

核磁共振在生物学研究中的应用

编者按: 经过 60 多年的发展,核磁共振波谱及成像技术的应用已经从物理学扩展到化学、生物学、医学、信息学等诸多领域,并发挥着越来越重要的作用。在本期发表的“核磁共振在生物学研究中的应用”专题中,作者们重点介绍了可大幅度提高灵敏度的超极化¹²⁹Xe 核磁共振波谱和成像技术及其在生物医学中的应用,核磁共振成像技术及其在脑功能研究中的应用,核磁共振波谱技术及其在生物大分子的三维结构、相互作用和动态过程研究中的应用。生物核磁共振是物理学与生物医学的综合交叉,涉及面广,本专题中不可能全面涉及,本刊今后将会陆续组织文章进行介绍。

核磁共振技术在生物研究中的应用 *

姜凌 刘买利[†]

(中国科学院武汉物理与数学研究所 武汉磁共振中心 波谱与原子分子物理国家重点实验室 武汉 430071)

摘要 核磁共振(nuclear magnetic resonance, NMR)是以原子核自旋的共振跃迁为探测对象的谱学方法。当自旋量子数不为零的原子核处于外磁场中时,会引起能级的 Zeeman 分裂。若再施加能量等于 Zeeman 能级差的射频场,则会诱发原子核自旋的共振跃迁,这种现象即为核磁共振。核自旋的共振频率与原子的类型有关,且受原子所处化学和物理环境的影响。此外,NMR 能量较低,不会影响探测对象(常为分子)的状态。因此,NMR 能够在无损条件下提供多种具有原子和分子分辨的物质组成、结构、形态、动态变化等丰富信息。

关键词 核磁共振,蛋白质溶液结构,蛋白质动力学,药物识别

Nuclear magnetic resonance applications in biological systems

JIANG Ling LIU Mai-Li[†]

(State Key laboratory of Magnetic Resonance and Atomic and Molecular Physics,
Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan 430071, China)

Abstract Nuclear magnetic resonance (NMR) spectroscopy is a state-of-the-art technology which has been widely applied in biological systems over the past decades. It is a powerful tool for macromolecular structure determination in solution, and has the unique advantage of being capable of elucidating the structure and dynamic behavior of proteins during vital biomedical processes. In this review, we introduce the recent progress in NMR techniques for studying the structure, interaction and dynamics of proteins. The methods for NMR based drug discovery and metabolomics are also briefly introduced.

Keywords nuclear magnetic resonance(NMR), protein solution structure, protein dynamics, drug screen

自 1945 年观察到凝聚态物质的 NMR 现象之后,经过 60 多年的发展,NMR 的研究领域和应用范围已经从物理学延伸到化学、生物学、医学、材料科学、信息学等几乎所有自然科学领域。先后有 5 次诺贝尔奖授予了核磁共振研究工作者,其中物理学奖两次(Rabi, 1944 年; Bloch 和 Purcell, 1952 年),

化学奖两次(Ernst, 1991 年; Wüthrich 等, 2002 年),生理和医学奖一次(Lauterbur 和 Mansfield,

* 国家自然科学基金(批准号:20921004;90813017)、国家重点基础研究发展计划(批准号:2009CB918600)资助项目

2011-02-28 收到

† 通讯联系人, Email: ml.liu@wipm.ac.cn

2003 年). 这表明 NMR 是一门学科交叉极其广泛的技术, 已经成为物理学、化学以及生命科学等诸多学科研究物质组成的成分、组织形态及其变化、脑功能、分子结构和动力学强有力的手段, 广泛应用于科学的研究的前沿领域。NMR 在生物医学领域的应用最为广泛也最受关注。例如:(1)蛋白质等生物大分子的三维结构和相互作用动力学的测定, 以生物大分子为靶标的药物筛选和药物分子结构优化, NMR 已成为蛋白质结构与功能研究和药物研发的强有力的工具;(2)以生物小分子为检测对象的代谢组学研究是近年来发展迅速的领域, 涉及人口健康和国民生活质量, 特别是和重大疾病的预防和治疗、新型高效药物的研发等密切相关。

本文将重点介绍 NMR 技术在上述这些领域的发展, 介绍该技术在蛋白质溶液高分辨结构解析、蛋白质相互作用、药物筛选以及代谢组学等方面的应用和最新进展。

1 概述

NMR 所提供的信息主要包括在原子核自旋的化学位移、耦合常数和弛豫速率等参数中。化学位移能够反映原子核所在的官能团及其在分子中的位置。耦合常数包括自旋—自旋耦合常数和偶极—偶极耦合常数。自旋—自旋耦合(又称 J-耦合或标量耦合)只存在于相隔 2—4 个化学键的核自旋之间。相隔 3 个化学键的 2 个核之间的 J-耦合常数的大小与这 3 个化学键组成的二面角的夹角有关, 是一种重要的分子结构参数。两个原子核之间的偶极—偶极耦合常数与两者之间的距离和在磁场中的取向有关, 它也是一种重要的分子立体结构参数。事实上, 多维 NMR 方法毫无例外地建立在原子核间的这两种耦合作用之上, 用于测定原子核或化学键之间的连接关系以及原子核之间的空间距离和取向。弛豫速率反映了某个自旋体系与周围环境和其他自旋体系之间的能量交换过程, 包含着分子内不同时间尺度的动力学信息。因此 NMR 方法检测到的是生物分子在溶液中的结构和运动状态, 可以在接近于生理环境的温度、盐浓度和 pH 值状态下检测生物分子的结构和相互作用, 可用于监测生物分子的动态变化, 研究蛋白质与金属离子、药物小分子、蛋白质或核酸的相互作用。

在蛋白质数据库(PDB)收录的生物大分子结构中, 用 X 射线晶体学、核磁共振波谱学(NMR)和电

子显微学(EM)方法测定的结构各占一定的比例。图 1 数据取自 2010 年 9 月的 PDB 数据库, 可见用 NMR 解析的结构数目逐年增加, 在 2007 年达到顶峰, 当年共有 965 个 NMR 结构得到发表, 占当年发表总结构数的 13.4%。NMR 技术最大的优点在于它能对溶液中和非晶态中的蛋白质进行测量, 实验条件更接近于生物大分子的生理环境, 适用于分析不能结晶或者结晶后的构像可能会改变的蛋白质。因此, NMR 技术成为除 X 射线以外最主要的蛋白质结构测定方法之一。

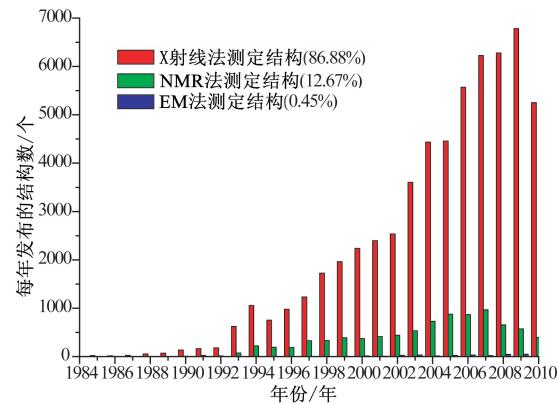


图 1 三种主要方法测定的生物大分子结构数逐年增加情况(图例中的 3 种颜色见本期《物理》网刊彩图)

更重要的是, 核磁共振方法能观察到蛋白质在很宽的时间尺度(从皮秒到秒量级)范围内的分子动力学过程等非常重要的信息, 如蛋白质不同结构域的相对运动, 蛋白与其他物质之间弱的或瞬态相互作用, 以及蛋白质折叠过程等, 这是其他方法很难观测到的。近几年来, NMR 研究不再局限于生物分子的结构测定, 研究重心逐渐向测定溶液状态下蛋白质相互作用、单个或多个结构域的动力学性质、捕获中间态或瞬态结构等研究方向转移, 以期更清晰地了解生命活动的奥秘, 从而显示了 NMR 方法旺盛的生命力。

2 核磁共振技术与方法的发展

20 世纪 50 年代, 第一台 30MHz 连续波商品核磁共振谱仪诞生。20 世纪 60 年代, 超导技术用于磁体制造, 使商品核磁共振谱仪的磁场逐步提高, 质子共振频率达到 100MHz。由于计算机技术的飞速进步, 脉冲傅里叶变换 NMR 技术应运而生, 它的出现使 NMR 领域开始了革命性的飞跃。20 世纪 70 年代后期起, NMR 理论和实验技术飞速发展, 核磁共振谱仪磁场不断提高, 二维谱、多量子跃迁、固体高

分辨等技术日益成熟。20世纪80年代,瑞士科学家Wüthrich教授开始将核磁共振技术用于蛋白质研究,并首次测出了液态蛋白质的三维结构^[1],Wüthrich教授对蛋白质三维结构的NMR测定方法的贡献,使他获得了2002年诺贝尔化学奖。

NMR的灵敏度和分辨率与磁场强度呈指数或线性关系。90MHz永磁型谱仪的信噪比仅为16:1,360MHz谱仪的信噪比超过了100:1,800MHz的信噪比又提高了一个数量级,最近1GHz已经被推向市场。此外,探头技术的发展也对NMR的发展提供了新的契机。1998年推出的低温探头(cryogenic probe)使得信噪比成倍增加,其中500MHz低温探头的信噪比已经超过了常规800MHz探头,而900MHz低温探头的信噪比更高达7500:1。除低温探头外,灵敏度较高的微量探头(micro-probe)和流动探头(flow-probe)也逐渐成为NMR谱仪的标准配置。使用高磁场强度的谱仪和新型探头的直接效果是提高了NMR分析的灵敏度和分辨率。对小分子代谢物的检测下限可以达到纳克量级,而蛋白质的检测下限可到微摩尔量级。

NMR磁体技术的发展,极大地推动了NMR在生命科学诸多领域中的广泛应用。而这些领域的发展又不断提出新的科学问题和技术挑战,促进NMR理论的发展及技术和方法的创新。例如,分子生物(遗传)工程技术的迅速发展使科学家可以成功制备¹⁵N和¹³C等稳定同位素标记的蛋白质样品,使多维异核NMR实验得以实现,实验的分辨率和灵敏度也大幅度提高,并在蛋白质等生物分子的NMR结构与功能研究中得到广泛应用。随着NMR研究的生物体系日渐复杂,促使NMR高场磁体和低温探头成为NMR实验室的常规装备,NMR可解析的蛋白质分子量也不断增大,进而推动了各种新的生物标记技术的产生。

科学问题的需求推动了NMR实验方法的不断创新。1997年,Wüthrich教授的研究组建立了简称为横向弛豫优化谱(TROSY)的二维实验方法,达到了窄化谱线、提高分辨率和灵敏度的目的^[2],是20世纪末NMR方法应用于生物研究的两个最重要的进展之一,在此基础上发展了一系列具有高分辨率、高灵敏度的多维NMR方法。其中交叉弛豫极化转移增强(CRINEPT)方法能使110kDa以上的蛋白质的NMR信号强度提高3倍^[3],可以用于研究分子量高达900kDa的蛋白质复合体中的部分结构^[4]。另一个重要进展是,美国Bax教授等发展的

在部分定向介质中测定残余的偶极-偶极耦合常数(RDC)的方法^[5]。加拿大Kay教授及其研究组在蛋白质结构的NMR测定方法研究方面做出了很多有影响的成果,他们发展的弛豫弥散技术可以检测出蛋白质折叠过程中占蛋白质总量1%和2%的中间体^[6]。近年来,利用顺磁弛豫增强(PRE)的NMR方法有了长足发展^[7,8]。这个方法首先是在蛋白质分子中引入顺磁探针,顺磁粒子附近自旋核的弛豫速率会因此而加快,通过比较顺磁探针引入前后这些自旋核的弛豫速率变化,获得核与核之间的空间距离关系。这一方法可以探测1.5—2.4nm范围的远程结构信息,特别适用于研究蛋白质分子间远程和瞬态相互作用。

由于传统的NMR实验耗费时间,快速核磁共振方法越来越受重视。近年来发展的几种快速NMR方法突破了传统的采样技术或数据处理模式,最大限度地挖掘低维谱的结构信息。2002年,Szyperski等人提出了“降维NMR”的思想(RD-NMR),即把n维NMR谱的信息压缩到n-1维谱中^[9]。随后出现的GFT-NMR方法^[10]、投影重建的方法(PR-NMR)^[11]和基于Hadamard变换^[12]的快速NMR方法,可以使实验时间成数十倍地减少,并克服了RD-NMR多重峰的弱点。随机采样模式是近来受到广泛关注的新型采样模式,该技术避开了传统的矩阵型采样模式,使采样点在间接维上随机分布,采样点数远小于传统方法,再通过强大的计算机重建技术还原成经典的核磁谱图。由于它不采用任何既定的采样图形(如辐射形、螺旋形、圆形等),谱图重建时由于随机噪音互相抵消,谱图中的噪声伪峰较少。由快速采样方法而发展起来的多种谱图重建的计算方法各有优劣,如最大熵重建Max-Ent^[13]、多维分解技术MDD^[14]、非均匀傅里叶变换NU-FT^[15]等。最近发展的利用Gridding快速傅里叶变换技术处理非均匀采样数据的方法比传统MFT方法的速度提高了60倍^[16]。

3 核磁共振技术在生物研究中的应用

3.1 蛋白质溶液结构解析

蛋白质是生命活动的主要承担者,其生物功能主要由三维结构决定。全面了解蛋白质等生物分子的精细三维结构是在原子分子水平上研究生命活动的基础,也是蛋白质科学的核心研究内容之一。蛋白质三维结构测定所需的样品要经过基因工程表达

(包括自旋标记)、化学分离和纯化,根据所用测定方法的需要制备成不同形态的样品。通过 NMR 实验收集一系列多核多维谱,进行谱峰归属,判别 NMR 信号与蛋白质所含各核的对应关系,根据这些信息,提取结构约束参数,再运用分子动力学计算获得蛋白质的三维结构图像。这其中每一个环节的进步和发展,无不涉及多个学科知识的综合运用。

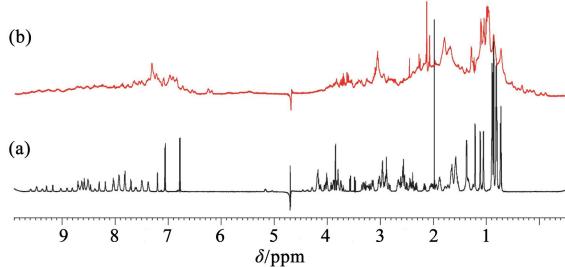


图 2 蛋白质的一维¹H-NMR 谱图,分子量分别为(a)5kDa 和(b)32kDa

图 2 是蛋白质的一维¹H-NMR 谱图,分子量分别为 5kDa(见图 2(a))和 32kDa(见图 2(b))。分子量增大明显降低了 NMR 谱图的分辨率,仅用¹H 谱不能解析生物大分子的信号。这是因为¹H 谱的化学位移范围比较狭窄(~ 15 ppm),NMR 信号易于重叠,而大多数蛋白质样品的¹³C 和¹⁵N 谱宽分别为 200ppm 和 30ppm 左右,信号更加分散。因此,分子量超过 5kD 的蛋白质测定溶液结构一般都需要进行同位素标记,即对蛋白质分子中的¹³C,¹⁵N 核进行同位素富集。蛋白质同位素标记技术是研究大分子蛋白质 NMR 结构的前提。

同位素富集分为全标记和选择性标记两种方式。全标记是将细菌培养基中的氮源或碳源化合物全部用¹⁵N,¹³C 标记的化合物替代,这样蛋白质的生物合成过程中仅摄取含有同位素标记的初始物,达到蛋白质全标记的目的。随着研究的蛋白质分子量增大,²H 全标记的样品也逐渐普及,从而可以有效地提高信号强度和分辨率。选择性标记是指仅对某些特定种类的氨基酸或者化学基团进行同位素标记,可以极大地减少信号的重叠程度。选择性标记需要对蛋白质的生物合成过程进行调控,细菌培养的过程通常比较复杂,成本也较高。现在较为通用的是对特定种类的氨基酸标记,如 Ala, Lys 等,一般需要使用缺陷型菌株,而且不同氨基酸的生物合成通路往往有交叉和重叠,要注意避免其他氨基酸被同时标记,造成混淆。另外,加拿大 Key 实验室发展的 VIL 甲基质子化²H,¹³C,¹⁵N 全三标记蛋白质制备技术^[17],日本 Kainosho 实验室发展的侧链质子选

择性²H 标记技术(SAIL),都是近年来受到广泛关注的新方法^[18],均成功地被用来解析大分子量(~ 100 kDa)蛋白质的 NMR 信号(见图 3)。

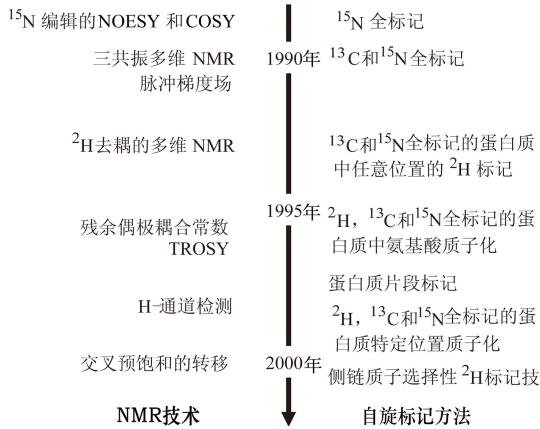


图 3 同位素标记技术及相应的 NMR 技术的发展历程(根据文献[19]重新绘制)

蛋白质结构解析的常规过程包括以下步骤:在获得适合结构测定的蛋白质样品后,首先需要采集多种多维 NMR 谱,在此基础上对蛋白质主链和侧链基团中各¹H/¹³C/¹⁵N 核进行归属。根据主链的化学位移(化学位移指数)可以得到蛋白质的二级结构。然后从有关 NMR 谱中提取几何约束参数,主要包括与核间距有关的核 Overhauser 效应(NOE,其数值与核间距的 6 次方成反比),自旋-自旋耦合常数(*J*,其数值与化学键二面角的关系符合 Karplus 公式),残余偶极-偶极耦合常数(*D*,其数值与化学键的空间取向有关)。最后,根据这些约束条件利用分子动力学计算,获得一组能量较低的三维结构。蛋白质三维 NMR 结构中不同区域的收敛程度有所不同,正好反映了蛋白质在溶液中的运动性,因此有别于晶体结构的单一构象。

在几何约束参数中,NOE 数量较多,并可给出一个距离范围(0.2—0.6nm),是蛋白质 NMR 结构解析最为重要的约束参数,如果平均每个氨基酸有 10—20 个 NOE 约束条件,主链的均方根偏差(rmsd)将达到 0.05nm,大概相当于精度为 0.2nm 的晶体结构。对于¹⁵N,¹³C 双标记的蛋白,可以得到大约每残基 20—25 个 NOE 约束,这时主链 rmsd 可以提高至 0.03—0.05nm。偶极-偶极耦合常数能够进一步提高结构测定的空间分辨率^[5, 20, 21],这对于分子量较大的蛋白质结构的 NMR 测定尤为重要。近年来,采用顺磁弛豫增强技术来获得长程结构约束在蛋白质复合体结构研究中应用越来越受重视^[7, 8]。最近,国立新加坡大学的杨代文教授提出了

一种基于 4D NOESY^[22] 实验的新的结构测定策略,通过区分残基内外的空间耦合关系等,使用少数几个 NMR 实验和相应的数据处理方法,就可测定出蛋白质的三维结构,单体的分子量可以达到 42kDa,选择性标记的多聚体可以达到 65kDa^[23].

相对来说,解析 20kDa 左右的蛋白质 NMR 结构比较容易实现,但是用核磁共振波谱法测定较大蛋白质($>30\text{kDa}$)的三维结构十分耗时,一般需要 3—5 个月,这是因为随着蛋白质分子量的增大,一方面,谱峰的数量随之急剧增大,需要采集更多的三维甚至更多维数的 NMR 实验才能完成谱峰的归属和获取足够的结构约束参数。另一方面,蛋白质分子运动性能减慢,谱峰增宽,灵敏度下降,还受到自旋扩散和分子局部运动的影响,这些都非常难以解析。针对大分子量蛋白质信号较弱、重叠严重的问题,解决办法之一是将蛋白质中的氢原子置换成氘原子,采用 TROSY 实验技术,可以显著提高谱图的分辨率和灵敏度^[2]。交叉弛豫极化转移(CRIPT)^[24]、交叉弛豫极化转移增强(CRINEPT)^[3]、交叉饱和转移(CST)^[25]等技术都有助于提高信号灵敏度,用于解决大分子量蛋白质的结构测定。

3.2 蛋白质复合体及其动力学的 NMR 研究

蛋白质在执行功能的过程中往往与其他蛋白或生物分子形成蛋白复合体。这种复合体是由多种成分经过适当的修饰和高度有序地精确装配而成,具有接受和执行重要的网络信息的功能。蛋白质复合体与遗传信息编码、酶催化、免疫应答、病毒感染、光合作用、药物(包括疫苗)设计等与人类重大疾病和创新药物研发密切相关。复合物的形成通常伴随着蛋白质局部结构的变化,其稳定性在亚微摩尔至亚摩尔量级,相应的动力学过程的差异也会有几个量级。因此蛋白复合体的三维结构远不只是一些单体蛋白结构的简单叠加,对任何方法而言,其结构测定往往比解析单个蛋白质结构更具挑战性。相对而言,对于溶液中蛋白质复合体的研究,特别是弱相互作用及动态过程的研究,NMR 独具特色和优势,发挥着非常重要的不可取代的作用。

解析蛋白复合物结构时最重要的是相互作用界面的确定,除了前面提到的 TROSY 和 CRINEPT 等技术外,¹³C/¹⁵N 过滤技术^[26]、化学位移微扰实验技术、残余偶极耦合技术、顺磁弛豫增强技术等均已运用于复合物的研究中。新型 NMR 实验和各种同位素标记技术的综合合理运用往往能达到意想不到的效果,近年来,670kDa 的 20S 蛋白水解酶结构^[27]

和 900kDa 的蛋白质复合体 GroEL/GroES 中部分结构^[4]的问世使 NMR 解析结构的上限又提高了一个数量级。

蛋白质复合物在行使功能时,有一个相互接触、结构变化、继而分离的过程,因此,生物体内弱的、暂态的相互作用非常普遍,结构域之间的相对运动显得尤为重要。PDZ1-2 是 NMDA 受体信号传导通路中的骨架蛋白 PSD-95 的 N 端 PDZ 结构域串中的前两个结构域。这种多结构域的蛋白质在信号通路和离子通道中极为普遍。这些蛋白质都很大,通常采用“各个击破”的策略,逐一研究单个结构域的结构和功能。PSD-95 的 PDZ1 和 PDZ2 相对较小,是研究结构域之间协同效应的理想模型。PDZ1-2 的三维溶液结构和动力学的 NMR 研究表明^[28],PDZ1 和 PDZ2 的孤立结构与其在 PDZ1-2 中的结构差别不大。它们与配体(cypin 肽段)形成复合体后,无论单独形成的复合体,还是共同形成的复合体,都能引起结构的变化,而且结构变化的程度也基本相同。纳秒量级的动力学测定(NMR 方法)表明,游离态 PDZ1-2 的两个结构域有一定的刚性。形成复合体后,PDZ1-2 的两个结构域之间的柔性大大增强,体系的熵增大,自由能降低(见图 4)。分子动力学模拟的结果与 NMR 实验完全符合。这一工作的结果表明,多结构域蛋白质可能通过结构和动力学重组来调控与配体的相互作用,结构域之间在行使功能时具有相关性。该结论有可能是多结构域蛋白质的共同特征^[28]。

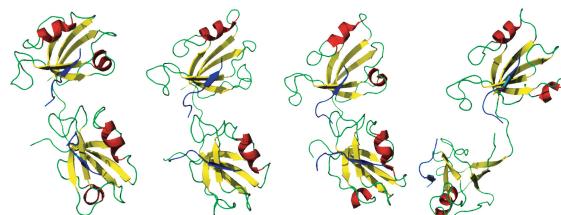


图 4 双结构域蛋白 PDZ1-2 通过结构和动力学重组来调控与配体(cypin 肽段)的相互作用(根据文献[28]重新制作)

NMR 弛豫弥散技术(relaxation dispersion)是近年来蛋白质动力学研究的重要进展之一,它通过改变自旋回波时间(CPMG)来测量横向弛豫速率受到的化学交换或相互作用等慢过程的影响,用于研究蛋白质中间态的结构和动力学过程^[6, 29]。2010 年,Key 小组在美国 *Science* 杂志上发表文章,用弛豫弥散技术确定了在蛋白折叠过程中寿命仅在毫秒量级的中间态结构^[30],发现脂肪酸结合蛋白 FBP11 的 FF 结构域在折叠时可以瞬时形成有完整结构的中间态,再缓慢过渡到稳定态结构。中间态的整体折

叠形式与稳定态类似,但含有更长的 H3 融合蛋白螺旋,而且 H3 和 H4 融合蛋白螺旋的相对倾斜角度略有不同(见图 5(a)). 顺磁弛豫增强技术也是近年来 NMR 领域的一大突破,它可以通过测定蛋白质的动力学结构变化,为复合物形成过程中低分布的激发态蛋白的存在提供有力的结构证据^[31]. Tang 等人在英国的 *Nature* 上发表文章指出,当顺磁探针处于一个动态的生物大分子系统中,如果稳态结构离顺磁探针较远,而瞬时激发态结构离顺磁探针较近时,瞬时激发态结构对 PRE 观察量的贡献就会比较显著,能观察到凸现出来的结构信息. 图 5(b)的红色部分是分布仅占 5% 的半关闭式激发态麦芽糖结合蛋白 MBP 的结构,与蓝色的自由稳态结构和绿色的结合稳态结构均有一定区别. PRE 方法界定的动力学过程的时间尺度是纳秒至微秒量级,与弛豫弥散技术(微秒至毫秒量级)互为补充.

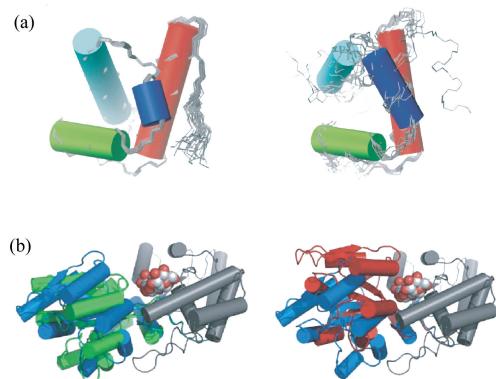


图 5 (a) 弛豫弥散技术测定 HYPA/FBP11 蛋白 FF 结构域的不同状态,左图为自由态结构,右图为瞬态结构^[30]; (b) 顺磁弛豫增强技术测定麦芽糖结合蛋白 MBP 溶液结构^[31]. 蓝色为开放式的自由态 MBP, 红色为半关闭式的自由态 MBP, 绿色为关闭式的结合态 MBP(见《物理》网刊彩图)

3.3 药物与蛋白质相互作用的 NMR 研究

用 NMR 技术研究药物与蛋白质的相互作用的方法灵活多变,可以根据检测对象分成药物检测和蛋白检测两种类型. 在以蛋白为检测对象的方法中,要进行比较的是药物加入前后靶蛋白的磁共振参数变化. 其中最著名的是 1996 年 Abbott 实验室 Shuker 等人提出 SAR-by-NMR 的概念^[32], 即通过比较加入药物前后的同位素标记蛋白质 NMR 谱图, 来确定化合物是否与蛋白质有相互作用, 该方法取得了显著成效. 这种方法可以在大型药物库中寻找多个可能的“hits”, 特别适合筛选结合较弱的药物^[33], 即离解常数 K_d 大于 $100 \mu\text{M}$ 的弱相互作用药物. 这种高通量筛选可以找到大量与靶蛋白相结合的结构片段, 并结合其他生化实验手段对药物进行

改造,从而指导药物设计和药物库的选择.

基于蛋白检测的 NMR 方法除了能筛选到不同结合强度的药物外,还可以观察到蛋白质与药物之间的 NOE 效应,确定复合物相互作用的界面,以及蛋白和药物的空间结构变化,从而有利于指导进一步的药物设计. 但是这种方法需要知道蛋白质的信号归属和结构信息,通常需要获得毫克量级的高纯度蛋白质样品,并辅以同位素标记,对蛋白质分子量也有一定的限制,且周期长,成本较高.

以药物为检测对象的方法比较的是蛋白质加入前后药物信号的变化情况. 这种方法不要求了解蛋白质的结构信息,不需要对蛋白质进行同位素标记,可以对分子量较大的蛋白质进行研究. 虽然这种方法不能给出详细的复合物结构信息,但是更适宜于进行药物的高通量筛选,是较为快速、低廉的手段. 小分子结合蛋白质后,其磁共振参数与大分子趋于一致,如化学位移、信号半峰宽、纵向及横向弛豫速率、自扩散系数等都产生显著变化,二维 NOESY 实验也是判断小分子是否与蛋白质结合的有效方法,它检测的是核间 Overhauser 效应(nuclear overhauser effect, NOE), NOE 反映的是核与核在空间的相对接近关系.

以药物为检测对象要求结合态和自由态小分子之间有快交换现象存在,常用的 NMR 方法包括转移 NOE 法^[34]、饱和转移差谱 STD 法^[35]、WaterLOGSY 法^[36]等. 其中转移 NOE 法灵敏度较低,适于测定离解常数在 100nM 与 1mM 之间的体系. 其中 STD 是研究药物与蛋白相互作用的非常普遍并行之有效的方法之一. 它采用选择性脉冲和 NOE 效应相结合,通过¹H 差谱来研究蛋白与配体的相互作用. 在图 6 中,首先采集混合溶液信号,获得谱图(a);然后选择性照射蛋白信号,这时与蛋白质有相互作用的小分子信号会同时被饱和,而没有相互作用的小分子信号则不受影响,获得谱图(b);谱图(a)和(b)两者的差谱将抵消没有相互作用的小分子信号,获得有相互作用的小分子谱图(c). 这种方法不受蛋白分子量限制,蛋白样品仅需微克量级,获得的谱图简单易懂,能快速判断药物是否与蛋白质有相互作用,受到普遍欢迎,近年来,它被用于检测药物与 8500kDa 人源鼻病毒蛋白 rhinovirus(HRV2)的相互作用^[37]以及药物与 RNA 的相互作用^[38]. 它和魔角旋转(MAS)技术相结合,还可测定多糖与固体麦胚凝集素的结合^[39]等等. WaterLOGSY 法与 STD 法非常类似,只是照射的是溶剂水,信号从水

传递给蛋白,再传递给结合的小分子.

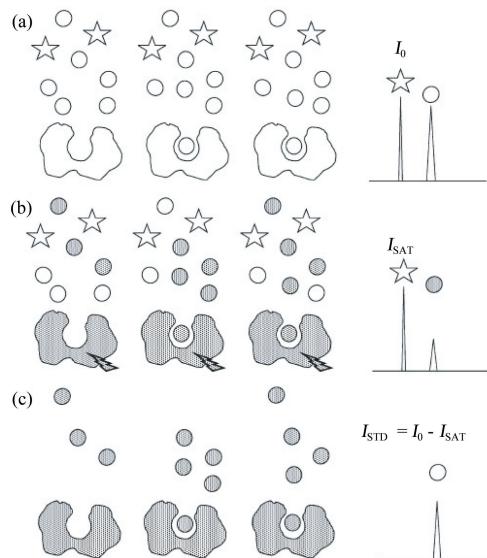


图 6 STD 差谱的示意图 (a)全谱;(b)对蛋白信号选择性照射谱;(c)STD 差谱

3.4 基于 NMR 的代谢组学简介

代谢组学是研究生物体内源性代谢物质的整体及其随内因和外因变化的科学,是系统生物学不可分割的一个重要组成部分. 遗传基因的变化,蛋白质活性的变化及其与环境的相互作用最终会在代谢水平上有所表现. 所以,对整体代谢物的研究包含了宿主基因的变化所导致的机体本身的代谢、寄生菌群的代谢、二者(指机体本身的代谢和寄生菌群的代谢)的共代谢以及二者代谢物质交换引起的变化. 1999 年,在对生物代谢复杂系统进行了近 20 年研究的基础上,Nicholson 等提出了代谢组学(metabolomics)的概念,并将其定义为:对生物系统因病理生理刺激或基因改变所导致的动态多参数代谢应答的定量测定^[40]. 2006 年,Tang 等将代谢组学的概念发展为:代谢组学是关于生物体内源性代谢物质的整体及其变化规律的科学^[41]. 同时提出了代谢组学的中心任务:检测、量化和编录生物内源性代谢物质的整体及其变化规律,揭示该变化规律与所发生的生物学事件(或过程)的本质. 如今,以代谢组学为基础的全局系统生物学(global systems biology)思想已经诞生,而且正处于快速发展阶段,已经在生命科学的许多领域得到应用.

代谢组学经过多年的发展,方法日趋成熟,其应用已经渗入生命科学的研究的诸多方面,并日益彰显出其强有力的科学潜能. 目前代谢组分析主要采用核磁共振和质谱检测的方法,研究对象通常包括体液、粪样、细胞、组织甚至整个生物体. 基于核磁共振

的代谢组学研究包括以下步骤:首先是实验设计和样品收集. 针对要解决的生物学问题,提出实验方案,建立模型. 采集较多的样本数量,以保证研究有一定的统计学意义. 其次是波谱数据采集和处理. 除了采集简单的一维核磁共振谱,还要采集横向弛豫时间加权谱和扩散加权谱,在不分离样品的情况下,选择性地检测小分子代谢物信号和大分子代谢物的信号^[42]. 接着进行数据处理,包括谱图处理、分段积分、数据归一化等步骤. 然后对得到的数据进行非监督方法分析,通常包括主成分分析(PCA)^[43] 和系统聚类分析(HCA)^[44]. 最后建立数据模型,并对建立的模型进行交叉验证. 在此基础上归属变化的代谢物,推断代谢途径中的生物化学反应,将代谢物整体动态变化与生物学问题联系起来.

4 展望

近年来,生物 NMR 发展迅速,新方法和新技术层出不穷,研究领域不断扩展. 如在蛋白质研究中,除了溶液中的蛋白结构和相互作用的 NMR 测定外,在固体和定向介质中,膜蛋白的结构及动力学测定,活细胞中蛋白质的结构研究等,已经成为新的研究热点. 膜蛋白是细胞内生物膜功能的主要体现者,它在各种重要生物学基本过程中起着非常关键的作用. 相对于可溶性蛋白,膜蛋白三维结构的测定困难重重. 在蛋白质数据库(PDB)中收录的 4 万多个结构中,膜蛋白三维结构不到 0.5%. 自第一个用魔角旋转 NMR 得到蛋白质三维结构之后^[45],用魔角旋转 NMR 研究蛋白质这个领域飞速向前发展^[46],显示了魔角旋转 NMR 技术在蛋白质结构研究中的潜力和良好发展前景. 2001 年出现的细胞内 NMR 使得细胞内蛋白质的原位 NMR 研究成为可能^[47]. 细胞内 NMR 的最大优势是原位,所揭示的蛋白质特性不受其他外界因素的影响. 目前这种技术已经在细胞内蛋白质的结构、动力学和相互作用^[48-50] 等的研究中表现出了巨大潜能.

蛋白质核磁共振研究面临的两大挑战是突破分子量上限和提高测定速率. 当蛋白质分子量增大时,共振峰的数目随之增多,谱峰重叠的几率相应增加,需要更高的分辨率加以区分. NMR 谱仪硬件技术和实验技术不断面临着挑战,高灵敏度和高分辨率 NMR 实验方法研究一直备受关注,并期待出现重大突破. 本文所列举的仅仅是核磁共振技术在蛋白质科学中的部分应用,探索生命奥秘的内在驱动力

将不断推进核磁共振实验技术向纵深发展,在生命科学中发挥更重要的作用。

活体波谱和成像是生物核磁共振研究中与生命科学,特别是与脑科学和重大疾病的临床诊断密切相关的前沿领域。基于超极化¹²⁹Xe磁共振波谱和成像及其在生物医学中的应用研究是近10年来发展极其迅速的领域,本期另外2篇专题文章对此进行了论述^[51,52],这里不再赘述。

参考文献

- [1] Williamson M P, Havel T F, Wuthrich K. *J. Mol. Biol.*, 1985, 182: 295
- [2] Pervushin K, Riek R, Wider G et al. *Proc. Natl. Acad. Sci. U S A*, 1997, 94: 12366
- [3] Riek R, Wider G, Pervushin K et al. *Proc. Natl. Acad. Sci. U S A*, 1999, 96: 4918
- [4] Fiaux J, Bertelsen E B, Horwich A L et al. *Nature*, 2002, 418: 207
- [5] Tjandra N, Bax A. *Science*, 1997, 278: 1111
- [6] Korzhnev D M, Salvatella X, Vendruscolo M et al. *Nature*, 2004, 430: 586
- [7] Barbieri R, Luchinat C, Parigi G. *Chem. Phys. Chem.*, 2004, 21: 797
- [8] Bertini I, Luchinat C, Piccioli M. *Methods Enzymol.*, 2001, 339: 314
- [9] Szyperski T, Yeh D C, Sukumaran D K et al. *Proc. Natl. Acad. Sci. U S A*, 2002, 99: 8009
- [10] Kim S, Szyperski T. *J. Am. Chem. Soc.*, 2003, 125: 1385
- [11] Kupce E, Freeman R. *J. Biomol. NMR*, 2003, 27: 383
- [12] Kupce E, Freeman R. *J. Magn. Reson.*, 2003, 163: 56
- [13] Stern A S, Li K B, Hoch J C. *J. Am. Chem. Soc.*, 2002, 124: 1982
- [14] Tugarinov V, Kay L E, Ibraghimov I et al. *J. Am. Chem. Soc.*, 2005, 127: 2767
- [15] Kazimierczuk K, Zawadzka A, Kozminski W. *J. Magn. Reson.*, 2009, 197: 219
- [16] Jiang B, Jiang X, Xiao N et al. *J. Magn. Reson.*, 2010, 204: 165
- [17] Tugarinov V, Kay L E. *J. Am. Chem. Soc.*, 2003, 125: 13868
- [18] Kainosho M, Torizawa T, Iwashita Y et al. *Nature*, 2006, 440: 52
- [19] Ohki S-y, Kainosho M. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 2008, 53: 208
- [20] Petri W, Griesinger C. *J. Am. Chem. Soc.*, 2000, 122: 3975
- [21] Tjandra N, Omichinski J G, Gronenborn A M et al. *Nat. Struct. Biol.*, 1997, 4: 732
- [22] Snyder D A, Xu Y, Yang D et al. *J. Am. Chem. Soc.*, 2007, 129: 14126
- [23] Xu Y, Zheng Y, Fan J S et al. *Nat. Methods*, 2006, 3: 931
- [24] Goldman M. *J. Magn. Reson.*, 1984, 60: 437
- [25] Takahashi H, Nakanishi T, Kami K et al. *Nat. Struct. Biol.*, 2000, 7: 220
- [26] Ikura M, Bax A. *J. Am. Chem. Soc.*, 1992, 114: 2433
- [27] Sprangers R, Kay L E. *Nature*, 2007, 445: 618
- [28] Wang W, Weng J, Zhang X et al. *J. Am. Chem. Soc.*, 2009, 131: 787
- [29] Sugase K, Dyson H J, Wright P E. *Nature*, 2007, 447: 1021
- [30] Korzhnev D M, Religa T L, Banachewicz W et al. *Science*, 2010, 329: 1312
- [31] Tang C, Schwieters C D et al. *Nature*, 2007, 449: 1078
- [32] Shuker S B, Hajduk P J, Meadows R P et al. *Science*, 1996, 274: 1531
- [33] Peng J W, Moore J, Abdul-Manan N. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 2004, 44: 225
- [34] Vaynberg J, Qin J. *Trends in Biotechnology*, 2006, 24: 22
- [35] Mayer M, Meyer B. *Angew. Chem. Int. Ed.*, 1999, 38: 1784
- [36] Dalvit C, Pevarello P, Tato M et al. *J. Biomol. NMR*, 2000, 18: 65
- [37] Benie A J, Moser R, Baumi E et al. *J. Am. Chem. Soc.*, 2003, 125: 14
- [38] Mayer M, James T L. *J. Am. Chem. Soc.*, 2002, 124: 13376
- [39] Klein J, Meinecke R, Mayer M et al. *J. Am. Chem. Soc.*, 1999, 121: 5336
- [40] Nicholson J K, Lindon J C, Holmes E. *Xenobiotica*, 1999, 29: 1181
- [41] Tang H R, Wang Y L. *Prog. Biochem. Biophys.*, 2006, 33: 401
- [42] Liu M, Nicholson J K, Lindon J C. *Anal. Chem.*, 1996, 68: 3370
- [43] Wold S, Esbensen K P G. *Chemometr. Intell. Lab.*, 1987, 2: 37
- [44] Ciosek P, Brzoka Z, Wroblewski W et al. *Talanta*, 2005, 67: 590
- [45] Castellani F, van Rossum B, Diehl A et al. *Nature*, 2002, 420: 98
- [46] McDermott A E. *Curr. Opin. Struct. Biol.*, 2004, 14: 554
- [47] Serber Z, Dotsch V. *Biochemistry*, 2001, 40: 14317
- [48] Bryant J E, Lecomte J T, Lee A L et al. *Biochemistry*, 2005, 44: 9275
- [49] Burz D S, Dutta K, Cowburn D et al. *Nat. Methods*, 2006, 3: 91
- [50] Selenko P, Wagner G. *Nat. Methods*, 2006, 3: 80
- [51] 李博, 吴瑞琪, 李安安等. 物理, 2011, 40(6): 374 [Li B, Wu R Q, Li A A et al. *Wuli (Physics)*, 2011, 40(6): 374 (in Chinese)]
- [52] 孙献平, 韩叶清, 罗晴等. 物理, 2011, 40(6): 381 [Sun X P, Han Y Q, Luo Q et al. *Wuli (Physics)*, 2011, 40(6): 381 (in Chinese)]