



蛋白质分子机器的结构动态的多尺度多方法研究

朱月玲^{1,3}, 龚洲¹, 唐淳^{2*}

1. 中国科学院精密测量科学与技术创新研究院, 武汉 430071

2. 北京大学化学与分子工程学院, 北京 100871

3. 中国科学院大学, 北京 100049

*通讯作者, E-mail: Tang_Chun@pku.edu.cn

收稿日期: 2023-12-26; 接受日期: 2024-02-02; 网络版发表日期: 2024-03-11

国家重点研发计划重点专项(编号: 2018YFA0507700)资助项目

摘要 蛋白质构成了生物体内最重要的分子机器, 对蛋白质分子机器动态特性的了解、运作机制的揭示有助于实现对其功能的调控. 本文对蛋白质分子机器结构动态的研究方法进行了全面而系统的综述, 从空间分辨率、时间尺度、离体与原位等角度, 对不同方法的特点和应用进行了深入的分析, 指出了它们的优势和不足以及多方法整合联用的策略. 文章还对未来蛋白质分子机器结构动态研究的发展方向进行了展望, 提出了前瞻性的观点.

关键词 蛋白质机器, 动态结构, 多方法整合, 时空尺度

1 引言

蛋白质作为生物体内最重要的分子机器, 在生命活动的方方面面都扮演着重要的角色. 从组成细胞结构的膜蛋白, 到支持细胞内代谢活动的各种酶, 再到实现细胞信号转导的各种调控因子, 几乎所有的生物过程都与蛋白质密切相关. 作为一种分子机器, 蛋白质特定的三维空间结构决定了其特定的功能, 对蛋白质结构的解析一直是前沿科学研究的重点方向^[1]. 蛋白质的结构并不是静止不动的, 而是处于不断的动态变化过程中^[2]. 蛋白质的动态体现在两个层次, 包括蛋白质的内部的结构动态变化, 和蛋白质与蛋白质、蛋白质与核酸、蛋白质与小分子, 以及蛋白质聚集集体动

态变化.

蛋白质分子机器的结构动态变化跨越了从皮秒到秒甚至更长的十多个数量级的时间尺度, 以及从亚纳米到微米的多个空间尺度. 这些动态变化的类型包括了化学键的振动转动、氨基酸侧链运动、局部结构运动、结构域整体运动以及蛋白质机器内不同亚基之间的动态互作等多种方式^[3]. 正是蛋白质结构的动态才使得蛋白质能够起行使特定的生物学功能, 而蛋白质分子机器的动态不足或者异常可能导致疾病的发生^[4]. 因此, 从多个不同的时空尺度解码蛋白质分子机器的结构动态变化, 对于深入理解蛋白质发挥功能的分子机制, 设计各种干预调控和诊疗手段, 意义重大.

借助于实验技术的发展, 人们对蛋白质结构的研

引用格式: Zhu Y, Gong Z, Tang C. An integrative approach to understand structural dynamics of protein molecular machines across spatiotemporal scales. *Sci Sin Chim*, 2024, 54: 788–801, doi: 10.1360/SSC-2023-0247

究和理解在不断深入. X射线晶体衍射、核磁共振和冷冻电子显微镜并称为解析高分辨率蛋白质结构的“三驾马车”. 此外, 小角散射、单分子荧光和化学交联质谱等实验方法也逐渐被应用于蛋白质的结构和动态研究. 近年来, AlphaFold、RoseTTAFold等人工智能深度学习算法也广泛应用于蛋白质结构预测的相关研究^[5,6], 很多蛋白质的预测结构与结构生物学实验结果高度吻合^[7]. 目前, AlphaFold2已经对包括人在内的48个物种, 超过2亿个蛋白质的结构做出了预测, 并发布了这些结构的数据库(<https://alphafold.ebi.ac.uk/>)^[8]. 这些技术方法的不断迭代^[9], 有力的推动了蛋白质研究的发展.

但需要注意的是, 目前已解析的高分辨率蛋白质结构所对应的样品分子大部分为重组表达、体外纯化的蛋白质. 而蛋白质是在细胞内发挥具体的生物学功能, 细胞内复杂的环境因素, 例如盐浓度、pH值、拥挤环境和其它分子的相互作用等都会影响蛋白质的结构和动态变化, 进而影响蛋白质的功能^[10]. 因此, 在细胞原位环境下研究蛋白质的结构动态和功能机制, 对于我们理解生命的基本过程和本质具有重要的意义^[11]. 然而细胞成分的复杂性和多样性限制了大部分传统结构研究方法的应用.

综上所述, 作为一种结构复杂且精巧的分子机器, 蛋白质通过多时空尺度的结构动态变化发挥多样的生物学功能. 本文将围绕蛋白质分子机器的结构动态变化, 从静态到动态, 从低分辨率到高分辨率, 从离体到细胞原位等不同的角度综述不同的研究方法是如何应用于蛋白质分子机器的结构研究的, 并进一步讨论这些研究方法的最新进展, 面临的挑战以及未来的发展趋势(图1和表1).

2 蛋白质结构生物学研究——稳态与动态

获得高分辨率的蛋白质结构是研究和理解蛋白质功能的基础和前提. X射线晶体衍射(X-ray crystallography)是目前获得高分辨率蛋白质结构最主要的方法^[12,13]. 目前, PDB数据库中已收录了208931个蛋白质相关的结构(截至2023年12月19日统计, 包括蛋白质单体及蛋白质复合物). 这些结构中, 有超过85%的结构是通过X射线晶体衍射方法解析的. X射线晶体衍射解析蛋白质结构的基本原理是给蛋白质分子拍一张高清

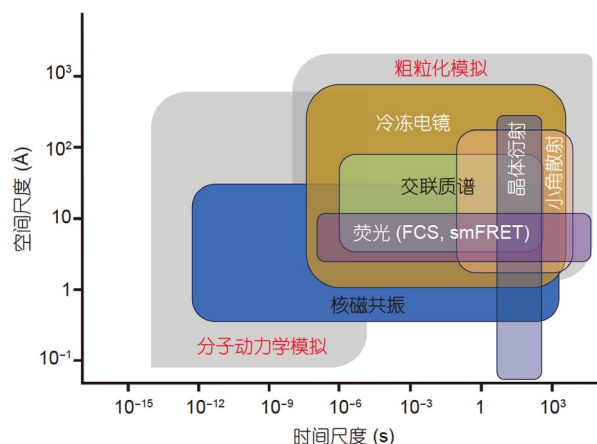


图 1 蛋白质分子机器只有通过结构的动态变化才能行使其特定的功能, 这种动态跨越了多个时间和空间尺度, 只有通过整合不同的实验手段和计算方法, 才能全面综合的了解蛋白质机器的运行机制, 调控和再造分子机器 (网络版彩图)

Figure 1 Protein molecular machines can only perform their specific functions through dynamic changes in structure. This dynamic behavior spans multiple time and spatial scales. Only by integrating different experimental and computational methods can we comprehensively understand the working mechanism, regulation, and regeneration of protein machines (color online).

晰度的照片. 因此, 获得高质量的蛋白质晶体是结构解析的关键步骤. 研究人员往往会通过筛选pH值、离子浓度、温度、及各种添加剂, 乃至对蛋白质分子做各种突变或截短来“削足适履”获得蛋白质晶体^[14]. 蛋白质分子在形成稳定晶体的同时, 往往会失去了其固有的运动特征. 因此, 蛋白质的晶体结构本质上是一种静态或稳态结构, 其反映的是蛋白质处于稳态时的结构特征. 近年来, 基于X射线自由电子激光(X-ray free electron laser)的时间分辨晶体学方法(time-resolved crystallography)可以在皮秒到纳秒的时间尺度内研究DNA损伤修复过程的蛋白质机器结构变化^[15,16]. 不过这类方法对硬件的要求高, 全球范围内只有少数几个设施可以开展相关研究. 值得一提的是, AlphaFold2等人工智能结构预测算法是以PDB数据库中的晶体结构作为主要训练素材的. 因此, AlphaFold2预测的结构主要反映的也是蛋白质的单一的静态结构信息^[17].

近年来, 冷冻电镜技术(Cryogenic Electron Microscopy, Cryo-EM)逐渐成为解析蛋白质结构的重要工具^[18]. 尤其是2013年以来, 电子直接探测器(Direct Detection Device, DDD)的发展使得冷冻电镜取得了革命性的进步, 打破了蛋白质侧链分辨率的屏障, 从而大幅

表 1 蛋白质结构研究的不同方法比较

Table 1 Comparison of different methods for protein structure and dynamic analysis

研究方法	空间分辨率	时间分辨率	优势	局限性
X射线晶体衍射	< 1 Å	无 ^{a)}	结构精度高, 不受体系大小限制	样品需要结晶
冷冻电镜	<10 Å	μs~s	适合研究超大复杂的蛋白质机器	不能研究小体系的分子
核磁共振	~1 Å	ps~s(或更长)	能在不同环境和细胞内研究蛋白质结构和动态	样品一般需要同位素标记
小角散射	~10 Å	μs~s	能在溶液状态下研究蛋白质的动态, 不需要标记或结晶	分辨率较低, 一般只能提供结构轮廓信息
交联质谱	~5~30 Å	> 1 s	能在不同环境和细胞内研究蛋白质结构和动态	~5~30 Å
荧光	~20~80 Å	ns~s(或更长)	能跟踪并检测单个分子的运动, 并且能应用于细胞内研究	单次标记获得的结构信息较少
分子模拟 ^{b)}	1~150 Å	ps~μs	能够研究动态变化和动态过程	需要实验验证

a) 样品需要结晶, 研究的是固定于晶格中结构状态; b) 这里仅指全原子的分子动力学模拟.

度的提高了冷冻电镜解析蛋白质结构的分辨率^[19]. 与X射线晶体衍射相比, Cryo-EM技术通过冷冻的方式固定了蛋白质不同的结构状态, 不需要事先对样品做结晶处理, 因而适合研究那些不容易结晶, 相对比较动态的大型体系或者蛋白质机器, 例如核糖体、剪切体, 蛋白酶体等^[20,21]. 通过优化样品的制备方式, 以及对图像的分类处理和分析, Cryo-EM可以研究蛋白质的动态结构. 但是这类研究往往依赖蛋白质本身的特性、样品制备的质量、以及照片的数量和衬度.

相比前述的技术方法, 核磁共振(Nuclear Magnetic Resonance, NMR)适合研究蛋白质的动态结构^[22,23]. 液体NMR技术能够在近生理的溶液环境开展研究, 而固体NMR针对膜蛋白、蛋白质相变聚集体的研究有独到的优势^[24,25]. 同时, NMR方法还能够研究和评估温度、离子浓度、pH值、拥挤环境、甚至是压力对蛋白质结构和动态变化的影响^[26]. 与X射线晶体衍射和冷冻电镜对蛋白质结构进行“拍照”不同, NMR对蛋白质结构的解析本质上是在进行“计算结构”. 通过不同的脉冲序列和检测方式, NMR能够获得不同维度的蛋白质结构以及动态变化的信息, 包括化学位移^[27]、距离^[28]、角度^[29]、弛豫速率^[30]、耦合常数^[31]等. 当蛋白质在溶液中发生结构的动态变化时, 这些参数也会随之变化. 通过化学交换相关的检测技术, 如ZZ-exchange, CPMG (Hahn and Carr–Purcell–Meiboom–Gill)自旋回波等, NMR脉冲序列能够针对不同时间尺度蛋白质结构动态变化开展研究^[32,33]. 此外, 通过弥散

加权核磁共振方法(Diffusion-Ordered Spectroscopy, DOSY)还可以检测蛋白质分子在溶液中的扩散速率, 从而获得蛋白质结构动态相关的分子尺寸信息^[34]. 核磁共振的参数是蛋白质分子机器的不同结构状态的参数的系综平均, 由于不同的参数的平均方式不尽相同, 将这些信息作为约束条件, 通过计算的方式就可以获得满足这些约束条件的蛋白质系综结构, 表征出蛋白质的动态变化^[35]. 例如, 近年发展起来的顺磁核磁增强技术, 通过引入顺磁标记实现对长程效应的检测, 检测的核磁信号和距离的负六次方的平均成正比, 因此能够非常灵敏的捕获顺磁标记与被测原子核距离变短的蛋白质的动态结构和瞬态结构^[36]. 顺磁核磁共振实验结合相关的计算方法可以检测丰度极低、持续时间极短的蛋白质动态结构和动态互作. 我们团队多年来利用核磁解析和表征蛋白质结构动态, 发展顺磁核磁共振方法首次发现了泛素蛋白存在弱的非共价二聚相互作用, 为共价泛素蛋白的形成提供了基础^[37]. 我们团队还解析了解离常数仅为25 mmol L⁻¹、存在时间小于1毫秒的极弱瞬态蛋白质相互作用复合体(图2). 这样, 当细胞内蛋白质的有效浓度很高时, 这种极弱相互作用使得不同蛋白质之间的信号得以快速传递^[38]. 除了蛋白质分子间的动态互作以外, 我们团队还通过设计刚性的顺磁探针, 检测和分析了细微的蛋白质环区的运动, 其平均运动幅度仅为0.6 Å, 并且证明这种细微的运动与配体的解离过程相关^[39].

相对于有确定结构的蛋白质分子机器, 固有无序

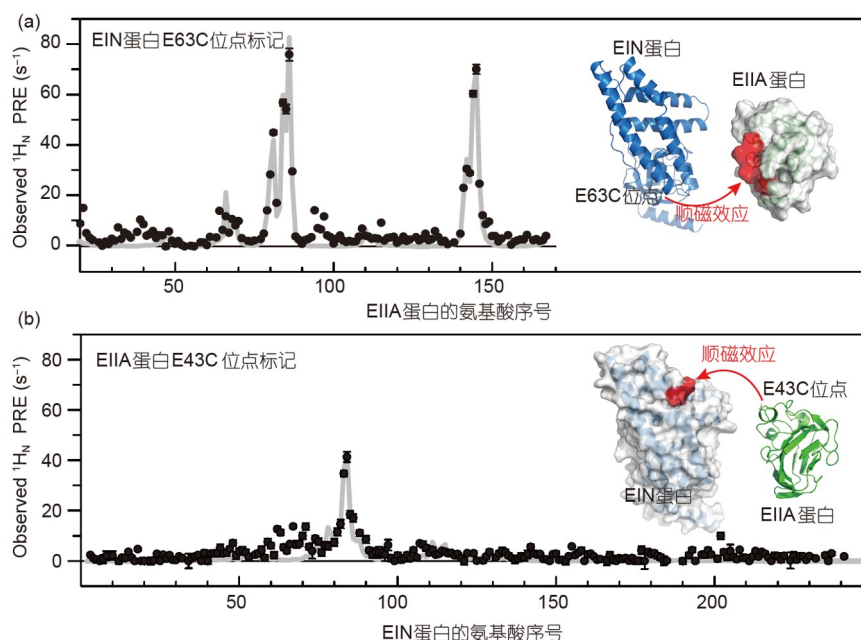


图 2 通过对蛋白质进行定点的顺磁标记和蛋白质之间所引起的顺磁弛豫增强分析, 可以捕获并表征极弱的蛋白质相互作用^[38]. 这里我们将顺磁探针定点标记在(a)EIN和(b)EIIA蛋白, 与之相互作用的EIIA和EIN蛋白的特定氨基酸残基就会被“点亮”, 表现为所检测到的PRE核磁值的增高 (网络版彩图)

Figure 2 By labeling proteins with paramagnetic probes and analyzing the paramagnetic relaxation enhancement (PRE) caused by protein-protein interactions, it is possible to capture and characterize ultra-weak interactions. Here, we site-specifically introduced paramagnetic probes on (a) EIN and (b) EIIA proteins. The specific amino acid residues of EIIA and EIN proteins that interact with them will be “lit up”, manifested as an increase in the PRE nuclear magnetic resonance (NMR) value detected (color online).

蛋白(Intrinsically Disordered Protein, IDP)是一类结构高度动态、缺乏稳定三级结构的蛋白质, IDP和IDR (Intrinsically Disordered Region, IDR, 即蛋白质中的固有无序区域)占整个人类蛋白质组的30-40%, 并在转录、翻译、信号转导等重要的生理过程中扮演着核心节点的角色^[40,41]. 因此, 表征IDP/IDR的结构, 揭示它们的动态变化及其与功能的关系具有重要的意义^[42,43]. IDP/IDR由于结构高度动态, 在X射线晶体衍射或冷冻电镜研究中要么无法解析, 要么在解析的结构中是缺失的. NMR是目前唯一能够解析IDP/IDR高分辨率结构的实验方法^[44]. NMR不仅能够解析IDP/IDR的结构, 动态变化和相互作用^[45], 近年来, NMR还用于分析相分离凝聚体中的IDP/IDR的结构状态, 揭示IDP/IDR自组装聚集的微观机制^[34,46].

除了实验方法以外, 以分子动力学模拟为代表的计算模拟方法也是研究蛋白质的结构, 分析其动态变化规律的重要手段. 分子动力学模拟通常从高分辨率的静态结构出发, 通过分子力场来表征结构的内在属性, 并在经典力学的约束下模拟并记录蛋白质的运动

状态和运动过程. 通过对蛋白质不同状态下的构象、相互作用、能量等的分析, 探究蛋白质动态变化的规律^[47]. 随着力场参数、算法和计算机硬件技术的不断发展, 尤其是GPU计算卡的发展应用, 分子动力学模拟的时间和空间尺度已经能够在一定程度上与实验方法匹配, Shaw团队还开发了专门用于分子模拟的硬件设备Anton^[48]. 此外, 粗粒化的模拟, 即忽略氨基酸残基、结构域甚至是整个分子的细节, 可以模拟再现更大的体系和更长的时长. 分子动力学模拟方法也与各种实验方法结合, 用来表征蛋白质结构的动态变化, 重构动态变化的过程, 还能进一步开展相关药物的设计和筛选工作^[49,50]. 尤其是对于IDP来说, 由于其体系高度动态, 将分子动力学模拟与NMR等实验方法结合表征IDP的结构变化, 以及IDP之间的动态相互作用是当今的一个重要发展动向^[51,52]. 与此同时, 各种人工智能算法的发展应用也进一步优化了IDP的结构研究策略^[53,54]. 总之, 实验和计算模拟方法可以从不同的角度来表征和研究静态到动态的蛋白质分子机器结构. 蛋白质的静态结构, 尤其是晶体结构往往是高分辨率

的, 而对于蛋白质的动态结构来说, 则往往需要采取一种折中的方法综合考虑信息量和结构分辨率的问题。

3 蛋白质动态结构的研究——低与高分辨率

前面提到了用于蛋白质结构解析的三种实验方法, 这些方法虽然都能够获得高分辨率的蛋白质结构信息, 但是在应用范围上往往会受到限制, 或者需要投入高昂的人力物力来获得实验样品。X射线晶体衍射需要优化条件获得高质量的晶体; 冷冻电镜需要制备高均一度的样品; 而NMR实验需要对实验样品做稳定同位素标记, 并由于谱峰重叠的问题, NMR方法大多数情况下只能解析分子量小于40 kDa的蛋白质结构^[55]。

另一方面, 在很多情况下, 我们往往并不需要从头来解析蛋白质的高分辨率结构。例如, 已经解析了包含多个结构域的蛋白质高分辨率结构, 我们更关注的是不同结构域的动态变化; 或者分别已知两个蛋白质单体的高分辨率结构, 我们想研究它们如何动态互作、所形成的复合物的结构。在这些情况下, 往往只需要一些低分辨率的结构信息, 结合高分辨率的局部结构, 就可能获得蛋白质的整体结构和动态变化。这些能够用于研究蛋白质结构和动态变化的低分辨率方法包括小角散射, 单分子荧光以及交联质谱等。以下分别对这些方法在蛋白质结构和动态研究方面的发展和应用展开讨论。

小角散射方法主要包括小角X射线散射(Small Angle X-ray Scattering, SAXS)和小角中子散射(Small Angle Neutron Scattering, SANS)。这两种方法在应用时都不需要对样品结晶, 也不需要任何修饰或标记。SAXS和SANS都能直接在溶液状态下检测, 获得蛋白质的整体形状轮廓, 并能根据散射信号分析蛋白质的折叠状态和尺寸大小等信息^[56]。由于溶液中所有的溶剂和溶质都会散射, 因此, 对散射背景信号的正确扣除是这些实验的关键步骤。由于中子散射对不同的原子核具有选择性, 在SANS实验中, 可以使用不同的样品制备策略, 将样品中的不同组分作为背景过滤掉。例如, 通过对多结构域蛋白中不同的结构域分别做氘代表达, 能够解析不同结构域对底物RNA的识别方式以及协同作用^[57]; 通过对目标蛋白质选择性氘代, 就能够研究大分子拥挤环境对蛋白质复合体的影响和调控

作用^[58]。SANS方法在这些应用中有效的实现了对背景信号(RNA, 大分子拥挤剂等)的过滤, 从而能够研究样品中的部分蛋白质结构信息。SAXS和SANS方法还能通过对散射信号的拟合与拆分来重现蛋白质的动态结构^[59]。目前有可供实验室安装使用的SAXS设备, 同时, 基于同步辐射光源的SAXS实验能够获得信噪比更好的散射数据, 所需的采样时间也更短。相比之下, SANS实验由于需要使用中子源, 目前在我国境内只有四川绵阳的中国绵阳研究堆(CMRR)和广东东莞的中国散列中子源(CSNS)等少数地方可以进行相关实验。

此外, 单分子荧光共振能量转移(single-molecule Fluorescence Resonance Energy Transfer, 即smFRET)被广泛应用于蛋白质的结构和动态变化研究。FRET需要在蛋白质上标记两个具有不同吸收波长的荧光基团(荧光探针或者荧光蛋白), 分别作为荧光的供体(donor)和受体(acceptor)。当特定波长的荧光激发donor时, donor会将一部分能量传递给acceptor, 激发acceptor发出荧光。通过检测和比较donor激发光强度和acceptor发射光强度就可以获得两个荧光探针之间的距离信息, 用于重构蛋白质的结构^[60]。smFRET指的是在单分子水平检测荧光共振能量转移效率的实验方法。在具体的实现方式上, smFRET还可以分为基于自由扩散的共聚焦显微镜检测方法(free diffusion with confocal microscopy)和基于固定样品的全内反射检测方法(Total Internal Reflection Fluorescence Microscopy, 即TIRF)^[61]。基于自由扩散的共聚焦显微镜检测是基于荧光标记的分子在溶液中的自由扩散, 使用共聚焦显微镜联合雪崩二极管对到达的光子进行计数。当标记的蛋白质分子通过激光聚焦点时, 荧光会被激发并被检测到。由于检测腔体积很小(在飞升量级)、样品浓度在皮摩尔量级, 因此在单位时间内只有最多一个分子通过激光检测点, 从而实现了单分子水平的测量。TIRF方法的原理是激光束从玻璃和溶液之间的界面入射, 当光束穿过界面并到达玻璃的上方时, 由于全内反射, 光束在玻璃和溶液的界面上形成一个薄的激发区域, 因此只有附近的荧光标记分子在这个激发区域内能被激发, 并且被宽场的CCD或CMOS相机所记录。TIRF方法需要将检测的蛋白质分子固定吸附在玻璃的表面, 因此除了荧光探针的标记以外, TIRF方法还需要额外标记生物素等分子来实现蛋白质分子的固

定^[62]。相比之下, 基于自由扩散的共聚焦检测方法在样品制备上则更加简单灵活, 并且能够检测到比TIRF更快的蛋白质结构动态。

smFRET技术对20~80 Å范围的距离变化比较敏感, 因此, 对于蛋白质的动态结构研究, smFRET可以在一个较宽范围内提供长程的距离约束^[63]。由于smFRET一次只标记一对荧光位点, 因此, 每一次实验仅能获得一组蛋白质结构的距离信息。虽然可以设计不同的荧光标记位点来获得多组不同的距离数据, 但总体而言, smFRET是一种低分辨率的结构研究方法。不过, smFRET能够检测到不同时间尺度蛋白质结构的动态变化。通过对荧光信号累计的统计分析, 还可以拟合获得蛋白质结构状态的比例信息。因此, smFRET能够与其他的结构研究方法, 尤其是NMR方法很好的结合, 在获得局部高分辨率结构信息的基础上叠加全局的距离变化信息, 从而更全面的表征蛋白质的动态结构^[64-66]。

交联质谱(Cross-linking combined with Mass Spectrometry, 即XL-MS)也是通过对距离的检测和分析来研究蛋白质结构的重要方法。XL-MS使用小分子交联剂与蛋白质发生交联反应形成共价连接, 然后通过酶切和高分辨质谱获得被交联的氨基酸位点信息。由于

小分子交联剂的结构和长度是已知的, 因此XL-MS技术可以提供被交联的氨基酸之间的距离信息^[67]。这些距离信息通常会与已知的PDB结构进行比对, 比对如果出现了所谓“过长交联对”, 也就是交联位点对应PDB结构上的距离超过了交联剂的最大臂长, 就提示蛋白质存在结构的动态变化^[68,69]。图3就展示了这种评判蛋白质是否存在结构动态的逻辑。

蛋白质分子上能够发生交联反应的通常是具有活性侧链的氨基酸(如Lysine, Glutamate, Aspartate, Cysteine), 因此, 交联获得的距离信息是比较“稀疏”的。另一方面, 两个氨基酸之间形成交联, 只能说明两个氨基酸之间的距离小于交联剂的最大臂长。通常使用的化学交联剂的最大臂长在20 Å左右, 因此这些被交联氨基酸之间的距离可以在小于20 Å的范围内变化, 这就限制了交联距离约束的精度。近期研究表明, 不同类型的交联剂达到其最大臂展长度时所处的状态出现的概率很低。在多数情况下, 交联剂在溶液中呈现出“打开”和“关闭”两种不同的构象状态, 并且在这两种不同状态之间快速切换, 而不是等概率的覆盖20 Å内的所有距离。这两种状态的概率分布也更有可能和蛋白质上被交联到的氨基酸的实际距离相吻合^[70], 对于序列上非临近的氨基酸更有可能是出现在距离较长的第二

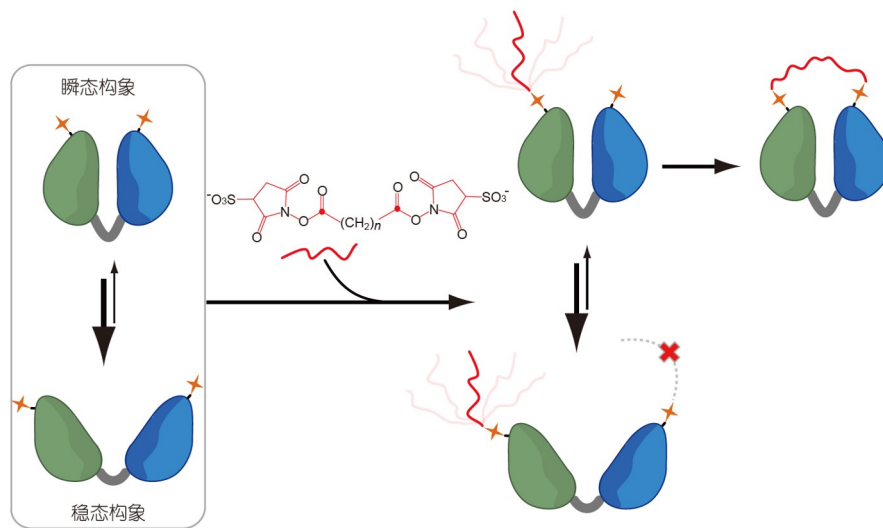


图 3 通过对交联反应的高分辨质谱和搜库匹配分析, 过长交联(over-length crosslink)往往提示蛋白质分子机器有结构动态运动, 即被交联的结构部分发生了由稳态构象转变成瞬态构象的转变^[68]。如图所示, BS3或BS2G交联剂首先与蛋白的一个结构域反应, 而第二步的交联反应是否能实现, 取决于两个结构域的开关运动(网络版彩图)

Figure 3 Over-length crosslinks often indicate structural dynamic motion of protein molecular machines, *i.e.*, the structural parts being crosslinked undergo a transition from stable conformation to transient conformation^[68]. As shown in the figure, the BS3 or BS2G crosslinker first reacts with one structural domain of the protein, and whether the second crosslinking reaction can be realized depends on the opening and closing motion between the two structural domains (color online).

个高斯分布内。因此, Tang研究团队基于交联剂的结构特点以及交联剂只能在蛋白质表面“爬行”的特点, 设计了一种新的距离函数, 将距离约束缩小到6 Å作用, 极大的提高XL-MS距离约束的精度^[71]。

XL-MS虽然只能获得低分辨率的、较稀疏的距离信息, 但可以通过与其它技术方法如X射线晶体衍射, 冷冻电镜的整合, 在获得局部高分辨率结构的基础上实现对大体系蛋白质机器的高分辨率结构表征, 揭示蛋白质机器发挥功能的动态变化^[72,73]。XL-MS技术不需要事先对样品进行修饰标记, 对样品的需求量少(通常在纳克或皮摩尔量级)。化学交联剂的长度和反应基团还可以根据研究的需要灵活设计, 并且还能在细胞内开展原位蛋白质分子机器的结构动态的研究。

以上这些不同的研究方法各具优势, 各有其适用的范围和应用的局限, 因此, 可以综合运用不同方法来对蛋白质分子机器的结构和动态变化开展研究, 优势互补、相互印证。XL-MS可以与晶体衍射或冷冻电镜结合, 用来研究复杂的蛋白质机器的动态结构^[74,75]。SAXS, XL-MS和smFRET这些低分辨率的研究方法也可以有机整合, 实现对多结构域蛋白动态结构的准确表征^[76]。将NMR与smFRET整合运用, 可以表征包括IDP在内的各种蛋白质分子的动态结构, 并且可以实现系统水平和单分子水平结构信息的相互验证^[64]。

4 蛋白质动态结构的研究——体外提纯与原位细胞体系

在细胞中, 蛋白质通过构象变化和相互作用执行着多样的生物功能。但目前对蛋白质的结构研究大多是在体外进行的, 也就是在经过优化的实验条件下对体外纯化的样品进行结构研究, 这与细胞内的生理环境存在很大的区别^[77]。细胞内部是一个拥挤的环境, 含有高达400克/升的各类大分子, 这就对蛋白质产生了显著的体积排斥效应。同时, 蛋白质还会与周围的细胞组分之间发生各种或强或弱的相互作用。细胞环境是否调控、如何调控蛋白质分子机器的结构动态是一个特别有意思也是具有挑战性的问题。因此, 发展方法以在原位、细胞内研究蛋白质结构及其动态变化, 意义重大。

自2001年Dötsch研究团队在大肠杆菌活细胞中首

次获得了NmerA蛋白的二维高分辨率NMR谱图以来, 细胞内NMR已经成为一种直接检测细胞水平蛋白质高分辨率结构的强大技术^[78,79]。除了常规的NMR谱图采集, 各种在体外广泛应用的顺磁核磁共振方法如顺磁弛豫增强、伪接触化学位移也被应用于各种动物细胞内的蛋白质结构研究, 实现了对诸如 α 突触核蛋白^[80]、GB1蛋白等在细胞内的结构^[81]的解析。由于NMR采样的周期较长, 非均一采样技术^[82]、培养基置换技术等方法被发展和应用到细胞内的NMR研究^[83], 以缩短核磁实验时间, 维持NMR实验期间细胞的活性状态。

还需要注意的是, 富含各种大分子和小分子的细胞内复杂环境会在NMR检测中产生很强的背景信号, 干扰对目标蛋白质的采样和分析。由于¹⁹F原子在天然生物系统中并不存在, 且¹⁹F原子核的磁共振信号对其周围的化学环境非常敏感, 呈现出很宽的化学位移分布范围。因此, ¹⁹F标记成为细胞内NMR研究蛋白质结构的理想方法^[84-86]。国内外多个研究团队开发了多种基于¹⁹F原子的探针和检测方法, 如Li研究团队^[87]开发了一种具有高反应选择性的水溶性探针wPSP-6F。探针上6个等效的氟原子和快速旋转的三氟甲基显著提高了NMR实验的信噪比, 使其能够在浓度低至10 μ M的情况下研究人源细胞中的蛋白质结构; 又如Luchinat研究团队^[88]开发了一种介质切换策略, 通过将氟化氨基酸直接导入到细胞内的蛋白质中, 从而无需进行蛋白质递送就可以实现在细胞内的¹⁹F直接检测。在这些标记方法的加持下, ¹⁹F-NMR检测方法已被广泛应用于研究细胞内的蛋白质结构和蛋白质-蛋白质相互作用, 揭示了细胞环境中的各种特异性和非特异性相互作用对蛋白质结构和动态变化的影响和调控^[89]。

相比于NMR研究细胞内的单个蛋白质(或蛋白质复合物)的结构和动态变化, XL-MS方法可以实现高通量的对细胞内整个蛋白质组进行结构分析^[67,90], 其流程如图4所示。通过开发具有增强膜通透性的新型交联剂^[91]、高效富集交联肽段的方法^[92,93]、发展可切割可碎裂交联剂(cleavable cross-linkers)的方法^[94], 以及检索算法的不断发展迭代^[95], 原位交联质谱技术能够获得细胞内几千个蛋白质的几万对交联数据, 为研究细胞内蛋白质结构、动态及相互作用提供了详细而全面的视角。为了尽可能降低交联剂对细胞状态的侵扰,

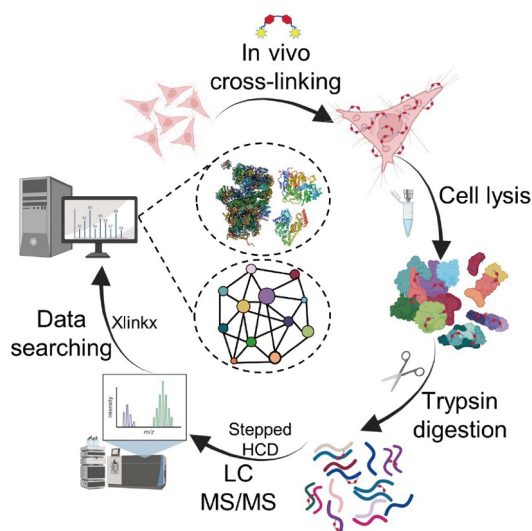


图 4 细胞内交联和质谱鉴定的基本流程^[98] (网络版彩图)

Figure 4 A schematic diagram for the workflow of *in-situ* cross-linking mass spectrometry [98] (color online).

提高交联的效率, Liu研究团队^[96]结合膜通透性和IMAC富集技术, 可以在30分钟内进行细胞内交联、最小化对细胞的干扰. 2022年, Zhang研究团队^[97,98]进一步发展BSP交联剂, 将细胞内交联时间缩短到5分钟. 2023年, 具有葡萄糖二糖骨架的新型TDS交联剂被开发应用, 这种交联剂无需辅助溶剂, 在捕获细胞内蛋白质结构和动态信息的同时, 最大限度的保持了细胞活力和蛋白质组的完整性.

在高通量高反应速率的基础上, 近年来还出现了多种亚细胞分离辅助的化学交联方法. 通过亚细胞器蛋白质标记辅助策略^[99], 开发定向交联剂递送-质谱等方法^[100], 有效的实现了针对线粒体等亚细胞器内蛋白质结构, 动态变化和相互作用的交联表征. 除了细胞水平以外, XL-MS方法还能直接在心肌组织^[101]、海马体^[102]、突触小泡膜^[103]等组织内开展针对蛋白质分子机器的活性和非活性构象的研究. 尤其需要指出的是, Zhang等人^[104]通过对细胞内交联位点的分析, 结合AlphaFold2对蛋白质组的预测精度打分, 实现了对蛋白质柔性动态区域的判定. 总之, 综合利用XL-MS技术方法, 可以获得关于蛋白质结构与功能在原位水平的完整图景, 有助于揭示蛋白质分子机器的运行机制.

除了NMR和XL-MS技术外, 各种细胞水平的成像方法如FRET、BRET(Bioluminescence Resonance Energy Transfer)也能有效的应用于细胞内的蛋白质结构

研究^[105]. 传统的基于荧光强度的FRET检测方法在应用于细胞内蛋白质研究时, 容易受蛋白质表达水平以及细胞内背景荧光信号的干扰. 而荧光寿命是荧光基团的本征属性, 不易受到周围环境的影响, 也和细胞内蛋白质的表达水平无关, 荧光寿命的变化也同样反映了荧光基团间的距离信息, 因此是一种理想的研究细胞内蛋白质结构与动态变化的方法^[106]. Zhang研究团队发展了基于荧光寿命的荧光共振能量转移检测方法(Fluorescence Lifetime Imaging Microscopy-Fluorescence Resonance Energy Transfer, FLIM-FRET). 该方法利用切割荧光蛋白技术实现了细胞水平不同位点的特异性荧光标记^[107], 并通过FLIM-FRET检测获得了多组分子间的距离信息, 在细胞水平解析了膜蛋白的同源二聚体结构.

5 蛋白质动态结构研究中的时间尺度

以上我们对蛋白质分子机器的结构研究, 从静态到动态、从体外到细胞内综述了不同研究方法的实现方式以及最新进展. 同时, 我们还从结构的分辨率的角度分析和比较了不同研究的特点以及它们的整合应用. 要全面的描述蛋白质分子机器的结构动态变化, 在空间三维结构的基础上, 还需要结合时间尺度, 这包括运动的速率、不同结构状态之间的交换速率及不同结构状态的寿命等.

NMR技术能对不同时间尺度内的蛋白质动态变化做出精细的刻画^[108]. 例如, 通过测量原子核的弛豫速率(relaxation rate)的变化就可以分析氨基酸是否在ps~ns尺度是否存在动态变化; 通过弛豫弥散(relaxation dispersion)和化学交换饱和转移(Cheical Exchange Saturation Transfer, CEST)可以分析蛋白质在μs~ms时间尺度的动态变化, 还能通过数据拟合获得交换速率以及不同构象的信息^[109~111]. 对于在ms~s时间尺度的交换, 常规的二维核磁谱图也可以用于分析蛋白质不同的结构状态^[112].

smFRET检测手段也可实现对不同时间尺度蛋白质动态结构的表征^[113]. 由于检测原理的不同, 基于自由扩散的共聚焦显微镜检测方法能够检测ns~ms尺度的结构动态变化, 因而常用来研究蛋白质分子机器的快速构象转变, 结构域之间的开合运动等^[114~116]. 相比之下, TIRF-FRET方法则涵盖了从ms~s甚至更长的时间

尺度,因而可以用来研究蛋白质的折叠,蛋白质-蛋白质相互作用等较慢的动态变化^[117]。此外,其它的基于荧光的实验方法,例如荧光相关光谱(Fluorescence Correlation Spectroscopy, FCS)等也能够不同的时间维度上获得诸如扩散系数,结合强度,聚集状态等蛋白质的动态信息^[118]。与smFRET相比, FCS只需要标记一个荧光探针,在实验操作上更为灵活方便,但不能获得距离这类直接反映蛋白质结构的特征信息,通常需要与其他方法联合使用。

由于交联过程中的化学反应需要较长的时间(至少是分钟级别),因此XL-MS通常认为表征的是处于热力学平衡状态下的蛋白质结构动态系综,而不考虑具体的时间尺度。近期, Lei和Dong研究团队^[119,120]开发了一种基于邻苯二甲醛反应基团的新型化学交联剂,能够将交联反应的时间缩短到秒量级。同时,作者结合分子动力学模拟,发现这种交联剂可以捕获到存在时间短的蛋白质去折叠过程中的动态中间体结构以及和蛋白质之间的动态相互作用。除了化学交联剂以外,近年来快速发展的各种光交联剂也广泛应用于蛋白质的结构研究。光交联技术是将光敏基团(如芳香族酮类,噻吩类,吡啶类等)作为交联反应基团,在受到特定波长的光照以后,这些光敏基团会释放出活化自由基或激发态物质,从而引发共价键的形成^[121,122]。与化学交联相比,光交联普遍具有更快的反应速率,并且触发时间可控,因此光交联方法在研究蛋白质分子机器的动态变化有很大的潜力。

综上所述,整合互补的谱学手段可以表征分子机器在不同时间尺度的结构动态,如图5所示。

6 总结与展望

本论文围绕蛋白质分子机器的结构动态的多尺度研究,从空间分辨率到时间分辨率,从离体到原位、细胞内等多个角度进行了综述。Cryo-EM技术使我们得到了大体系的蛋白质机器的复杂结构,X射线晶体衍射能够解析高分辨率的蛋白质结构,而NMR在近生理的溶液环境甚至是细胞原位环境研究蛋白质的动态结构。此外,小角散射、单分子荧光和交联质谱虽然只能获得低分辨率的结构信息,但是在蛋白质的动态结构表征方面也各具有独到的优势。XL-MS能在细胞水平提供蛋白质相互作用的界面信息。由于蛋白质在细胞中的有效浓度更高,XL-MS有望捕获瞬态的蛋白质弱相互作用,以及蛋白质遭遇复合物。此外,通过与细胞内NMR方法的结合,XL-MS有望在细胞原位水平进行高时空分辨的捕获的表征,如蛋白质机器的组装。

对于蛋白质动态结构的研究,大部分的检测方法都会用到探针。NMR实验中有各种顺磁探针,smFRET需要用到各种荧光探针,交联剂在本质上也是一种探针。这些连接到蛋白质分子上的各种探针都不可避免的会对蛋白质的结构,动态变化甚至功能造成影响或者干扰。对于NMR和smFRET来说,往往还需要对特定位点做突变以保证标记位点的特异性。因此,在应用这些方法时,需要仔细的评估探针对于结构状态以及蛋白质功能的影响,同时还应该结合其它方法进行交叉验证。

目前有专门的数据库PDB-Dev来专门收录通过各种整合方法获得的生物大分子结构^[123,124]。随着技术方法的不断发展,蛋白质分子机器的结构和动态变化研

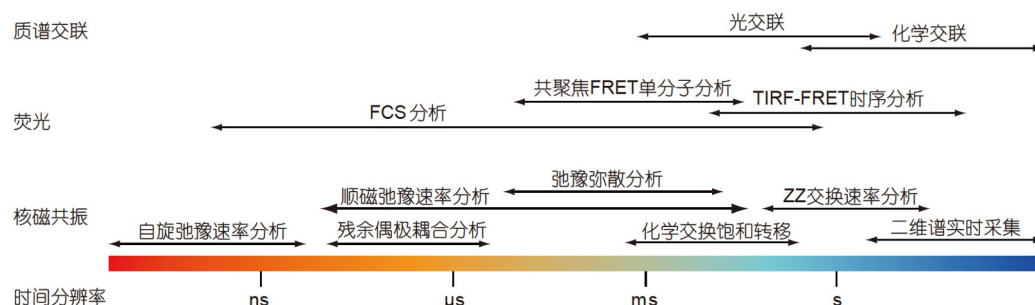


图5 可通过对质谱交联、荧光、核磁共振等谱学手段的整合,表征分析蛋白质分子机器在不同时间尺度的结构动态。其中,核磁共振中的不同测量方法覆盖超过10个数量级^[63](网络版彩图)

Figure 5 The structural dynamics of protein molecular machines on different time scales can be characterized and analyzed by integrating mass spectrometry cross-linking, fluorescence, and nuclear magnetic resonance (NMR) spectroscopic techniques. Among them, different measurement methods in NMR cover more than 10 orders of magnitude ^[63] (color online).

究也将逐渐从研究体外的纯化蛋白质发展到直接研究细胞内、亚细胞器内、甚至组织内的蛋白质体系。各种实验与计算方法的不断发展也将不断推动从原子到生命系统的研究。在这个过程中, 各种方法的结构表征也将加速产生海量的数据。这些数据也将成为未来人工智能算法训练的基础。因此, 有效整理各种实验数据, 对不同数据格式之间的兼容以方便有效的整合分析, 设立统一的规范和标准以提升数据的质量, 甚至是针对不同的实验技术发展相应的人工智能方法来对数据进行核查和评判, 以进一步提升数据的泛化和拓展能力。

需要指出的是, 目前大多数方法研究蛋白质动态结构还是对热力学平衡状态的检测, 往往缺少一个净的能量输入和随时间变化的动力学信息。在蛋白质分子机器行使功能时, 会经历多个非平衡状态, 如过渡态或者激发态。但由于非平衡状态寿命短占比低, 因

此难以检测到。在NMR研究中, 动态核极化(Dynamic Nuclear Polarization, DNP)是一种能大幅提高NMR信号的技术。DNP通过将样品中的核极化度增强到远高于热平衡状态的水平来实现对信号的大幅增强^[125,126], 从而可以极大的提升NMR方法的检测灵敏度。发展DNP-NMR技术, 并在此基础上结合其他的磁共振检测和分析技术, 有望在原子水平捕获和监测蛋白质的各种动态结构, 如处于非平衡状态的酶的激发态结构。利用Pressure Jump NMR技术, 还能够分析ms时间尺度的蛋白质聚集、蛋白质折叠中间态等^[127]。此外, 结合微流控技术, 对样品体系的微型化也是近年来一个趋势^[128]。对非平衡态下蛋白质分子机器动态结构的研究, 对蛋白质结构动态变化的高分辨率的实时监测与分析, 是今后蛋白质分子机器研究方法的重要发展方向, 并将进一步加深我们对蛋白质分子机器行使功能的机制的认识。

参考文献

- 1 McCammon JA, Karplus M. *Acc Chem Res*, 1983, 16: 187–193
- 2 Karplus M, Kuriyan J. *Proc Natl Acad Sci USA*, 2005, 102: 6679–6685
- 3 Marsh JA, Teichmann SA. *Annu Rev Biochem*, 2015, 84: 551–575
- 4 Sekhar A, Kay LE. *Annu Rev Biophys*, 2019, 48: 297–319
- 5 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. *Nature*, 2021, 596: 583–589
- 6 Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN, Schaeffer RD, Millán C, Park H, Adams C, Glassman CR, DeGiovanni A, Pereira JH, Rodrigues AV, van Dijk AA, Ebrecht AC, Opperman DJ, Sagmeister T, Buhlheller C, Pavkov-Keller T, Rathinaswamy MK, Dalwadi U, Yip CK, Burke JE, Garcia KC, Grishin NV, Adams PD, Read RJ, Baker D. *Science*, 2021, 373: 871–876
- 7 Tejero R, Huang YJ, Ramelot TA, Montelione GT. *Front Mol Biosci*, 2022, 9: 877000
- 8 Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, Židek A, Green T, Tunyasuvunakool K, Petersen S, Jumper J, Clancy E, Green R, Vora A, Lutfi M, Figurnov M, Cowie A, Hobbs N, Kohli P, Kleywegt G, Birney E, Hassabis D, Velankar S. *Nucleic Acids Res*, 2022, 50: D439–D444
- 9 Li EH, Spaman LE, Tejero R, Huang YJ, Ramelot TA, Fraga KJ, Prestegard JH, Kennedy MA, Montelione GT. *J Magn Reson*, 2023, 352: 107481
- 10 Theillet FX. *Chem Rev*, 2022, 122: 9497–9570
- 11 Bonomi M, Heller GT, Camilloni C, Vendruscolo M. *Curr Opin Struct Biol*, 2017, 42: 106–116
- 12 Edwards MJ, White GF, Butt JN, Richardson DJ, Clarke TA. *Cell*, 2020, 181: 665–673.e10
- 13 Zhang B, Li J, Yang X, Wu L, Zhang J, Yang Y, Zhao Y, Zhang L, Yang X, Yang X, Cheng X, Liu Z, Jiang B, Jiang H, Guddat LW, Yang H, Rao Z. *Cell*, 2019, 176: 636–648.e13
- 14 Kermani AA. *FEBS J*, 2021, 288: 5788–5804
- 15 Maestre-Reyna M, Wang PH, Nango E, Hosokawa Y, Saft M, Furrer A, Yang CH, Putu EPGN, Wu WJ, Emmerich HJ, Caramello N, Franz-

- Badur S, Yang C, Engilberge S, Wranik M, Glover HL, Weinert T, Wu HY, Lee CC, Huang WC, Huang KF, Chang YK, Liao JH, Weng JH, Gad W, Chang CW, Pang AH, Yang KC, Lin WT, Chang YC, Gashi D, Beale E, Ozerov D, Nass K, Knopp G, Johnson PJM, Cirelli C, Milne C, Bacellar C, Sugahara M, Owada S, Joti Y, Yamashita A, Tanaka R, Tanaka T, Luo F, Tono K, Zarzycka W, Müller P, Alahmad MA, Bezold F, Fuchs V, Gnau P, Kiontke S, Korf L, Reithofer V, Rosner CJ, Seiler EM, Watad M, Werel L, Spadaccini R, Yamamoto J, Iwata S, Zhong D, Standfuss J, Royant A, Bessho Y, Essen LO, Tsai MD. *Science*, 2023, 382: eadd7795
- 16 Christou NE, Apostolopoulou V, Melo DVM, Ruppert M, Fadini A, Henkel A, Sprenger J, Oberthuer D, Günther S, Pateras A, Mashhour AR, Yefanov OM, Galchenkova M, Reinke PYA, Kremling V, Scheer TES, Lange ER, Middendorf P, Schubert R, De Zitter E, Lumbao-Conradson K, Herrmann J, Rahighi S, Kunavar A, Beale EV, Beale JH, Cirelli C, Johnson PJM, Dworkowski F, Ozerov D, Bertrand Q, Wranik M, Bacellar C, Bajt S, Wakatsuki S, Sellberg JA, Huse N, Turk D, Chapman HN, Lane TJ. *Science*, 2023, 382: 1015–1020
- 17 Wayment-Steele HK, Ojoawo A, Otten R, Apitz JM, Pitsawong W, Hömberger M, Ovchinnikov S, Colwell L, Kern D. *Nature*, 2024, 625: 832–839
- 18 Yip KM, Fischer N, Paknia E, Chari A, Stark H. *Nature*, 2020, 587: 157–161
- 19 Liao M, Cao E, Julius D, Cheng Y. *Nature*, 2013, 504: 107–112
- 20 Dong Y, Zhang S, Wu Z, Li X, Wang WL, Zhu Y, Stoilova-McPhie S, Lu Y, Finley D, Mao Y. *Nature*, 2019, 565: 49–55
- 21 Bertram K, Agafonov DE, Liu WT, Dybkov O, Will CL, Hartmuth K, Urlaub H, Kastner B, Stark H, Lührmann R. *Nature*, 2017, 542: 318–323
- 22 Stiller JB, Otten R, Häussinger D, Rieder PS, Theobald DL, Kern D. *Nature*, 2022, 603: 528–535
- 23 Pérez LM, Ielasi FS, Bessa LM, Maurin D, Kragelj J, Blackledge M, Salvi N, Bouvignies G, Palencia A, Jensen MR. *Nature*, 2022, 602: 695–700
- 24 Mandala VS, Williams JK, Hong M. *Annu Rev Biophys*, 2018, 47: 201–222
- 25 Zhuo XF, Wang J, Zhang J, Jiang LL, Hu HY, Lu JX. *J Am Chem Soc*, 2020, 142: 3412–3421
- 26 Stiller JB, Jordan Kerns S, Hoemberger M, Cho YJ, Otten R, Hagan MF, Kern D. *Nat Catal*, 2019, 2: 726–734
- 27 Kay LE. *J Mol Biol*, 2016, 428: 323–331
- 28 Jensen MR, Ruigrok RW, Blackledge M. *Curr Opin Struct Biol*, 2013, 23: 426–435
- 29 Müntener T, Joss D, Häussinger D, Hiller S. *Chem Rev*, 2022, 122: 9422–9467
- 30 Fenwick RB, Oyen D, Wright PE. *Phys Chem Chem Phys*, 2016, 18: 5789–5798
- 31 Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KFA, Becker S, Meiler J, Grubmüller H, Griesinger C, de Groot BL. *Science*, 2008, 320: 1471–1475
- 32 Palmer AG, Koss H. *Methods Enzymol*, 2019, 615: 177–236
- 33 Palmer III AG. *J Magn Reson*, 2014, 241: 3–17
- 34 Murthy AC, Dignon GL, Kan Y, Zerbe GH, Parekh SH, Mittal J, Fawzi NL. *Nat Struct Mol Biol*, 2019, 26: 637–648
- 35 Liu Z, Gong Z, Jiang WX, Yang J, Zhu WK, Guo DC, Zhang WP, Liu ML, Tang C. *eLife*, 2015, 4: e05767
- 36 Clore GM, Iwahara J. *Chem Rev*, 2009, 109: 4108–4139
- 37 Liu Z, Zhang WP, Xing Q, Ren X, Liu M, Tang C. *Angew Chem Int Ed*, 2012, 51: 469–472
- 38 Xing Q, Huang P, Yang J, Sun JQ, Gong Z, Dong X, Guo DC, Chen SM, Yang YH, Wang Y, Yang MH, Yi M, Ding YM, Liu ML, Zhang WP, Tang C. *Angew Chem Int Ed*, 2014, 53: 11501–11505
- 39 Liu Z, Gong Z, Guo DC, Zhang WP, Tang C. *Biochemistry*, 2014, 53: 1403–1409
- 40 Cumberworth A, Lamour G, Babu MM, Gsponer J. *Biochem J*, 2013, 454: 361–369
- 41 van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT, Kim PM, Kriwacki RW, Oldfield CJ, Pappu RV, Tompa P, Uversky VN, Wright PE, Babu MM. *Chem Rev*, 2014, 114: 6589–6631
- 42 Wright PE, Dyson HJ. *Nat Rev Mol Cell Biol*, 2015, 16: 18–29
- 43 Phillips AH, Kriwacki RW. *Curr Opin Struct Biol*, 2020, 60: 1–6
- 44 Camacho-Zarco AR, Schnapka V, Guseva S, Abyzov A, Adamski W, Milles S, Jensen MR, Zidek L, Salvi N, Blackledge M. *Chem Rev*, 2022, 122: 9331–9356
- 45 Mittag T, Forman-Kay JD. *Curr Opin Struct Biol*, 2007, 17: 3–14
- 46 Abyzov A, Blackledge M, Zweckstetter M. *Chem Rev*, 2022, 122: 6719–6748
- 47 Brini E, Simmerling C, Dill K. *Science*, 2020, 370: eaaz3041

- 48 Dror RO, Young C, Shaw DE. Anton, A Special-Purpose Molecular Simulation Machine. In: *Encyclopedia of Parallel Computing*. Boston: Springer, 2011, 60–71
- 49 Mondal A, Lenz S, MacCallum JL, Perez A. *Curr Opin Struct Biol*, 2023, 81: 102609
- 50 Ayaz P, Lyczek A, Paung YT, Mingione VR, Iacob RE, de Waal PW, Engen JR, Seeliger MA, Shan Y, Shaw DE. *Nat Commun*, 2023, 14: 1885
- 51 Chan-Yao-Chong M, Durand D, Ha-Duong T. *J Chem Inf Model*, 2019, 59: 1743–1758
- 52 Yu L, Bruschweiler R. *PLoS Comput Biol*, 2022, 18: e1010036
- 53 Gupta A, Dey S, Hicks A, Zhou HX. *Commun Biol*, 2022, 5: 610
- 54 Ramanathan A, Ma H, Parvatikar A, Chennubhotla SC. *Curr Opin Struct Biol*, 2021, 66: 216–224
- 55 Evenäs J, Tugarinov V, Skrynnikov NR, Goto NK, Muhandiram R, Kay LE. *J Mol Biol*, 2001, 309: 961–974
- 56 Meisburger SP, Thomas WC, Watkins MB, Ando N. *Chem Rev*, 2017, 117: 7615–7672
- 57 Sonntag M, Jagtap PKA, Simon B, Appavou MS, Geerlof A, Stehle R, Gabel F, Hennig J, Sattler M. *Angew Chem Int Ed*, 2017, 56: 9322–9325
- 58 Gong Z, Yang J, Qin LY, Tang C, Jiang H, Ke Y, Dong X. *J Phys Chem B*, 2022, 126: 4840–4848
- 59 Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. *Nucleic Acids Res*, 2016, 44: W424–W429
- 60 Agam G, Gebhardt C, Popara M, Mächtel R, Folz J, Ambrose B, Chamachi N, Chung SY, Craggs TD, de Boer M, Grohmann D, Ha T, Hartmann A, Hendrix J, Hirschfeld V, Hübner CG, Hugel T, Kammerer D, Kang HS, Kapanidis AN, Krainer G, Kramm K, Lemke EA, Lerner E, Margeat E, Martens K, Michaelis J, Mitra J, Moya Muñoz GG, Quast RB, Robb NC, Sattler M, Schlierf M, Schneider J, Schröder T, Sefer A, Tan PS, Thurn J, Tinnefeld P, van Noort J, Weiss S, Wendler N, Zijlstra N, Barth A, Seidel CAM, Lamb DC, Cordes T. *Nat Methods*, 2023, 20: 523–535
- 61 Krainer G, Keller S, Schlierf M. *Curr Opin Struct Biol*, 2019, 58: 124–137
- 62 Metskas LA, Rhoades E. *Annu Rev Phys Chem*, 2020, 71: 391–414
- 63 Tang C, Gong Z. *J Mol Biol*, 2020, 432: 2913–2929
- 64 Vedel IM, Papagiannoula A, Naudi-Fabra S, Milles S. *Curr Opin Struct Biol*, 2023, 82: 102659
- 65 Liu Z, Dong X, Yi HW, Yang J, Gong Z, Wang Y, Liu K, Zhang WP, Tang C. *Cell Discov*, 2019, 5: 19
- 66 Gomes GNW, Krzeminski M, Namini A, Martin EW, Mittag T, Head-Gordon T, Forman-Kay JD, Gradinaru CC. *J Am Chem Soc*, 2020, 142: 15697–15710
- 67 Piersimoni L, Kastiris PL, Arlt C, Sinz A. *Chem Rev*, 2021, 122: 7500–7531
- 68 Ding YH, Gong Z, Dong X, Liu K, Liu Z, Liu C, He SM, Dong MQ, Tang C. *J Biol Chem*, 2017, 292: 1187–1196
- 69 Gong Z, Liu Z, Dong X, Ding YH, Dong MQ, Tang C. *Biophys Rep*, 2017, 3: 100–108
- 70 Gong Z, Ye SX, Nie ZF, Tang C. *J Phys Chem B*, 2020, 124: 4446–4453
- 71 Gong Z, Ye SX, Tang C. *Structure*, 2020, 28: 1160–1167.e3
- 72 Ferber M, Kosinski J, Ori A, Rashid UJ, Moreno-Morcillo M, Simon B, Bouvier G, Batista PR, Müller CW, Beck M, Nilges M. *Nat Methods*, 2016, 13: 515–520
- 73 Russel D, Lasker K, Webb B, Velázquez-Muriel J, Tjioe E, Schneidman-Duhovny D, Peterson B, Sali A. *PLoS Biol*, 2012, 10: e1001244
- 74 Schmidt C, Urlaub H. *Curr Opin Struct Biol*, 2017, 46: 157–168
- 75 Shi Y, Pellarin R, Fridy PC, Fernandez-Martinez J, Thompson MK, Li Y, Wang QJ, Sali A, Rout MP, Chait BT. *Nat Methods*, 2015, 12: 1135–1138
- 76 Liu Z, Gong Z, Cao Y, Ding YH, Dong MQ, Lu YB, Zhang WP, Tang C. *Biochemistry*, 2018, 57: 305–313
- 77 Boob M, Wang Y, Gruebele M. *J Phys Chem B*, 2019, 123: 8341–8350
- 78 Serber Z, Dötsch V. *Biochemistry*, 2001, 40: 14317–14323
- 79 Li C, Liu M. *FEBS Lett*, 2013, 587: 1008–1011
- 80 Theillet FX, Binolfi A, Bekei B, Martorana A, Rose HM, Stuijver M, Verzini S, Lorenz D, van Rossum M, Goldfarb D, Selenko P. *Nature*, 2016, 530: 45–50
- 81 Pan BB, Yang F, Ye Y, Wu Q, Li C, Huber T, Su XC. *Chem Commun*, 2016, 52: 10237–10240
- 82 Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Wälchli M, Smith BO, Shirakawa M, Güntert P, Ito Y. *Nature*, 2009, 458: 102–105
- 83 Tanaka T, Ikeya T, Kamoshida H, Suemoto Y, Mishima M, Shirakawa M, Güntert P, Ito Y. *Angew Chem Int Ed*, 2019, 58: 7284–7288

- 84 Huang Y, Reddy KD, Bracken C, Qiu B, Zhan W, Eliezer D, Boudker O. *J Am Chem Soc*, 2023, 145: 8583–8592
- 85 Huang Y, Wang X, Lv G, Razavi AM, Huysmans GHM, Weinstein H, Bracken C, Eliezer D, Boudker O. *Nat Chem Biol*, 2020, 16: 1006–1012
- 86 Zhu W, Guseman AJ, Bhinderwala F, Lu M, Su XC, Gronenborn AM. *Angew Chem Int Ed*, 2022, 61: e202201097
- 87 Chai Z, Wu Q, Cheng K, Liu X, Zhang Z, Jiang L, Zhou X, Liu M, Li C. *Angew Chem Int Ed*, 2023, 62: e202300318
- 88 Pham LBT, Costantino A, Barbieri L, Calderone V, Luchinat E, Banci L. *J Am Chem Soc*, 2023, 145: 1389–1399
- 89 Speer SL, Zheng W, Jiang X, Chu IT, Guseman AJ, Liu M, Pielak GJ, Li C. *Proc Natl Acad Sci USA*, 2021, 118: e2019918118
- 90 Chavez JD, Weisbrod CR, Zheng C, Eng JK, Bruce JE. *Mol Cell Proteomics*, 2013, 12: 1451–1467
- 91 Wheat A, Yu C, Wang X, Burke AM, Chemmama IE, Kaake RM, Baker P, Rychnovsky SD, Yang J, Huang L. *Proc Natl Acad Sci USA*, 2021, 118: e2023360118
- 92 An Y, Zhao Q, Gao H, Zhao L, Li X, Zhang X, Liang Z, Zhang L, Zhang Y. *Anal Chem*, 2022, 94: 3904–3913
- 93 Tan D, Li Q, Zhang MJ, Liu C, Ma C, Zhang P, Ding YH, Fan SB, Tao L, Yang B, Li X, Ma S, Liu J, Feng B, Liu X, Wang HW, He SM, Gao N, Ye K, Dong MQ, Lei X. *eLife*, 2016, 5: e12509
- 94 Liu F, Rijkers DTS, Post H, Heck AJR. *Nat Methods*, 2015, 12: 1179–1184
- 95 Yang B, Wu YJ, Zhu M, Fan SB, Lin J, Zhang K, Li S, Chi H, Li YX, Chen HF, Luo SK, Ding YH, Wang LH, Hao Z, Xiu LY, Chen S, Ye K, He SM, Dong MQ. *Nat Methods*, 2012, 9: 904–906
- 96 Jiang PL, Wang C, Diehl A, Viner R, Etienne C, Nandhikonda P, Foster L, Bomgardner RD, Liu F. *Angew Chem Int Ed*, 2022, 61: e202113937
- 97 Gao H, Zhao L, Zhong B, Zhang B, Gong Z, Zhao B, Liu Y, Zhao Q, Zhang L, Zhang Y. *Anal Chem*, 2022, 94: 7551–7558
- 98 Chen J, Zhao Q, Gao H, Zhao L, Chu H, Shan Y, Liang Z, Zhang Y, Zhang L. *Angew Chem Int Ed*, 2023, 62: e202212860
- 99 Sciolino N, Liu A, Breindel L, Burz DS, Sulchek T, Shekhtman A. *Commun Biol*, 2022, 5: 451
- 100 Chen Y, Zhou W, Xia Y, Zhang W, Zhao Q, Li X, Gao H, Liang Z, Ma G, Yang K, Zhang L, Zhang Y. *Nat Commun*, 2023, 14: 3882
- 101 Chavez JD, Lee CF, Caudal A, Keller A, Tian R, Bruce JE. *Cell Syst*, 2018, 6: 136–141.e5
- 102 Gonzalez-Lozano MA, Koopmans F, Sullivan PF, Protze J, Krause G, Verhage M, Li KW, Liu F, Smit AB. *Sci Adv*, 2020, 6: eaax5783
- 103 Wittig S, Ganzella M, Barth M, Kostmann S, Riedel D, Pérez-Lara Á, Jahn R, Schmidt C. *Nat Commun*, 2021, 12: 858
- 104 Zhang B, Gong Z, Zhao L, An Y, Gao H, Chen J, Liang Z, Liu M, Zhang Y, Zhao Q, Zhang L. *Angew Chem Int Ed*, 2023, 62: e202301345
- 105 Wright SC, Lukasheva V, Le Gouill C, Kobayashi H, Breton B, Mailhot-Larouche S, Blondel-Tepaz É, Vieira NA, Costa-Neto C, Héroux M, Lambert NA, Parreiras-e-Silva LT, Bouvier M. *Proc Natl Acad Sci USA*, 2021, 118: e2025846118
- 106 Yang J, Gong Z, Lu YB, Xu CJ, Wei TF, Yang MS, Zhan TW, Yang YH, Lin L, Liu J, Tang C, Zhang WP. *J Mol Biol*, 2020, 432: 4596–4611
- 107 Jiang WX, Dong X, Jiang J, Yang YH, Yang J, Lu YB, Fang SH, Wei EQ, Tang C, Zhang WP. *Sci Rep*, 2016, 6: 20568
- 108 Anthis NJ, Clore GM. *Quart Rev Biophys*, 2015, 48: 35–116
- 109 Hansen AL, Lundström P, Velyvis A, Kay LE. *J Am Chem Soc*, 2012, 134: 3178–3189
- 110 Vallurupalli P, Bouvignies G, Kay LE. *J Am Chem Soc*, 2012, 134: 8148–8161
- 111 Niu X, Ding J, Zhang W, Li Q, Hu Y, Jin C. *J Magn Reson*, 2018, 293: 47–55
- 112 Kleckner IR, Foster MP. *Biochim Biophys Acta*, 2011, 1814: 942–968
- 113 Lerner E, Barth A, Hendrix J, Ambrose B, Birkedal V, Blanchard SC, Börner R, Chung HS, Cordes T, Craggs TD, Deniz AA, Diao J, Fei J, Gonzalez RL, Gopich IV, Ha T, Hanke CA, Haran G, Hatzakis NS, Hohng S, Hong SC, Hugel T, Ingargiola A, Joo C, Kapanidis AN, Kim HD, Laurence T, Lee NK, Lee TH, Lemke EA, Margeat E, Michaelis J, Michalet X, Myong S, Nettels D, Peulen TO, Ploetz E, Razvag Y, Robb NC, Schuler B, Soleimaninejad H, Tang C, Vafabakhsh R, Lamb DC, Seidel CA, Weiss S. *eLife*, 2021, 10: e60416
- 114 Soranno A, Buchli B, Nettels D, Cheng RR, Müller-Spätth S, Pfeil SH, Hoffmann A, Lipman EA, Makarov DE, Schuler B. *Proc Natl Acad Sci USA*, 2012, 109: 17800–17806
- 115 Gopich IV, Szabo A. *Proc Natl Acad Sci USA*, 2012, 109: 7747–7752
- 116 Kalinin S, Peulen T, Sindbert S, Rothwell PJ, Berger S, Restle T, Goody RS, Gohlke H, Seidel CAM. *Nat Methods*, 2012, 9: 1218–1225
- 117 Mazal H, Haran G. *Curr Opin BioMed Eng*, 2019, 12: 8–17
- 118 Wang Z, Zhang H, Jian L, Ding B, Huang K, Zhang W, Xiao Q, Huang S. *Biophys Rep*, 2022, 8: 100–118
- 119 Wang JH, Tang YL, Gong Z, Jain R, Xiao F, Zhou Y, Tan D, Li Q, Huang N, Liu SQ, Ye K, Tang C, Dong MQ, Lei X. *Nat Commun*, 2022, 13: 1468
- 120 Wang JH, Gong Z, Dong X, Liu SQ, Tang YL, Lei XG, Tang C, Dong MQ. *Biophys Rep*, 2022, 8: 239–252

- 121 Ziemianowicz DS, Ng D, Schryvers AB, Schriemer DC. *J Proteome Res*, 2019, 18: 934–946
- 122 Stahl K, Graziadei A, Dau T, Brock O, Rappsilber J. *Nat Biotechnol*, 2023, 41: 1810–1819
- 123 Sali A, Berman HM, Schwede T, Trewhella J, Kleywegt G, Burley SK, Markley J, Nakamura H, Adams P, Bonvin AMJJ, Chiu W, Peraro MD, Di Maio F, Ferrin TE, Grünwald K, Gutmanas A, Henderson R, Hummer G, Iwasaki K, Johnson G, Lawson CL, Meiler J, Marti-Renom MA, Montelione GT, Nilges M, Nussinov R, Patwardhan A, Rappsilber J, Read RJ, Saibil H, Schröder GF, Schwieters CD, Seidel CAM, Svergun D, Topf M, Ulrich EL, Velankar S, Westbrook JD. *Structure*, 2015, 23: 1156–1167
- 124 Vallat B, Webb B, Fayazi M, Voinea S, Tangmunarunkit H, Ganesan SJ, Lawson CL, Westbrook JD, Kesselman C, Sali A, Berman HM. *Acta Crystallogr D Struct Biol*, 2021, 77: 1486–1496
- 125 Mao J, Aladin V, Jin X, Leeder AJ, Brown LJ, Brown RCD, He X, Corzilius B, Glaubit C. *J Am Chem Soc*, 2019, 141: 19888–19901
- 126 Narasimhan S, Scherpe S, Lucini Paioni A, van der Zwan J, Folkers GE, Ovaa H, Baldus M. *Angew Chem Int Ed*, 2019, 58: 12969–12973
- 127 Charlier C, Courtney JM, Alderson TR, Anfinrud P, Bax A. *J Am Chem Soc*, 2018, 140: 8096–8099
- 128 Plata M, Hale W, Sharma M, Werner JM, Utz M. *Lab Chip*, 2021, 21: 1598–1603

An integrative approach to understand structural dynamics of protein molecular machines across spatiotemporal scales

Yueling Zhu^{1,3}, Zhou Gong¹, Chun Tang^{2*}

¹Innovation Academy for Precision Measurement Science and Technology Chinese Academy of Sciences, Wuhan 430071, China

²College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

³University of Chinese Academy of Sciences, Beijing 100049, China

*Corresponding author (email: Tang_Chun@pku.edu.cn)

Abstract: Proteins constitute the most important molecular machines within biological organisms. Understanding the dynamic properties of protein molecular machines and revealing their working mechanisms are crucial for the regulation of their functions. This article provides a comprehensive and systematic review of the research methods for the structural dynamics of protein molecular machines. It delves into the characteristics and applications of different methods from perspectives such as spatial resolution, time-scale, *in vitro* versus *in vivo*. Additionally, the review highlights the advantages and limitations of individual methods as well as strategies for an integrated approach. It also looks forward to the future development on the structural dynamics of protein molecular machines and provides authoritative viewpoints.

Keywords: protein machinery, structural dynamics, integration of multiple methods, spatiotemporal scale

doi: [10.1360/SSC-2023-0247](https://doi.org/10.1360/SSC-2023-0247)