

基金项目: 国家重大科技基础设施运维项目(2022000014); 中国科学院关键技术人才项目(2021000022); 国家自然科学基金项目(32200988)

*通信作者: E-mail: qinwenming@sari.ac.cn

and data processing. These upgrades have made the platform faster and more efficient for solving structures. Due to its high-brightness synchrotron source and advanced automation systems, this facility has helped researchers make important discoveries in fields like drug design, virus research, gene editing, and plant immunity. So far, it has supported the publication of over 2,000 scientific papers and contributed nearly 5,000 high-quality biomolecular structures to the Protein Data Bank (PDB). In the future, the crystallography beamline will continue to innovate and work together with other structural biology methods, such as cryo-electron microscopy, nuclear magnetic resonance, small-angle scattering, and AI-based structure prediction. These combined approaches will help scientists move beyond static snapshots of molecules to better understand how they function and change, offering strong support for life science research and precision drug discovery.

Key words: X-ray crystallography; synchrotron radiation; protein structure determination; high-throughput data collection; automated platform; structure-based drug discovery

X 射线晶体学是一种基于 X 射线衍射原理, 解析晶体中原子级排列结构的重要技术。当晶体样品受到 X 射线照射时, 入射射线与晶体内部电子发生相互作用, 产生规律性的衍射图样。通过对这些衍射信号的计算分析, 可以准确重建分子三维结构信息, 分辨率可达到原子级别^[1]。作为结构生物学领域的核心技术手段之一, X 射线晶体学广泛应用于蛋白质、核酸等生物大分子及其复合物的高分辨率结构解析, 是理解大分子功能机制的重要方法^[2]。在科学研究中, 该技术为揭示酶催化机制、信号转导过程中受体-配体识别、蛋白质突变导致的结构变化及其功能影响等关键问题提供了直观的结构依据。在新药研发领域, X 射线晶体学也已经成为结构指导药物发现 (structure-based drug discovery, SBDD) 的核心技术手段, 广泛应用于先导化合物筛选、结合模式确认及分子优化设计等关键环节^[3]。通过高精度解析蛋白质与小分子配体的复合物结构, 能够准确确定结合位点信息和蛋白质与配体之间的相互作用方式, 为药物作用机制研究和药效优化提供可靠的结构信息。

根据蛋白质数据库 (Protein Data Bank, PDB) 统计^[4], 全球约 80% 的生物大分子三维结构数据通过 X 射线晶体学的方法解析获得 (图 1)。X 射线晶体学在分子量较小或中等的蛋白质、蛋白质复合物以及小分子配体复合物研究中有着明显的优势, 并且具备数据采集流程标准化、数据采集速度快、自动化程度高、适合高通量筛选等特点, 已成为药物设计与精准医学研究中不可替代的重要手段。与此同时, X 射线晶体学与其他结构生物学技术形成了良好的互补体系。核磁共振 (NMR) 技术适用于溶液状态下分子的动态行为和构象变化分析^[5], 冷冻电镜 (Cryo-EM) 则在解析大型复合体、柔性蛋白及难

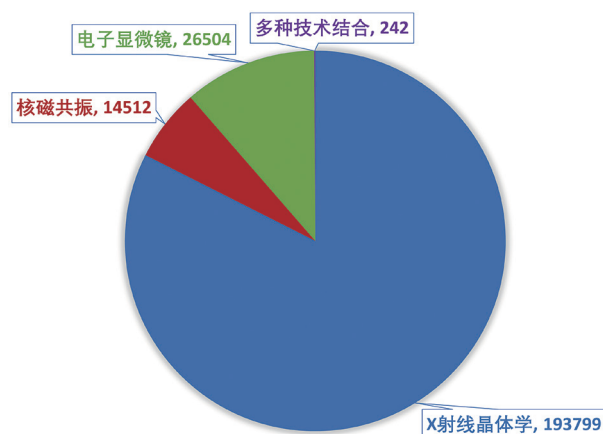


图1 截至2025年4月26日, 不同方法解析生物大分子结构数量(<https://www.rcsb.org/stats>)

结晶样品方面具有独特优势^[6]。多种技术手段的协同应用正在推动结构生物学研究从静态结构解析逐步迈向动态机制解析, 进一步拓展了生物大分子功能研究的深度与广度。

总体而言, X 射线晶体学作为目前技术最为成熟、应用最广泛的高分辨率结构解析方法, 未来将在生命科学研究与药物开发等方面持续发挥关键作用, 并在多技术融合与自动化、智能化发展的推动下, 不断拓展其在结构生物学领域的应用。

1 生物大分子晶体学线站的发展现状与趋势

X 射线晶体学在解析生物大分子高分辨率三维结构中的应用与发展, 高度依赖于 X 射线光源技术的持续演进。自同步辐射光源应用于生物大分子结构研究以来, 光源性能不断提升, 分辨率、数据采集效率以及动态机制解析等方面的技术瓶颈也在持续突破。第一代同步辐射光源源自高能物理加速器

的副产辐射,采用“寄生模式”获取X射线。其获得的射线亮度低、束流稳定性差、发散度大,难以满足生物大分子高分辨率结构解析的需求。第二代光源通过专门设计储存环和弯转磁体作为辐射源,显著改善了束流的亮度和稳定性,逐步应用于生物大分子结构研究。第三代同步辐射光源在储存环设计上进一步优化,大量引入插入器(如多极扭摆器、波荡器等)装置,实现了X射线亮度和发散度等性能数量级的提升。高亮度、微束聚焦及快速采集能力显著增强,使得膜蛋白、微晶、复杂大分子复合物等高难度样品的研究成为可能^[7,8]。

目前,全球范围许多第三代同步辐射光源内均有建设成熟的生物大分子晶体学(macromolecular crystallography)线站。典型代表包括瑞士光源(SLS)的X06DA线站^[9]、日本SPring-8的BL32XU线站^[10]、英国Diamond光源的I03和I04线站^[11],以及中国上海同步辐射光源(SSRF)的BL17U1、BL18U1和BL19U1线站^[12,13]等。依靠第三代光源在束流亮度、稳定性及微束聚焦能力方面的优势,这些生物大分子晶体学线站在全球生物大分子结构解析领域占据了重要地位,尤其在中小分子体系、配体-靶点复合物研究以及片段化合物筛选(fragment-based screening, FBS)等方面展现出广泛的应用和持续的技术领先性。根据蛋白质数据库(PDB)的统计,目前大部分已解析的晶体结构数据仍主要来源于第三代同步辐射光源平台。

随着第四代高亮度、高相干性X射线光源体系的加速建设,基于储存环的衍射极限储存环(diffraction-limited storage ring, DLSR)^[14]和基于直线加速器的自由电子激光(X-ray free electron laser, XFEL)正逐步成为动态结构生物学研究的重要技术支撑^[15]。DLSR通过进一步降低电子束发射度,在束流亮度上相较第三代光源提升2~3个数量级,空间和时间相干性显著增强,特别适合高分辨率成像及微束聚焦实验。XFEL则在峰值亮度方面实现了8~10个数量级的提升,具备飞秒级脉冲宽度和高度相干性,能够在辐射损伤发生前完成数据采集,采用“衍射前毁灭”(diffraction-before-destruction)策略,突破了传统同步辐射光源在时间分辨能力上的限制。该技术在捕捉生物大分子反应过程中短寿命中间态、解析超快动力学变化等方面具有独特优势,正为结构生物学在超快动力学和动态机制研究领域开辟新的技术路径^[16,17]。

尽管第四代光源在特定前沿领域展现出巨大潜

力,然而第三代同步辐射光源在常规高分辨率晶体学实验中的成熟度、稳定性、通量优势及性价比依然不可替代。特别是在药物发现中的高通量片段筛选、常规蛋白质结构解析及自动化数据采集领域,第三代光源的软硬件体系已高度成熟,实验流程标准化,配套设施完善,具备极高的实验效率与广泛的用户适配性^[18,19]。在满足大规模样品筛选和结构解析需求方面,第三代光源晶体学线站仍将在未来相当长一段时间内保持强有力的竞争优势,继续支撑生物大分子结构生物学与药物开发的主力应用场景。因此,未来生物大分子晶体学的发展将呈现出第三代光源持续主导常规高通量解析领域、第四代光源补充超快动态研究的多元协同格局。多技术融合将进一步拓展结构生物学研究的广度和深度,助力生命科学与精准医疗不断向前推进。

2 蛋白设施晶体学线站介绍

国家蛋白质科学研究(上海)设施(以下简称“蛋白质设施”)是我国生命科学领域的重要国家重大科技基础设施,旨在构建国际一流的结构生物学研究体系。设施集成了多种先进的大型仪器平台,为获取生物大分子的高分辨率、动态及原位结构信息提供系统化、规模化的技术支持,服务于科研院所、高校及生物医药企业的前沿研究与创新。蛋白质设施的晶体结构分析系统由位于上海光源的三条同步辐射X射线蛋白质晶体学专用线站组成,包括BL19U1、BL18U1和BL17B1,于2015年7月正式通过国家验收并开放运行。这三条线站在功能定位和技术构建上各具特色,协同支持不同类型、不同尺寸、不同难度的蛋白质和复合物的晶体结构研究,广泛服务于生命科学中的关键科学问题。BL19U1蛋白质复合物晶体结构线站主要面向超大分子复合物晶体结构解析,理论上支持晶胞尺寸可达3 000 Å的大尺寸晶体测定^[20]。针对复合物晶体分子量大、晶胞大、衍射弱等特点,该线站配置了高亮度、高稳定性、低发散度的光束系统,确保高分辨率衍射数据的采集,特别适用于结构复杂的大型蛋白质复合物研究。根据目前公开的结构数据,BL19U1线站已成功测定一些晶胞参数较大的蛋白质复合物结构,其中一个轴的长度甚至大于600 Å。例如PDB编号为5XL3蛋白晶体,其晶胞参数为 $a=100.5$ Å, $b=100.5$ Å, $c=685.1$ Å^[21];PDB编号为8YQ0蛋白晶体,其晶胞参数为 $a=108$ Å, $b=108$ Å, $c=659.6$ Å^[22]。BL18U1蛋白质微晶体结构线站针对晶体尺寸微小、

衍射能力较低的问题, 设计了兼备高通量和微聚焦能力的 X 射线光束系统, 适用于晶体尺寸在 10~20 μm 之间的蛋白质, 特别是膜蛋白及难结晶蛋白质的结构解析。依托该线站目前有一批包括 G 蛋白偶联受体 (GPCR) 等膜蛋白结构已经被解析^[23-25]。BL17B1 高通量蛋白质晶体结构线站以自动化、高通量为核心目标, 面向大规模样品的快速筛选与结构测定, 支持多波长异常散射法 (MAD)、单波长异常散射法 (SAD) 等多种晶体学方法, 具备高能量分辨率与宽波长范围调节能力。表 1 是三条晶体学线站的性能指标参数。

在开放运行的十年间, 蛋白质设施晶体学线站持续推进性能的优化与技术升级。图 2 展示了 BL19U1 实验站内关键设备变化, 图 2A 为初期设备布局, 图 2B 为升级后的设备布局。围绕高效数据采集与实验自动化需求, 线站在探测器系统、样品更换速度、衍射仪精度及控制软件平台等多个核心环节完成了系统性改进。这些技术革新显著提升了线站的整体性能和实验运行效率, 为高质量晶体结构解析提供了坚实的技术保障。

在硬件方面, 三条线站的探测器均采用 Pilatus

系列高频探测器, 从而实现了无快门 (shutter-free) 数据采集模式, 显著提升了数据收集速度。采集效率从早期的每小时 2~3 套数据提升至现在的 20 套以上。线站均配备自动化样品抓取系统, 支持远程进行样品更换, 无须人工换样, 进一步提升了实验的效率和自动化水平。线站还配备了高精度的衍射仪 MD2, 具备样品对中与旋转控制功能, 确保了衍射数据采集过程中的高重复性和高质量数据获取。近年来, BL19U1 衍射仪已经从原先的 MD2 型号升级至性能更优的 MD2S 型号, 自动对中系统的响应速度更快、定位精度更高, 整体样品视频更加清晰, 且整体运动平台的速度与稳定性得到了优化, 从而进一步保障了样品的高质量采集。BL19U1 机械手系统的样品装载量由原先的 90 颗扩展至 500 颗以上, 采用双抓手设计, 支持同步上样与下样, 单次换样时间从 1 min 以上缩短至 10 s 以内, 显著提升了样品处理能力和实验通量。此外, BL19U1 线站还引入了湿度控制装置及快速喷嘴交换系统 (REX), 可与冷冻气流系统无缝集成, 并通过统一的控制平台灵活切换多种实验模式。

值得一提的是, 目前该线站已集成室温下原位

表1 蛋白质设施三条晶体学线站的性能和指标

晶体学线站	BL17B1	BL18U1	BL19U1
光子能量范围(keV)	5~23	5~18	7~15
能量分辨率($\Delta E/E$)	3×10^{-4}	2×10^{-4}	2×10^{-4}
聚焦光斑尺寸(H×V) (μm^2)	$\leq 150\times 180$	$\leq 10\times 7$	$\leq 130\times 80$
光子通量(phs/s 12 keV @300mA)	$\geq 3\times 10^{11}$	$\geq 5.0\times 10^{11}$	$\geq 1\times 10^{12}$
光束发散角(H×V) (mrad^2)	$\leq 1.5\times 0.2$	$\leq 0.70\times 0.25$	$\leq 0.1\times 0.1$
探测器型号	Pilatus3 2M	Pilatus3 6M	Pilatus3 6 M

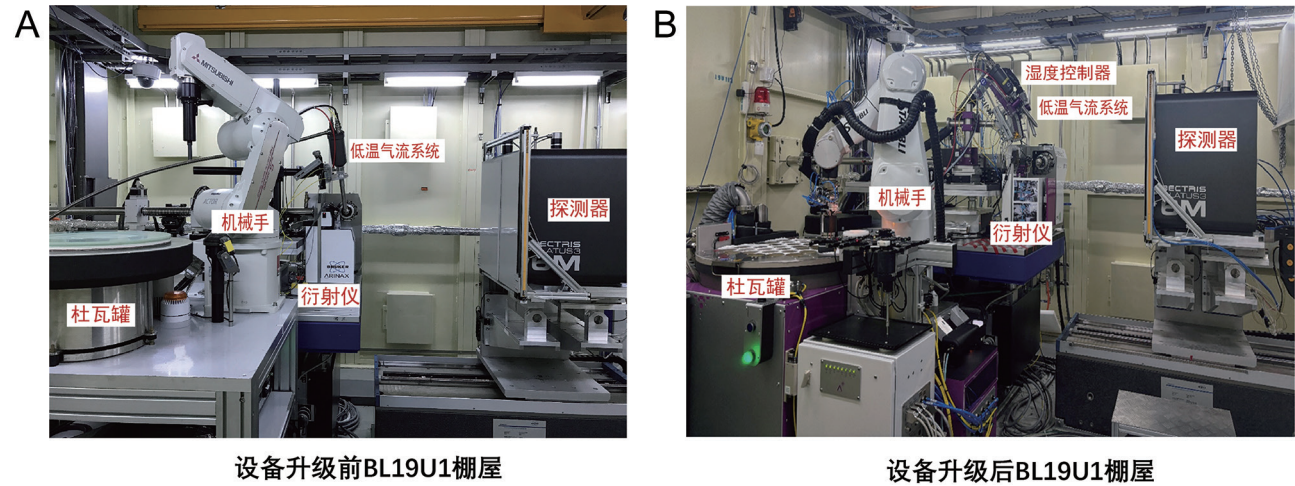


图2 BL19U1设备升级前后变化^[12]

数据采集装置,可在结晶板直接对晶体样品进行室温条件下的衍射实验。这一配置有效避免了传统样品转移及冷冻处理过程中可能导致的晶体损伤,提升了数据采集的可靠性与效率。特别是在基于晶体片段库筛选等高通量实验中,原位采集方式大幅缩短了筛选流程中晶体筛选与上样环节所需的时间。如图3所示,该装置已集成至MXCuBE3数据采集系统,用户可通过统一的操作界面实现便捷的实验控制与数据采集。

在软件方面,三条晶体学线站已由早期数据采集的BLU-ICE系统更新为MXCuBE3系统。作为基于网页的用户操作界面,MXCuBE3在交互体验、系统兼容性与扩展性方面均实现了显著提升,能够灵活适配多种实验设备,实现高效的数据采集与过程控制^[26,27]。该系统通过Exporter协议对MD2S衍

射仪的各项功能(如状态监测、自动对中、螺旋扫描采集等)进行控制,并采用Exporter或Socket协议对机械臂执行样品抓取与自动换样操作。通过EPICS协议进行探测器的控制。此外,MXCuBE3支持远程实验模式,研究人员可通过网络安全登录,实现远程样品加载、实验设置及数据采集,大幅提升了线站的使用灵活性和服务覆盖范围。系统还集成了样品信息的预导入功能,简化了用户操作流程,有效避免因手动输入导致的错误,提高了整体实验效率和数据采集的安全性(图4)。

为进一步简化用户数据处理流程,提升线站自动化程度,线站还部署了高效自动化数据处理管道,集成多套国内外主流晶体衍射数据处理软件,包括AutoPX^[28]、autoProc^[29]、xia2-XDS^[30]和xia2-DIALS^[31]等。用户在数据采集完成后,可实时获得

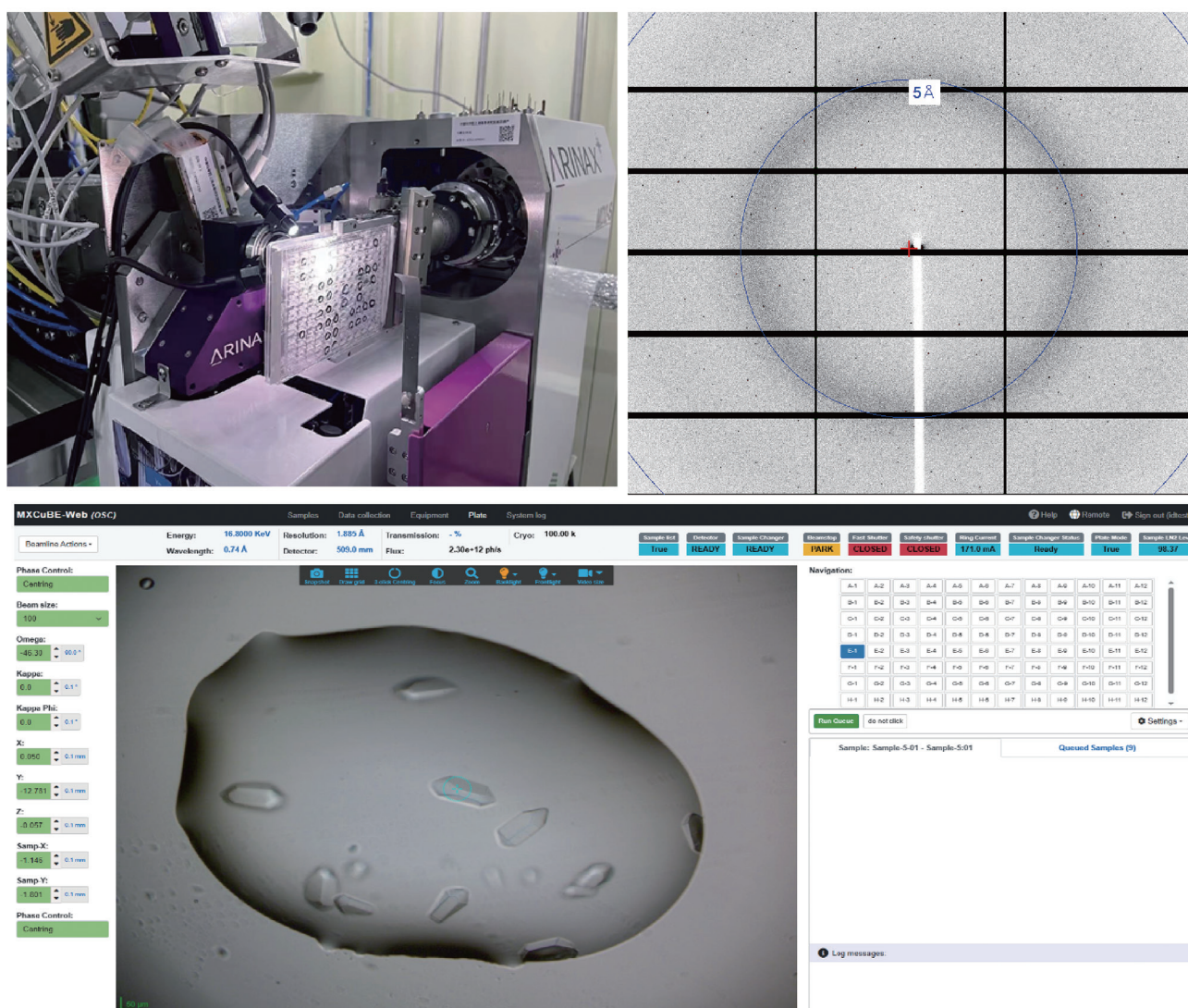
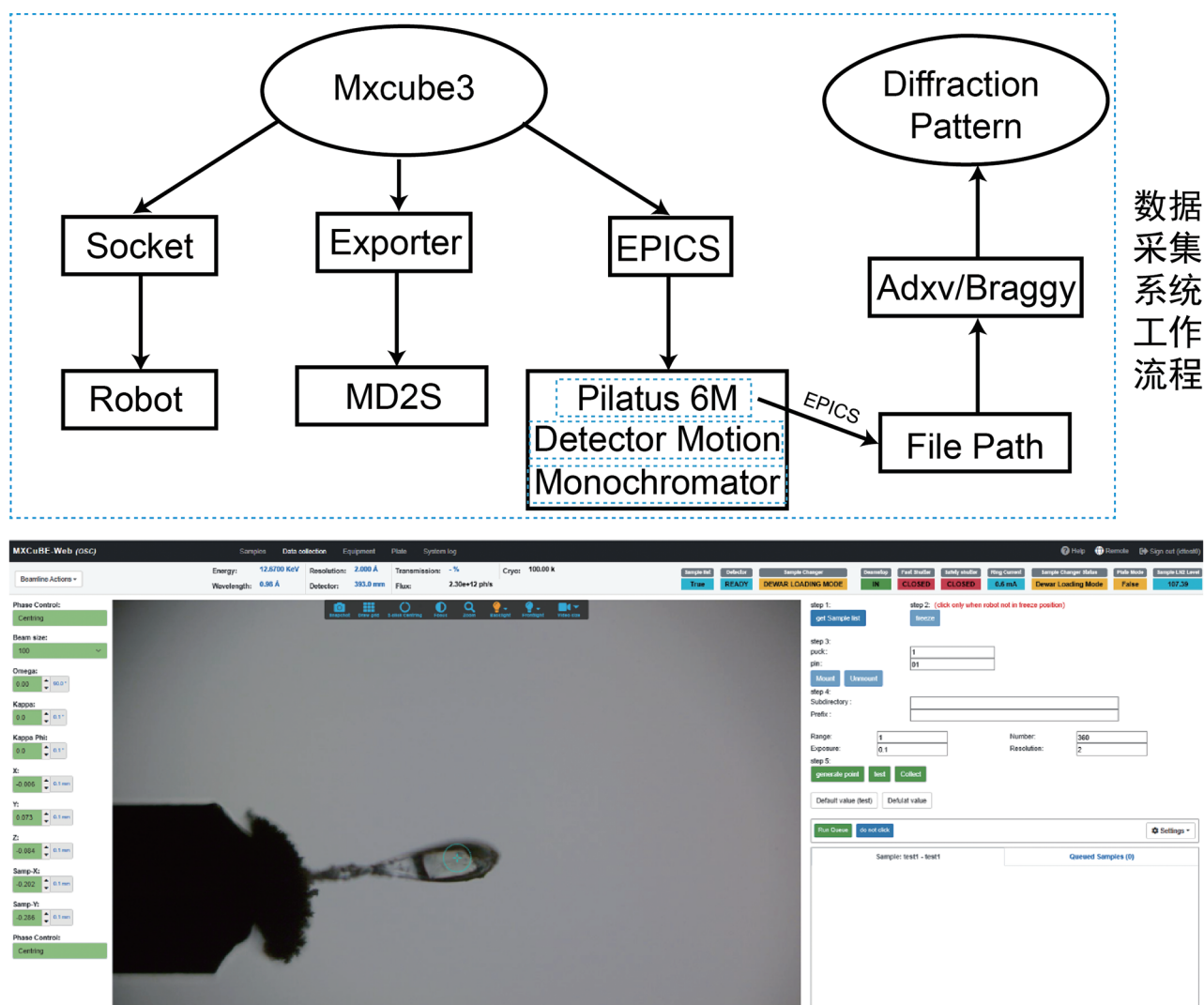


图3 基于结晶板的原位数据采集系统

图4 MXCuBE3系统集成与工作流程示意图^[20]

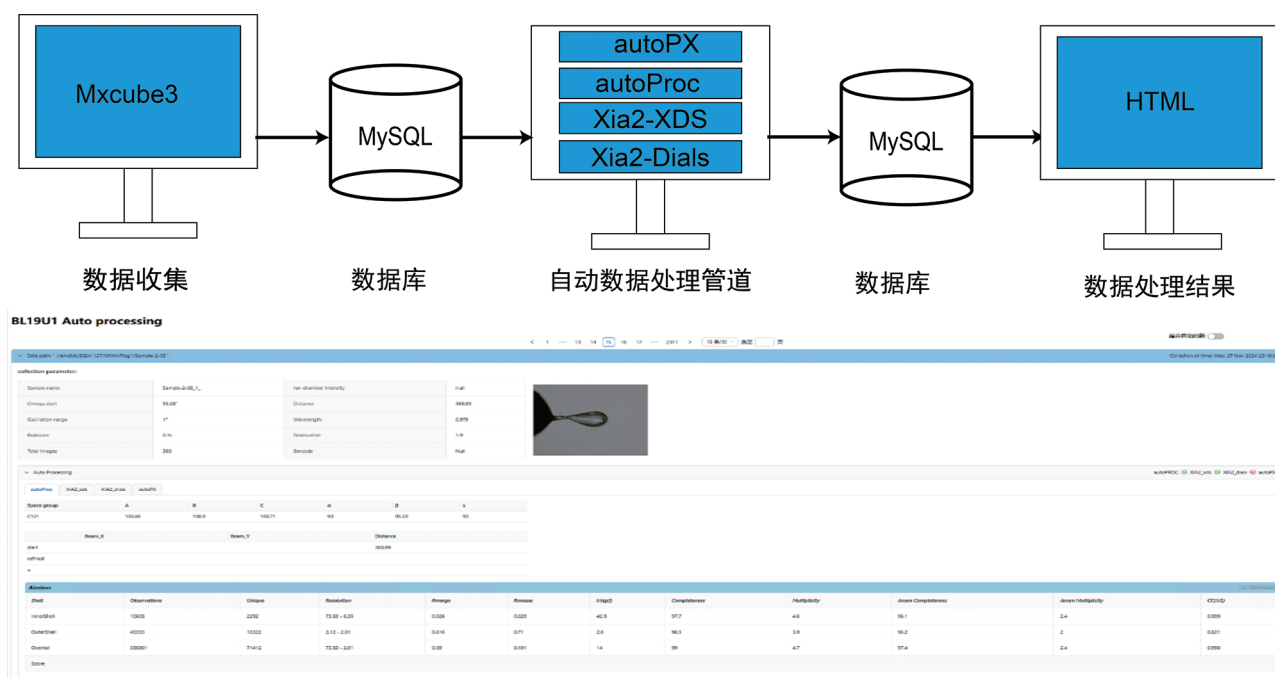
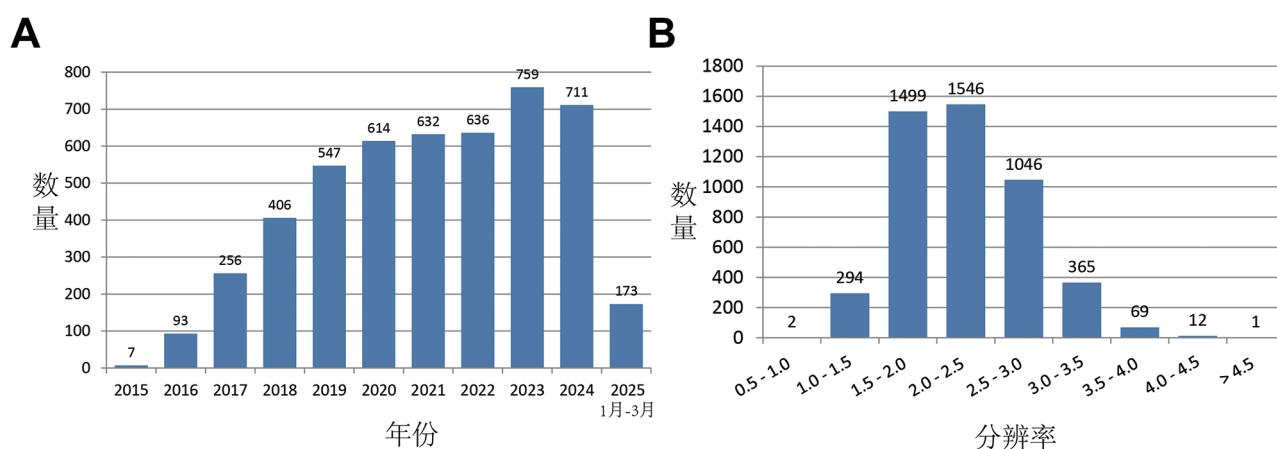
结构因子和数据质量分析日志,及时评估晶体质量,从而显著降低了实验门槛,提高了结构解析的整体效率与可重复性(图5)。

3 技术发展的应用与成果产出

目前,蛋白质设施晶体结构分析系统已成为我国生命科学领域高水平科研产出的重要支撑平台。截至2025年3月,平台用户累计在PDB数据库中发布的结构数据已达到4834个。图6A展示了2015年至2025年(截至3月)PDB数据库中发布的结构数量随年份的变化趋势。数据显示,自2015年以来,结构发布数量逐年增长,尤其自2018年起,增长势头尤为显著。2023年发布的结构数量达到峰值759个,2024年略有下降,仍维持在高水平(711个)。2025年截至3月已发布了173个结构,预计全年数据仍将保持高水平。这一增长趋势与线站性

能的不断提升密切相关,技术创新和设施优化为结构发布量的增加提供了强有力的支撑。图6B展示了这些结构的分辨率分布情况。从图中可见,绝大多数结构的分辨率集中在2.0~2.5 Å(1546个)、1.5~2.0 Å(1499个)以及2.5~3.0 Å(1046个)之间,表明当前收集的数据具有较高的结构质量。此外,还有一定数量的结构分布在3.0~3.5 Å(365个)范围内,而只有13个(占比0.27%)结构分辨率低于4.0 Å,进一步反映出整体解析的结构质量较高。这些高分辨率的结构对于当前人工智能(AI)在蛋白质结构预测、蛋白质-药物相互作用预测以及化合物的虚拟筛选等领域具有重要意义,为相关研究提供了坚实的结构数据基础。

其中,BL19U1线站蛋白质复合物晶体结构线站表现尤为突出。自开放运行以来,其用户提交至PDB数据库的新解出结构数量持续增长,相关科研

图5 自动数据处理流程以及处理结果展示^[20]

图A显示2015—2025年(截至3月)期间三条线站发布结构的年度分布趋势, 呈持续增长态势; 图B展示对应结构的分辨率分布情况, 主要集中在 1.5~2.5 Å 范围内, 表明整体结构解析质量较高。

图6 BL18U1、BL19U1和BL17B1三条蛋白质晶体学线站近年来在PDB数据库中发布的结构数量及分辨率分布统计图

成果产出亦屡创新高。根据 biosync 网站对 PDB 数据库中投递的结构进行分类分析, BL19U1 线站多次实现年度新解出结构数量在全球 140 余条同类线站中排名第一的佳绩, 充分展现了我国同步辐射设施在生物结构研究中的国际竞争力 (<https://biosync.rcsb.org>)。

在科研产出方面, 截至 2025 年 4 月, 依托该平台获取数据的用户累计共发表 SCI 论文 2 000 余篇, 其中影响因子大于 10 分的高水平论文超过

1 000 余篇, 在国际顶级期刊 *Cell*、*Nature*、*Science* 上发表论文 70 余篇, 显著提升了我国在结构生物学领域的国际影响力。基于晶体学线站所采集的高质量结构数据, 用户在病毒感染机制、重大疾病防控、药物开发、基因编辑机制、植物免疫研究、酶催化机制研究以及蛋白质结构设计等多个方向取得了具有广泛国际影响的科研成果。例如在埃博拉病毒入侵机制研究中, 高福院士团队解析了埃博拉病毒膜融合蛋白与宿主受体的识别模式, 阐明了其分

子入侵机制, 为抗病毒药物的结构设计提供了新靶点^[32]。该成果入选 2016 年度“中国医学科技十大新闻”及“中国生命科学十大进展”, 为埃博拉病毒疫情防控提供了理论依据。在抗结核药物开发方面, 饶子和院士团队率先解析了结核分枝杆菌关键靶标蛋白 MmpL3 与候选药物 SQ109 的复合物高分辨率结构, 首次揭示其药物作用机制与跨膜转运功能, 推动新一代抗结核药物的结构优化与研发^[33]。该研究成果入选 2019 年度“中国生命科学十大进展”。在新冠疫情暴发初期, 饶子和 / 杨海涛团队与合作者在国际上首次解析了新冠病毒主蛋白酶与抑制剂复合物的高分辨率晶体结构, 揭示了其抑制剂精确靶向机制, 为抗新冠药物研发奠定了坚实基础^[34]。该成果入选 2020 年度“中国生命科学十大进展”。此外, 晶体学平台还承担了多个新冠病毒关键靶标蛋白的快速解析任务, 助力筛选并验证 50 余种潜在药物分子, 其中来瑞特韦 (RAY1216)^[35] 和先诺欣 (simnoretvir)^[36] 等已经上市。在基因编辑调控领域, 黄志伟团队解析了 SpyCas9-sgRNA 与 Anti-CRISPR 蛋白 AcrIIA4 形成的复合物晶体结构, 首次揭示其抑制机制, 拓展了对 CRISPR-Cas9 与噬菌体防御系统“军备竞赛”的分子层面理解, 同时为开发可调控的基因编辑系统提供了结构基础^[37]。该成果入选 2017 年度“中国十大医学科技新闻”。在冠状病毒跨物种传播机制研究中, 通过结合晶体学、表面等离子共振 (SPR) 与病毒感染模型, 高福团队解析了 RaTG13 冠状病毒刺突蛋白与 ACE2 受体的结合机制, 揭示了其潜在的跨物种传播能力, 深化了对新型冠状病毒起源及传播规律的认识^[38], 该成果入选 2021 年度“中国生命科学十大进展”。在植物免疫领域, 多项研究基于晶体学展开, 如通过解析受体 RXEG1 (apo-RXEG1)、受体 - 配体识别 (RXEG1-XEG1) 和受体 - 配体 - 共受体复合物 (RXEG1-XEG1-BAK1) 等多种不同状态的结构, 结合生物化学和植物细胞生物学等手段进一步阐明了其发挥功能的具体机制^[39]。该研究阐明了细胞膜受体蛋白 RXEG1 识别病原菌核心致病因子 XEG1 激活植物免疫的作用机制, 首次揭示了细胞膜受体蛋白具有激活免疫活性和直接抑制致病因子 XEG1 酶活的双重免疫功能, 对认识、合理利用和精准改造植物免疫受体, 提高作物广谱抗性具有重要的指导意义, 相关成果于 2022 年在 *Nature* 杂志发表。还解析了植物免疫蛋白 PGIP 与病原菌果胶酶 PG 的复合物结构。该研究揭示了多聚半乳糖醛酸酶抑

制蛋白 PGIP 和多聚半乳糖醛酸酶 PG 互作激活植物免疫的全新分子机制, 该研究对于认识植物免疫系统、合理利用和精准改造植物免疫元件提高作物抗性具有重要的指导意义^[40]。相关研究成果于 2024 年在 *Science* 杂志上发表。在酶催化机制方面, 周佳海课题组及其合作者首次表征了自然界中催化 Alder-ene 反应的酶及其催化氧杂 Diels-Alder (DA) 反应的同源蛋白, 解析了这两类酶及其复合物的高分辨率晶体结构, 并基于结构信息和计算指导通过定点突变实现周环选择性的逆转, 阐明两类酶如何利用几乎相同的活性位点实现周环选择性的精准控制, 研究成果于 2020 年发表在 *Nature* 杂志上^[41]。在蛋白质结构设计领域, 晶体学线站作为设计验证和结构指导的重要平台, 发挥了关键支撑作用, 依托高分辨率晶体结构解析, 助力用户开发并验证蛋白质从头设计模型 SCUBA-D (SCUBA-diffusion)^[42], 相关研究成果于 2022 年发表在 *Nature* 杂志, 充分体现了晶体学技术在蛋白质设计领域的重要价值。

4 面临的挑战与未来发展方向

当前, X 射线晶体学虽已进入技术高度成熟阶段, 但在冷冻电镜 (Cryo-EM)、人工智能蛋白质结构预测 (如 AlphaFold) 等新兴技术迅速发展的背景下, 正面临多维度的竞争与挑战。在大型分子复合物、柔性蛋白以及天然难以结晶体系的研究中, Cryo-EM 凭借样品制备简便、无须结晶以及对大分子复杂体系的良好适配性, 展现出强大的解析能力^[43]。而基于深度学习的结构预测工具, 如 AlphaFold^[44], 在提供初步结构模型、补全缺失结构区段等方面, 也表现出越来越高的准确性和实用性。这些技术的崛起, 在一定程度上分担了 X 射线晶体学在部分领域的结构解析功能, 但并未撼动其在高分辨率结构测定中的核心地位。特别是在中小分子体系、配体 - 靶点复合物解析、小分子药物设计与优化、催化机制揭示等关键环节, X 射线晶体学依然拥有不可替代的技术优势。

随着自动化技术和信息化水平的不断提升, 生物大分子晶体学平台正加速向高通量和智能化方向发展。新一代蛋白质晶体学线站在样品换样、晶体定位与自动对中、快速数据采集及高效数据处理等关键环节不断推进技术创新, 实现了从样品上机到数据产出的高度自动化和流程标准化。这些技术进步显著提升了样品处理能力和实验通量, 满足了大规模结构筛选和高效结构解析的需求。在此基础上,

基于结构的片段化合物筛选等应用场景得到了广泛应用,特别是在新药研发早期阶段展现出重要价值。通过自动换样系统、精准的晶体对中技术、连续数据采集及自动化数据分析流程的整合,可以在短时间内完成大量配体结合状态的结构测定,为靶向化合物筛选与优化提供了强有力的结构依据。目前,国际多家同步辐射光源已在这一领域建立了相应的技术平台^[45]。例如,英国 Diamond 光源的 I04-1 线站依托 XChem 平台,搭建了高通量片段筛选体系,可在一周内完成上千个片段化合物的筛选实验,极大地加速了基于结构的药物发现进程^[45]。这一技术模式也为生物大分子晶体学在药物开发中的应用持续拓展了新路径。

此外,生物大分子晶体学也在不断推进与其他科学技术的融合发展。生物大分子晶体学平台与 Cryo-EM、NMR、质谱分析、AI 结构预测等多模态技术的深度协同正在加速形成,不同方法在空间、时间分辨率及适用样品类型上的互补优势,将进一步丰富大分子复合体结构解析和动态过程研究方法和体系。未来,围绕生物大分子结构与功能机制解析,基于多技术整合的全景式结构生物学将成为重要的发展方向,推动生命科学和新药研发不断向更深层次拓展。

[参 考 文 献]

- [1] Smyth MS, Martin JH. X ray crystallography. *Mol Pathol*, 2000, 53: 8-14
- [2] Garman EF. Developments in X-ray crystallographic structure determination of biological macromolecules. *Science*, 2014, 343: 1102-8
- [3] Maveyraud L, Mourey L. Protein X-ray crystallography and drug discovery. *Molecules*, 2020, 25: 1030
- [4] Burley SK, Bhikadiya C, Bi C, et al. RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Res*, 2023, 51: 488-508
- [5] Hu Y, Cheng K, He L, et al. NMR-based methods for protein analysis. *Anal Chem*, 2021, 93: 1866-79
- [6] Benjin X, Ling L. Developments, applications, and prospects of cryo-electron microscopy. *Protein Sci*, 2020, 29: 872-82
- [7] Grabowski M, Cooper DR, Brzezinski D, et al. Synchrotron radiation as a tool for macromolecular X-ray crystallography: a XXI century perspective. *Nucl Instrum Meth B*, 2021, 489: 30-40
- [8] Jiao Y, Bai ZH, Li X. Accelerator physics and technology of the fourth generation synchrotron radiation light source. *Physics*, 2024, 53: 8
- [9] Bingel-Erlenmeyer R, Olieric V, Grimshaw JPA, et al. SLS crystallization platform at beamline X06DA-A fully automated pipeline enabling X-ray diffraction screening. *Cryst Growth Des*, 2011, 11: 916-23
- [10] Hirata K, Ueno G, Nisawa A, et al. New micro-beam beamline at SPring-8, targeting at protein micro-crystallography. *Aip Conf Proc*, 2010, 1234: 901-4
- [11] Materlik G, Rayment T, Stuart DI. Diamond Light Source: status and perspectives. *Philos Trans A Math Phys Eng Sci*, 2015, 373: 20130161
- [12] Zhang WZ, Tang JC, Wang SS, et al. The protein complex crystallography beamline (BL19U1) at the Shanghai Synchrotron Radiation Facility. *Nucl Sci Tech*, 2019, 30: 170
- [13] Wang QS, Zhang KH, Cui Y, et al. Upgrade of macromolecular crystallography beamline BL17U1 at SSRF. *Nucl Sci Tech*, 2018, 29: 010102
- [14] Yao Z, Rogalinski J, Asimakopoulou EM, et al. New opportunities for time-resolved imaging using diffraction-limited storage rings. *J Synchrotron Radiat*, 2024, 31: 1299-307
- [15] Pellegrini C. X-ray free-electron lasers: from dreams to reality. *Phys Scripta*, 2016, T169: 014004
- [16] Chapman HN. X-ray free-electron lasers for the structure and dynamics of macromolecules. *Annu Rev Biochem*, 2019, 88: 35-58
- [17] Branden G, Neutze R. Advances and challenges in time-resolved macromolecular crystallography. *Science*, 2021, 373: eaba0954
- [18] Kaminski JW, Vera L, Stegmann DP, et al. Fast fragment-and compound-screening pipeline at the Swiss Light Source. *Acta Crystallogr D*, 2022, 78: 328-36
- [19] Owen RL, Juanhuix J, Fuchs M. Current advances in synchrotron radiation instrumentation for MX experiments. *Arch Biochem Biophys*, 2016, 602: 21-31
- [20] Xiao QJ, Wu TT, Bao KW, et al. Upgrade of crystallography beamline BL19U1 at the Shanghai Synchrotron Radiation Facility. *J Appl Crystallogr*, 2024, 57: 630-7
- [21] Song H, Qi JX, Xiao HX, et al. Avian-to-human receptor-binding adaptation by influenza A virus hemagglutinin H4. *Cell Rep*, 2017, 20: 1201-14
- [22] Zhang L, Zhao X, Hu JY, et al. PRPS2 enhances RNA m6A methylation by stimulating SAM synthesis through enzyme-dependent and independent mechanisms. *Nat Commun*, 2025, 16: 3966
- [23] Xiao QJ, Chen XY, Wang C, et al. Mechanistic insights into proton-coupled substrate translocation of nucleoside proton symporters. *J Biol Chem*, 2025, 301: 108357
- [24] Xiao Q, Sun B, Zhou Y, et al. Visualizing the nonlinear changes of a drug-proton antiporter from inward-open to occluded state. *Biochem Biophys Res Commun*, 2021, 534: 272-8
- [25] Peng X, Yang L, Liu Z, et al. Structural basis for recognition of antihistamine drug by human histamine receptor. *Nat Commun*, 2022, 13: 6105
- [26] Mueller U, Marjolein T, Jie N, et al. MXCuBE3: a new era of MX-beamline control begins. *Synchrotron*

- Radiation News, 2017, 30: 22-7
- [27] Oscarsson M, Beteva A, Flot D, et al. MXCuBE2: the dawn of MXCuBE collaboration. *J Synchrotron Radiat*, 2019, 26: 393-405
- [28] Wang L, Yun Y, Zhu Z, et al. AutoPX: a new software package to process X-ray diffraction data from biomacromolecular crystals. *Acta Crystallogr D Struct Biol*, 2022, 78: 890-902
- [29] Vonrhein C, Flensburg C, Keller P, et al. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D Biol Crystallogr*, 2011, 67: 293-302
- [30] Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr*, 2010, 66: 125-32
- [31] Winter G, Waterman DG, Parkhurst JM, et al. DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct Biol*, 2018, 74: 85-97
- [32] Wang H, Shi Y, Song J, et al. Ebola viral glycoprotein bound to its endosomal receptor niemann-pick C1. *Cell*, 2016, 164: 258-68
- [33] Zhang B, Li J, Yang X, et al. Crystal structures of membrane transporter MmpL3, an anti-TB drug target. *Cell*, 2019, 176: 636-48
- [34] Jin Z, Du X, Xu Y, et al. Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature*, 2020, 582: 289-93
- [35] Chen X, Huang X, Ma Q, et al. Preclinical evaluation of the SARS-CoV-2 M^{pro} inhibitor RAY1216 shows improved pharmacokinetics compared with nirmatrelvir. *Nat Microbiol*, 2024, 9: 1075-88
- [36] Jiang X, Su H, Shang W, et al. Structure-based development and preclinical evaluation of the SARS-CoV-2 3C-like protease inhibitor simnotrelvir. *Nat Commun*, 2023, 14: 6463
- [37] Dong D, Guo MH, Wang SH, et al. Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. *Nature*, 2017, 546: 436-9
- [38] Liu KF, Pan XQ, Li LJ, et al. Binding and molecular basis of the bat coronavirus RaTG13 virus to ACE2 in humans and other species. *Cell*, 2021, 184: 3438-51
- [39] Sun Y, Wang Y, Zhang X, et al. Plant receptor-like protein activation by a microbial glycoside hydrolase. *Nature*, 2022, 610: 335-42
- [40] Xiao Y, Sun G, Yu Q, et al. A plant mechanism of hijacking pathogen virulence factors to trigger innate immunity. *Science*, 2024, 383: 732-9
- [41] Ohashi M, Jamieson CS, Cai Y, et al. An enzymatic Alder-ene reaction. *Nature*, 2020, 586: 64-9
- [42] Huang B, Xu Y, Hu X, et al. A backbone-centred energy function of neural networks for protein design. *Nature*, 2022, 602: 523-8
- [43] Nakane T, Kotecha A, Sente A, et al. Single-particle cryo-EM at atomic resolution. *Nature*, 2020, 587: 152-6
- [44] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 2021, 596: 583-9
- [45] Douangamath A, Powell A, Fearon D, et al. Achieving efficient fragment screening at XChem facility at diamond light source. *J Vis Exp*, 2021, 171: e62414