

核磁共振波谱研究蛋白质三维结构及功能

施蕴渝, 吴季辉

(中国科学技术大学生命科学学院, 合肥微尺度物质科学国家实验室, 安徽合肥 230027)

摘要:文中总结了中国科学技术大学生命科学学院核磁共振波谱实验室十多年来的工作。我们的研究主要集中于研究人和其他真核生物基因表达调控相关蛋白质以及细胞连接处相关蛋白质。在这两个体系中许多蛋白质与人类健康及疾病相关, 有的可能是潜在的 药物作用靶标。我们主要关注用核磁共振波谱方法研究蛋白质-蛋白质相互作用的结构基础。核磁共振适合研究在接近生理条件下的分子相互作用, 特别是适合研究低亲和力的瞬态的复合物。它可以提供蛋白质相互作用界面, 复合物结构, 以及蛋白质相互识别过程动力学的信息。文中给出了一些例子。我们也研究蛋白质内部动力学, 包括皮秒-纳秒时间尺度, 与毫秒-微秒时间尺度的动力学。与圆二色谱及荧光光谱结合, 核磁共振可以详细表征蛋白质的折叠与去折叠。文中给出的核磁共振应用的最后一个例子是用计算机虚拟筛选, 核磁筛选, 我们发现了一个人的双功能的磷酸酶的一种新类型的抑制剂, 并研究了该抑制剂对细胞功能的影响。这一策略有可能用于早期药物的发现。

关键词:核磁共振波谱; 蛋白质-蛋白质相互作用; 蛋白质结构; 蛋白质动力学; 折叠与去折叠; 配基筛选

中图分类号: Q71

文献标识码: A

Protein structure and function studied by NMR

SHI Yun-yu, WU Ji-hui

(School of Life Science, and Hefei National Laboratory for Physical Sciences at Microscale, University of Science and Technology of China, Hefei 230027, China)

Abstract: More than 10 years of research in the Laboratory of Nuclear Magnetic Resonance (NMR) at the School of Life Science, University of Science and Technology of China is reviewed. Our researches have focused on two systems; proteins related to the regulation of gene expression in humans and other eukaryotes, and proteins existing in the cell junction in humans. The majority of proteins selected from these two systems are related to human health and diseases, and some are potential drug targets. We were interested in using NMR to study structural basis of protein-protein interactions. NMR was highly suited

收稿日期: 2008-06-30; 修回日期: 2008-07-10

基金项目: 国家重点基础研究发展(973)计划(2002CB713806, 2006CB806507, 2006CB910201, 2006AA02A315)和国家自然科学基金(30121001, 30570361, 30670426)资助。

作者简介: 施蕴渝(通讯作者), 中国科学院院士/教授。1965年毕业于中国科学技术大学物理系生物物理专业。1965~1970年在卫生部中医研究院担任实习研究员, 1970年至今在中国科学技术大学任教, 期间曾分别在意大利罗马大学及 CNRS 结构化学研究所、荷兰格罗宁根大学、法国 CNRS 结构生物学实验室和理论化学实验室进修或合作研究。现任中国科学院生物与医学学部常委, 教育部高等学校生物科学与工程教学指导委员会主任委员, 中国科学技术大学校学术委员会副主任。在 PNAS, J. Biol. Chem., Biochemistry 等国内外学术刊物发表论文 110 余篇。主要用多维核磁共振波谱及计算生物学研究与重大疾病或重要生理功能相关的蛋白质结构, 动力学与功能关系, 以及蛋白质相互作用。E-mail: yyshi@ustc.edu.cn
吴季辉, 男, 1956年生, 教授。研究方向: 分子生物工程。E-mail: wujihui@ustc.edu.cn

for investigating molecular interactions under approximately physiological conditions and was particularly suited for the study of low-affinity, transient complexes. It can provide information of protein interaction surface, complex structure, and dynamic properties during protein recognition. Several examples were given in this paper. NMR was also used to study dynamic properties of protein both in pico-second to nano-second and in micro-second to mili-second time scales. We have studied protein folding and unfolding by NMR together with fluorescence and circular dichroism experiments. Proteins in unfolded states were characterized in detail by NMR. The last example of NMR application is the identification of a novel inhibitor of a human dual-specific phosphatase and the cellular effects of this compound were also studied. Our results demonstrate that our screening strategy, which combines both virtual and NMR-based methods, is feasible and might be employed in the early stage of drug discovery.

Key words: NMR spectroscopy; protein-protein interaction; protein structure; protein dynamics; folding and unfolding; ligand screening

0 引言

X 射线晶体衍射,核磁共振波谱,电镜三维重构是结构生物学的三种主要研究方法.至今,核磁共振波谱技术是能够在原子分辨率下测定溶液中生物大分子三维结构的唯一方法.过去认为核磁共振只适合研究分子量较小的蛋白质的结构,近年来由于核磁共振波谱技术以及标记方法的发展,核磁共振方法可以测定的蛋白质的分子量,从原理上将不受限制.目前可以测量 40~50 kD 分子量的蛋白质.而且这个限制正被突破,用核磁共振波谱方法测定结构的最大的单链蛋白质的分子量已达到 82 kD^[1].核磁技术不仅仅局限于结构测定,它在生物学中的应用还包括如下方面:

(I) 研究生物大分子及其复合物在溶液中的三维结构和功能;

(II) 研究动态的生物大分子之间以及与配基的相互作用;

(III) 研究生物大分子的动态行为;

(IV) 用固体核磁共振或液体核磁共振技术研究膜蛋白的结构与功能;

(V) 研究蛋白质折叠,折叠动力学;

(VI) 用于药物筛选与设计;

(VII) 研究代谢组学;

(VIII) 研究活细胞中的蛋白质-蛋白质相互作用;

(IX) 核磁成像用于认知科学研究.

中国科学技术大学生命科学学院生物核磁共振波谱实验室在过去十多年来建立了用核磁共振波谱技术研究蛋白质结构与动力学的实验平台,开展基因克隆、表达,蛋白质分离、纯化、性质鉴定,核磁共

振波谱实验,结构解析,动力学,以及蛋白质-蛋白质相互作用,蛋白质与配基相互作用研究.重点研究真核基因表达调控相关蛋白质,细胞连接处相关蛋白质,以及蛋白质-蛋白质相互作用的三维结构基础和动力学.其中许多蛋白质具有重要生理功能,与疾病相关.一些蛋白质是潜在的药物作用靶标.

1 研究蛋白质-蛋白质相互作用的三维结构基础

研究蛋白质-蛋白质,蛋白质-核酸,蛋白质-配基相互作用对于生物学功能研究十分重要.核磁共振特别适合从结构生物学角度研究蛋白质-蛋白质相互作用,确定蛋白质相互作用界面.酵母双杂交、GST-pull down、免疫共沉淀或者 far western blot,电泳迁移实验,表面等离子共振谱(SPR)、等温量热(ITC),超离心沉降平衡,噬菌体展示,动态光散射,荧光去偏振等实验可以在离体情况下研究蛋白质相互作用;荧光共振能量转移的显微镜分析,荧光共振能量转移的流式细胞仪分析,荧光相关谱,激光共聚焦可研究细胞内的蛋白质共定位;蛋白质交联等方法可研究活细胞中的蛋白质相互作用;除此而外还有使用二维凝胶或二维液相色谱与同位素标识和质谱结合的蛋白质组学方法;生物信息学方法可以研究和预测蛋白质相互作用.但是如果需要知道二者相互作用的界面,回答两个蛋白质相互作用涉及的氨基酸残基,结构的信息是必不可少的.细胞中蛋白质相互作用是动态的,低亲和力的(解离常数 K_d 常常大于 10^{-4} M),许多复合物瞬时存在,不稳定.对于这种动态的复合物,很难解出晶体结构,而核磁共振波谱技术特别适合研究瞬时存在,动态的复合物.

通过滴定实验,对化学位移扰动、可以在接近生理条件下确定蛋白质相互作用界面.该方法十分灵敏, K_d 为 10^{-2} M 的非常弱的蛋白质-蛋白质相互作用也能被检测^[2,3].

1.1 真核基因表达调控相关蛋白质的三维结构与功能,以及分子相互识别的蛋白质三维结构基础

真核基因表达调控是生物学中的重大前沿问题,细胞分化,细胞重要生理功能的调控,肿瘤及许多其他疾病发生都与真核基因表达调控密切相关.真核基因表达调控发生在各个水平,包括转录,mRNA 剪切,翻译等.近年来 ncRNA 的发现,表明 ncRNA 在真核基因表达调控中起重要作用.在过去几年中,对于上述体系,我们开展了如下系列研究:

(I) 与表观遗传调控相关的蛋白质相互识别的结构生物学基础

组蛋白的翻译后修饰是表观遗传的重要内容之一,它建立并维持基因表达的顺序.组蛋白修饰的错误会引起发育异常和肿瘤.组蛋白的修饰(乙酰化、甲基化、磷酸化、SUMO 化、泛素化等)及其组合构成了组蛋白密码.这些密码可以被含有特定结构域(如 Bromodomain 和 Chromodomain 结构域)的蛋白解读,从而启动下游生物学反应如染色质凝聚、DNA 修复或转录激活/抑制. Bromodomain 是一类与乙酰化组蛋白特异识别的蛋白质结构域.含有 Bromodomain 的蛋白质包括三大类:第一类是依赖 ATP 的染色质重塑复合物,它们利用 ATP 水解的能量局部破坏或改变组蛋白与 DNA 的结合;第二类是组蛋白乙酰转移酶(acetyltransferase, HAT),组蛋白去乙酰酶(deacetylase, HDAC),它们通过决定核小体组蛋白 N 端乙酰化水平调节基因转录;第三类是 BET 家族的蛋白质. BET 家族蛋白的共同特点是 N 端包括两个 Bromodomain, C 端有一个 ET 结构域.其中 Bromodomain 与乙酰化的组蛋白结合,而 ET 结构域参与其他的蛋白质-蛋白质相互作用. BET 家族蛋白与组蛋白的相互作用在细胞分裂期仍然存在,而此时大部分转录调节因子都与高度浓缩的染色体分离.

在过去几年中,我们分别测定了人的 Brg1, Brd2, Brd4, Brd7 的 Bromodomain 结构域的三维结构,研究了它们与组蛋白的相互识别.它们的 Bromodomain 都形成左手四螺旋束,在分子的一端由 ZA-loop 和 BC-loop 上的疏水残基构成一个疏水

口袋,提供与乙酰化赖氨酸侧链结合的界面.另外 Brd2, Brd4 在 ZA-loop 上有额外的二级结构单元,可能与它们对乙酰化小肽的选择特异性相关.

其中 Brg1 是人的染色质重塑复合物 SWI/SNF 复合物的核心催化亚基,它利用 ATP 水解的能量重塑染色质,帮助转录因子和 RNA 聚合酶 II 在基因的启动子区结合.因此, Brg1 在细胞周期调控,细胞分化和肿瘤发生中发挥重要作用.我们发现 Brg1 的 Bromodomain 与 H4AcK12 具有较强的结合能力,并确定二者结合界面,构建复合物模型;通过突变实验确定了结合界面的关键残基^[4]. 论文已为综述文章 Oncogene(2007)引用^[5].

Brd2 及 Brd4 属于 BET 家族,二者的 Bromodomain 特异识别乙酰化组蛋白. Brd2 可以通过 ET domain 与 Kaposi 肉瘤病毒的核心抗原 LANA 相互作用,并与其他蛋白形成复合物维持 Kaposi 病毒基因组在细胞中的稳定性. Brd2 还可以通过 ET domain 与 E2Fs 形成复合物,进而激活一些 E2Fs 调控的细胞周期调节基因. Brd4 的 CTD 结构域 C 末端的 20 个氨基酸残基组成的肽段直接与乳突淋巴瘤病毒 E2 蛋白 N 端的反式激活结构域(1-201)相互作用, Brd4 还可以与 Kaposi 肉瘤相关疱疹病毒相互作用. Brd4 可以与 P-TEFb (positive transcription elongation factor b) 结合,激活 P-TEFb. 上述 Brd2 及 Brd4 的功能均与它们 Bromodomain 特异识别乙酰化组蛋白有关. 我们发现 Brd2 第二个 Bromodomain 对乙酰化组蛋白 H4-AcK12 具有较强的结合能力, $K_d = 2.9$ mM. Brd4 第二个 Bromodomain 对乙酰化组蛋白 H4-AcK5/K12, H3-AcK9/K14 具有较强的结合能力, K_d 均为 mM 量级. 我们还分别确定了它们的结合界面,并通过突变实验确定了结合界面的关键残基^[6,7].

Brd7 蛋白是含有 bromodomain 的人源蛋白,它的功能与鼻咽癌密切相关. 我们应用溶液核磁共振方法测定了 BRD7 蛋白 bromodomain 的结构,并通过化学位移扰动的方法研究了它与来自于组蛋白 H3 和 H4 的多个乙酰化位点的结合,发现 BRD7 蛋白 bromodomain 能与 H3 上 K9, K14 以及 H4, K8, K16 等乙酰化位点结合,并得到了它们的解离常数^[8].

(II) 与前体 mRNA 剪切过程相关的蛋白质相互识别的结构生物学基础

前体 mRNA 剪切过程对于绝大多数的真核基

因表达都是不可缺少的. 前体 mRNA 的剪切在真核生物体内由剪切体复合物负责.

已经证明 SKIP (spliceosomal component and transcriptional coregulator) 在剪切体活化激活过程中加入 45S snRNP, 参与了剪切体活化过程. 在这过程中 SKIP 募集 PPIL1. PPIL1 是脯氨酸顺反异构酶超家族中 cyclophilin 家族的成员. 我们用多维核磁共振波谱测定了 PPIL1 的三维结构. 用 GST pull-down 和 Western blot 实验, 证实 PPIL1 与剪切体蛋白质 SKIP N 端(残基 59-129, SKIPN) 可以结合, 其结合常数为 10^{-7} mM 量级. 用化学位移扰动实验确定了 PPIL1 与剪切体蛋白质 SKIP 的 SKIPN 的结合界面, 这是一个全新的蛋白质结合界面. 脯氨酸顺反异构酶抑制剂 cyclosporine A 的加入不影响 SKIP 结合, 这证实 PPIL1, SKIP 和 cyclosporine A 可以形成三元复合物^[9,10]. 核磁谱及 CD 光谱, 差示扫描微量热(DSC) 结果表明, 自由状态下的 SKIPN 是一个天然无结构的蛋白 motif, 当 SKIPN 与 PPIL1 结合后, SKIPN 的前 21 个氨基酸(残基 59-79, G21K) 转变为有结构的状态. 我们用 NMR 的方法测定了 PPIL1-G21K 复合物的溶液结构. 结果表明, 复合物状态下和自由状态下 PPIL1 的结构类似, 二级结构及其走向没有改变, PPIL1 上与 G21K 结合的氨基酸在整体上有些移动. 在复合物状态下的 G21K 主链走向呈现一个“钩状”的结构, G21K 可以分为三个部分, 其中两个部分可以观测到 NOE. 复合物的结构表明, PPIL1 与 G21K 的结合主要依靠疏水相互作用和静电相互作用. NMR 滴定和突变研究推测, G21K 上 Glu 可能与 PPIL1 上 Arg 的胍基形成盐桥, 此相互作用对于维持 PPIL1-G21K 的结构起到关键的作用. PPIL1-SKIP 复合物结构的测定以及突变的研究为进一步阐述剪切体的激活提供了结构的基础.

(III) 蛋白质修饰物的三维结构基础, 以及它们与相关蛋白质的相互识别

SUMO 是在所有真核生物中高度保守的蛋白质修饰物, 具有极为重要的生理功能. 我们测定了人的 SUMO3 的三维结构, 并用化学位移干扰方法研究了它和酶 2Ubc9 的相互作用界面, 通过 HADDOCK 方法得到复合物的结构模型^[11,12].

Urm1 是另一个蛋白质修饰物, 在真核生物中高度保守. 我们用 NMR 测定了酵母 Urm1 在溶液中的三维结构, 对泛素超家族的蛋白质三维结构比

较和分子进化树分析. 我们发现它虽然具有蛋白质修饰物的功能, 但是无论是序列, 还是结构都更接近原核生物中的 MoeB. 分子静电表面, 疏水表面分析表明 Urm1-Uba4 相互作用与 MoeB-MoeB 相互作用的相似性. Urm1 是连接真核生物中的蛋白质修饰物与原核生物中载硫蛋白的分子, 它在泛素超家族中十分保守, 保留了泛素及类泛素蛋白质修饰物共同祖先的结构特征, 可以被称为分子化石. 我们的结构测定为泛素及类泛素蛋白质修饰物起源于原核生物中载硫蛋白的假说提供了有力支持^[13].

(IV) 具有新的折叠类型的一个新的锌指蛋白酵母 Kti11p 的三维结构的测定

Kti11p 是真核生物中高度保守的含有 CSL 锌指的蛋白质. 它最早是因影响 *Kluyveromyces lactis* 毒素 Zymocin 的功能而被发现的. Zymocin 的靶位点是一个 7 亚基蛋白复合物 TOT, 其中的 6 个亚基被鉴定为 RNAP II 的转录延伸因子. Kti11p 通过调节转录延伸因子的功能来影响 RNAP II 的转录. Kti11p 在鼠中的同源物为 Dph3p 与翻译延伸因子 2(EF-2) 翻译后修饰即白喉酰胺的生物合成有关. 我们用多维核磁共振波谱解出了 Kti11p 的三维结构, 并用 EXAFS 方法确定锌离子与 4 个 Cystidin 配位. Kti11p 是 CSL 锌指蛋白质家族解出结构的第一个成员, 是一种新的折叠, 属于一种新的锌指结构^[14]. 我们还测定了 Kti11p 人的类似蛋白 DESR1 的三维结构^[15].

1.2 细胞紧密连接与黏性连接处蛋白质的结构功能, 以及分子相互识别的蛋白质三维结构基础

多细胞生物的细胞连接, 包括紧密连接, 黏性连接, 桥粒和间隙连接. 多细胞生物的细胞连接, 对于保持组织的完整性, 对于细胞极化, 迁移, 形态发生, 突触(神经突触, 免疫突触) 形成具有重要的意义, 也与发育生物学及肿瘤的浸润转移关系密切. 肿瘤的浸润转移实质是上皮细胞-间质细胞的转变 (epithelial-mesenchymal transition or EMT). 在此转变过程中细胞结构改变, 失去极性, 以及失去与邻近细胞的接触, 获得运动性. 细胞黏附分子的抑制剂有可能发展成阻止肿瘤转移的药物. 细胞紧密连接, 间隙连接与离子, 小分子通过, 与血脑屏障有关. 细胞连接不仅仅是物理连接, 起到保持组织完整性的目的, 而且也是细胞信号转导的中心. 在细胞连接处有许多蛋白质和蛋白质复合物, 包括跨膜蛋白、adaptor 分子、转录因子、调节蛋白等, 涉及许多信号

传导通路. 它们之间彼此相互作用, 保持细胞连接的完整功能. 我们重点研究了近年来新发现的细胞-细胞黏附系统 Nectin-AF-6/afadin 系统, 以及 claudin-ZO 蛋白(ZO-1, ZO-2, ZO-3)系统.

(I) AF-6 蛋白质的 PDZ 结构域的溶液结构, 以及它与 neurexin 及 Bcr 的相互作用

人 AF-6 基因定位在染色体 6q27, 染色体易位形成的 MLL-AF6 融合蛋白出现在急性髓系白血病中. AF-6 属于细胞紧密连接与黏性连接复合体的组成部分. 它也是 RAS 的效应蛋白, 而且受到 Ser/Thr 蛋白激酶 Bcr 的调控. AF-6 的 C 端含有一个 PDZ 结构域, 该结构域涉及 AF-6 与其他许多蛋白质的相互作用. 我们测定了 AF-6 的 PDZ 结构域的三维结构, 用 ITC 和 BIACORE 以及核磁共振波谱研究了它与 neurexin 及 Bcr 的相互作用, 说明分子相互识别的专一性^[16], 进一步利用同位素滤波技术获得了蛋白质复合物界面 NOE, 解出了复合物的三维结构. 结果表明 AF-6 的 PDZ 结构域兼有第一类与第二类 PDZ 的特征. 研究结果提示有必要对 PDZ 分类及现有机理进行更深入的研究^[17].

(II) ZO 蛋白质结构与功能的研究

ZO(Zonula Occludens proteins)蛋白质(ZO1, ZO2, ZO3)是一种存在于细胞紧密连接, 黏性连接和间隙连接中的组成蛋白. 它们与膜蛋白及细胞骨架蛋白都有联系.

前人已证明 ZO1 可以通过第二个 PDZ 结构域形成同二聚体或者与 ZO2、ZO3 形成异二聚体, 然而其潜在的结构基础尚不清楚. 我们通过 GST pull-down 和免疫共沉淀实验证明 ZO1-PDZ2 和 ZO2-PDZ2 同二聚体与异二聚体的形成. 我们用 NMR 方法解析了 ZO2-PDZ2 的溶液结构, 揭示 PDZ 结构域形成了一种新颖的二聚模式——three-dimensional (3D) domain swapping. 这是首次在 PDZ 结构域中发现的 domain swapping 结构. 我们通过分析提出 ZO 蛋白 PDZ 结构域 domain swapping 结构形成的原因, 并通过实验得到证实. 我们推测此类结构可推广至 ZO1-PDZ2、ZO3-PDZ2 的同二聚体以及 ZO1-PDZ2/ZO2-PDZ2 ZO1-PDZ2/ZO3-PDZ2 异二聚体. 最近 ZO1-PDZ2 同二聚体的晶体已被解出, 我们实验室也解析了 ZO1-PDZ2 同二聚体的核磁结构, 证明了我们的推测.

化学交联及动态光散射实验还揭示 ZO1-PDZ2 和 ZO2-PDZ2 在溶液中均可形成寡聚体, 这种 ZO

由 PDZ 介导的寡聚化可能为 Claudins 的聚合, 亦即紧密连接的形成提供了结构基础^[18].

(III) GOPC 蛋白质的 PDZ 结构域的三维结构以及它与 neuroligin 的相互作用

人 GOPC (Golgi-associated PDZ- and coiled-coil motif-containing protein)是第一个被报道的含有 PDZ 结构域的高尔基体囊泡运输相关蛋白. GOPC 包含两个 coiled-coil 区域和一个 PDZ 结构域. 我们运用 NMR 方法测定了人 GOPC 蛋白质 PDZ 结构域的溶液结构, 结果显示人 GOPC 蛋白质 PDZ 结构域具有典型的 PDZ 结构域特征, 属于典型的第一类 PDZ 结构域; NMR 化学位移扰动实验表明, hGOPC-PDZ 能够与 $\beta 1$ 肾上腺素受体和神经胶质素 neuroligin C-末端相互作用. 我们确定了它们之间的相互作用界面, 它们之间的作用属于弱相互作用(解离常数为 10^{-4} M); 结合分子动力学模拟的方法, 我们构建了 GOPC PDZ 结构域/neuroligin C-terminal 多肽的复合物模型. 从细节上提供了结合位点上残基间的相互作用模式以及成键方式. 我们还研究了它与 Frizzles 8 的相互作用^[19].

2 蛋白质的动力学研究

蛋白质的柔性及运动性与功能有密切关系. 例如酶的柔性及运动性对于其功能态的热力学稳定性及催化功能十分重要. 由于蛋白质是一个残基间相互耦联的动力学系统, 配基结合将会引起信号在蛋白质内部传递, 引起别构效应. 核磁共振特别适合研究蛋白质的动力学, 得到的动力学信息可以具体确定到所研究的基团. 通过自旋-晶格弛豫时间, 自旋-自旋弛豫时间, 以及 NOE 的测量可以获得皮秒-纳秒的动力学信息. 近年来还发展了测定微秒-毫秒的动力学的方法——RIP, CPMG 及 relaxation dispersion experiment^[20~22].

(I) AF-6 PDZ 结构域自由态, 以及与 Bcr C 末端肽结合过程的动力学

我们将 AF-6 PDZ 结构域自由态的结构与复合物的结构进行比较, 发现远离肽结合沟, 相距一定距离的表面 2 具有构象重排. 通过 relaxation dispersion 实验检测 AF-6 的 PDZ 结构域在结合 Bcr 前后毫秒时间尺度的化学交换, 我们发现与 Bcr C 末端肽的结合引起变构信号在 PDZ 结构域内部的动力学网络中传递, 从而引起远离肽结合沟, 相距一定距离的表面毫秒时间尺度的动力学变化. 这些

变化不能以简单的二态模型加以解释^[23]。

(II) LARG 蛋白质 PDZ 结构域在自由和复合物状态下结构和动力学的研究

白血病相关的 Rho 鸟苷酸交换因子(LARG)特异性地激活 RhoA,在白血病的形成过程中发挥着重要作用. LARG PDZ 结构域与 plexin B1 末端序列的相互作用可以激活 RhoA,将胞外信号转化为胞内信号,调控细胞骨架的重排.到目前为止,没有关于两者相互作用的结构和动力学方面的信息.我们通过核磁共振波谱(NMR)方法解析了 LARG PDZ 在自由和复合物状态下的溶液结构. LARG PDZ 和 plexin B1 末端小肽形成复合物以后,结构发生了较为明显的变化,特别是 $\beta\text{B}/\alpha\text{B}$ 结合沟的底端部分;这种构象变化与其他 PDZ 在结合小肽前后发生构象变化的部位是不一样的^[24].为了寻找 LARG PDZ 结构变化以及柔性区域氨基酸残基信号消失的原因,我们探索了 LARG PDZ 的动力学性质.通过简化的谱密度函数来分析主链弛豫数据,自由状态的 LARG PDZ 显示出较大幅度 and 范围的内部运动,包括皮秒-纳秒时间尺度和微秒-毫秒时间尺度的运动. $\beta\text{B}/\beta\text{C}$ loop 和 $\beta\text{E}/\alpha\text{B}$ loop 区域的氨基酸残基在两种构象或者多种构象之间进行交换,引起谱宽增宽而导致信号消失.通过 CPMG 实验,我们分析了处于自由状态、低小肽浓度状态(蛋白质:小肽浓度比为 10 : 1)和处于复合物状态的 LARG PDZ 的构象交换对于横向弛豫速率的贡献情况,以及构象交换的速率.在自由状态下,蛋白质的大部分氨基酸残基显示出毫秒量级的内部运动,它们主要集中在 $\beta\text{B}/\alpha\text{B}$ 结合沟上或者周围.在加入少量小肽之后,毫秒量级的运动更加明显,结合沟附近的氨基酸残基的构象交换对于横向弛豫的贡献更加明显,交换速率降低.加入过量的 plexin B1 末端小肽使 99% 以上的蛋白质处于结合状态,蛋白质和小肽之间的相互作用限制了蛋白质运动的空间自由性,没有明显的毫秒量级的内部运动.

突变和等温滴定热(ITC)实验证实了 $\beta\text{B}/\beta\text{C}$ 和 $\beta\text{E}/\alpha\text{B}$ loop 区域的氨基酸残基的构象会影响 LARG PDZ 和小肽之间的相互作用.暗示了 LARG PDZ 动力学和结构变化之间的关系; $\beta\text{B}/\beta\text{C}$ loop 和 $\beta\text{E}/\alpha\text{B}$ loop 区域在不同构象之间的交换以及构象的易变性可以较好的辅助 $\beta\text{B}/\alpha\text{B}$ 结合沟进行构象的变化而容纳配体.在结合沟附近的氨基酸残基经历着明显的毫秒量级的内部运动,这种构象的柔性

能够很好地帮助蛋白质进行不同的构象变化来结合不同的小肽.

3 研究蛋白质折叠与去折叠

无结构的蛋白质(disordered proteins),包括去折叠状态的蛋白质,熔融球蛋白,折叠中间态,变性蛋白质,还包括内源的天然无结构的蛋白质.这些无结构的蛋白质对于了解蛋白质的折叠,蛋白质的聚合,蛋白质的纤维化十分重要.它们与淀粉状纤维变性病如 Prion 病,帕金森氏病,埃茨海默病等疾病有关.目前估计有三分之一的真核生物蛋白质含有超过连续 30 个残基组成的无结构的序列.与肿瘤相关的信号蛋白质这个比例还要高.一些内源性无结构的蛋白质在蛋白质-蛋白质相互作用过程中起了关键作用.在蛋白质结合过程中会发生无序向有序结构的转换. NMR 是少数几种能提供无结构或有部分结构的蛋白质以及蛋白质折叠过程的多方面信息的方法^[25].非天然态蛋白质分子的 NMR 研究对科学家们来说是个挑战.去折叠和部分折叠的蛋白多肽链在多种构象之间互相转换,再加上其内在的活性,使得化学位移分散度很差,共振峰的认证非常困难.但是,最近在 NMR 技术上的进展,使去折叠和部分折叠蛋白质分子的主链共振峰能得到完全的认证,进而使细致地了解这些构象状态的结构和动力学特征成为可能.可以根据所谓的 secondary chemical shifts,得到蛋白质二级结构的信息.脉冲场梯度核磁扩散实验,可以得到平移扩散系数,并由此计算蛋白质的流体力学半径.由于寻常的 NOE 效应限于 5 Å 范围,所以在无结构的蛋白质中 NOE 常常只在残基内部被观察到.用顺磁探针标记,通过顺磁弛豫增强效应,可以得到 20~25 Å 长程信息.残存的偶极-偶极相互作用,原理上也可以用于无结构的蛋白质或者部分非折叠态的蛋白质,给出主链拓扑的有用信息.运用上述方法可以获得折叠中间态或熔融球蛋白的结构信息.

人上游结合因子(human upstream binding factor, hUBF)是一种在 RNA 聚合酶 I 催化的转录激活过程中扮演重要角色的核转录因子.它包含六个 HMG Box(high mobility group box)结构域,其中第一个(hUBF HMG Box1)能特异性地单独与核糖体启动子结合并对整个蛋白质分子与 DNA 的特异性结合起关键作用^[1~3]. hUBF HMG Box1 和 HMG Box5 已在本实验室克隆、表达和纯化,其溶

液三维空间结构也用 NMR 方法得到了确定^[26,27], 它们含有大约 90 个氨基酸残基, 其天然结构是一个由三段 α 螺旋组成的扭曲的 L 形, 三段螺旋中间是一个由它们包围而成的疏水核心, 该疏水核心的稳定性由其周边非极性残基侧链广泛的疏水相互作用所维持. 我们对 hUBF HMG Box1 的平衡去折叠进行了系统研究, 圆二色谱 (circular dichroism, CD)、ANS (1-anilino-8-naphthalene sulfonate) 荧光光谱和核磁共振 (nuclear magnetic resonance, NMR) 等方法被用来监测 hUBF HMG Box1 在盐酸胍、热和酸诱导的去折叠过程中发生的结构变化. 结果显示, 当盐酸胍浓度增加, hUBF HMG Box1 分子协同性地发生去折叠, 没有明显的中间体出现. 然而, 在酸滴定过程中, 蛋白质分子首先在 pH 2 左右最大限度地去折叠形成 U (unfolded) 态, 然后在 0.2 和更低的 pH 下复折叠形成一个紧密的部分折叠态 (A 态). 该局部折叠态具有某些典型熔球态的特征^[28,29]. 对于 hUBF HMG Box5, 通过圆二色谱、荧光光谱、差示扫描量热 (DSC) 实验研究了温度改变的去折叠与重折叠. 实验表明不能用简单的协同的两态模型来说明. 我们发现 Box-5 蛋白质的热变性过程可以被描述成两个不同的转变过程, 而且暗示了在 55 °C 附近存在一个部分折叠的熔球态 (molten globule). 核磁共振波谱结果表明: Box-5 蛋白质的热变性的第一个转变过程不仅包括 Box-5 蛋白质中小翼 (minor wing) 部分的解开, 而且还包括螺旋 3 羧基端的部分去螺旋化, 以及 Box-5 蛋白质中大翼 (major wing) 部分的非极性残基的水合过程; Box-5 蛋白质的热变性的第二个转变过程则对应于这一熔球态的去折叠过程. 明显的, 第二个转变过程要比第一个转变过程更加协同. 差示扫描量热 (DSC) 结果表明, 这一 Box-5 蛋白质熔球态 (Molten Globule) 的去折叠过程所需的焓变约占 Box-5 蛋白质去折叠全部焓变的 69.5%. 所以, 相对于去折叠态 (U state), 这一中间态有相对高的热稳定性. 我们还发现, 局域结构的倾向性以及非天然无规卷曲区的相互作用对 Box-5 蛋白质熔球态的稳定起了关键的作用. 此外, 虽然在这一 Box-5 蛋白质热变性中间态中, α 螺旋和无规则末端区域的高频运动 (皮秒到纳秒时间尺度) 存在明显的差异性, 但是其主链结构的低频运动 (纳秒到微秒时间尺度) 更加均一化 (homogeneous).

按照统计物理学的观点, 处于自然状态下的蛋

白质溶液中, 虽然绝大部分蛋白质以自然折叠状态存在, 仍然有很少量的部分折叠状态和完全非折叠状态的蛋白质存在. 寻常实验方法难以检测到这些部分折叠状态和完全非折叠状态的蛋白质. 在远低于变性转换区的温度或变性剂浓度下进行的 H/D 交换实验被称为自然态 H/D 交换实验 (NHX). 实验中以蛋白结构中氨基质子和溶剂质子的交换为探针, 通过改变溶液的环境, 如加入少量的变性剂, 改变温度等, 非自然态折叠构象的组成比例可以改变. 单个残基的稳定性可以通过各自氨基质子交换的速率以及质子交换对变性剂的敏感程度检测出来. 通过氢键打开方式以及交换所伴随的能量变化的不同也可以得到蛋白质二级结构稳定性高低的信息^[30]. 我们用这种方法系统研究了 hUBF HMG Box5 的去折叠过程. 我们进行了变性剂浓度在 0~1.6 M GdmCl 范围内的自然态下氘交换实验, 并与核磁共振波谱结合. 实验结果表明, 处于蛋白三段螺旋间的疏水核心也正是氘交换慢核心的所在, 而三段螺旋对蛋白自然态结构稳定性的贡献看似并不相同, 螺旋 3 似乎贡献更少一些. 另外自然态下氘交换实验还成功检测到了蛋白折叠过程中部分折叠形式的存在, 并给出了残基不同稳定性的丰富信息^[31]. 类似方法还用于 MICAL-1 CH 结构域的去折叠过程研究^[32].

4 基于蛋白质三维结构的计算机虚拟筛选与核磁实验结合, 发现作为潜在先导化合物的小分子配基, 研究其功能意义

VHR 是从人成纤维细胞的 cDNA 克隆表达产物中发现的具有 185 个残基的蛋白质, 是一种双特异性的磷酸酶. 研究发现 ERK1 和 ERK2 是 VHR 的天然底物, VHR 通过催化磷酸化 JNK 脱磷酸, 起到负调控的作用. ZAP-70 酪氨酸激酶能有效地使 VHR 中的 Tyr138 残基磷酸化, 从而促进 VHR 对 MAPK 途径的抑制. 缺失 VHR 的细胞会在 G1-S 和 G2-M 期静息, 并且有衰老迹象. 筛选出特异性抑制 VHR 活性的分子不仅对于药物开发有帮助, 对于更深一步了解 VHR 的生理作用及具体机制也有很重要的意义. 我们首先运用分子对接程序 DOCK4.0 对含有 80 000 个小分子的库进行虚拟筛选, 选出 25 个小分子. 继而通过扩散编辑 NMR 方法评估所选小分子的结合能力, 发现了 4 个能够结

合 VHR 的配体. 并以对硝基苯酚磷酸盐为底物测定计算了它们抑制 VHR 去磷酸的能力, 发现其中一个小分子 GATPT 具有新颖的结构, 强烈并且竞争性抑制 VHR 的酶活 ($IC_{50} = 2.92 \mu M$). 这是 VHR 的一种新类型的抑制剂. AutoDock 计算显示的复合物结构模型表明这个小分子和蛋白质活性口袋紧密结合. 一系列实验结果显示抑制剂 GATPT 能够影响细胞内 VHR 的天然底物 phospho-ERK 和 phospho-JNK 的表达量, 而且它对细胞周期的影响和 RNA 干扰细胞内 VHR 表达所导致的细胞周期变化一致. VHR 的小分子抑制剂 GATPT 的发现为研究蛋白质酪氨酸磷酸酶 VHR 的生理功能提供了一种新的途径^[33]. 文章被作为封面文章发表于 Chembiochem.

5 结语

综上所述, 核磁共振波谱技术在生物学中, 特别是在结构生物学, 蛋白质科学研究中有广泛应用. 在某些研究领域, 它可以起到其他技术无法替代的作用. 生物核磁共振波谱技术是交叉学科, 涉及磁共振波谱学、分子生物学、生物化学和计算生物学等学科. 我们重点研究的是人的蛋白质的三维结构与功能. 人的蛋白质的结构解析难度很大, 成功率为 3%~5% (原核生物为 10%~15%). 近年来我们实验室解析了一批人的蛋白质或蛋白质结构域的三维结构, 已有 23 个人的蛋白质结构存入 PDB 数据库.

“十五”期间我们作为中国科学院结构基因组研究的牵头单位, 两次应邀在国际结构基因组大会作大会报告, 并被应邀撰写综述文章^[34]. 我们还参与国际综述文章的撰写^[35].

核磁共振波谱技术作为一种重要的结构生物学研究方法, 今后在蛋白质结构、动力学和功能研究方面, 必将与蛋白质晶体学, 冷冻电镜, 计算生物学一起发挥作用. 特别是在研究蛋白质的柔性, 运动性与功能关系, 研究内源无序的蛋白质方面发挥重要作用.

参考文献 (References)

- [1] Kay L E. NMR studies of protein structure and dynamics[J]. *J Magn Reson*, 2005, 173: 193-207.
- [2] Takeuchi K, Wagner G. NMR studies of protein interactions[J]. *Curr Opin Struct Biol*, 2006, 16: 109-117.
- [3] Vaynberg J, Qin J. Weak protein-protein interactions as probed by NMR spectroscopy [J]. *Trends Biotechnol*, 2006, 24: 22-27.
- [4] Shen Weiqun, Xu Chao, Huang Wei, et al. Solution Structure of Human Brg1 Bromodomain and Its Specific Binding to Acetylated Histone Tails[J]. *Biochemistry*, 2007, 46(8): 2100-2110.
- [5] Mujtaba S, Zeng L, Zhou M M. Structure and acetyllysine recognition of the bromodomain[J]. *Oncogene*, 2007, 26: 5521-5527.
- [6] Huang Hongda, Zhang Jiahai, Shen Weiqun, et al. Solution structure of the second bromodomain of Brd2 and its specific interaction with acetylated histone tails [J]. *BMC Struct Biol*, 2007, 7(1): 57.
- [7] Liu Ying, Wang Xiqing, Zhang Jiahai, et al. Structural basis and binding properties of the second bromodomain of Brd4 with acetylated histone tails[J]. *Biochemistry*, 2008, 47(24): 6403-6417.
- [8] Sun H B, Sun Hongbin, Liu J X, et al. Solution structure of BRD7 bromodomain and its interaction with acetylated peptides from histone H3 and H4[J]. *Biochem Biophys Res Commun*, 2007, 358(2): 435-441.
- [9] Xu Chao, Wu Jihui, Xu Yingqi, et al. Backbone and side chain assignments of human Peptidylprolyl isomerase Like 1 (hPPI1) [J]. *Journal of Biomolecular NMR*, 2005, 31(2): 179-180.
- [10] Xu Chao, Zhang Jiahai, Huang Xiaojuan, et al. Solution structure of human peptidyl prolyl isomerase like protein 1 and insights into its interaction with SKIP[J]. *Journal of Biological Chemistry*, 2006, 281(23): 15900-15908.
- [11] Ding H, Xu Y, Chen Q, et al. Solution Structure of Human SUMO-3 C47S and Its Binding Surface for Ubc9[J]. *Biochemistry*, 2005, 44(8): 2790-2799.
- [12] Ding Husheng, Yang Yuedong, Zhang Jiahai, et al. Structural basis for SUMO-E2 interaction revealed by a complex model using docking approach in combination with NMR data[J]. *Proteins: Structure, Function, and Bioinformatics*, 2005, 61(4): 1050-1058.
- [13] Xu Junjie, Zhang Jiahai, Wang Li, et al. Solution structure of Urm1 and its implications for the origin of protein modifiers[J]. *Proc Natl Acad Sci USA*, 2006, 103: 11625-11630.
- [14] Sun Jianping, Zhang Jiahai, Wu Fangming, et al. Solution structure of Ktillp from *Saccharomyces cerevisiae* reveals a novel zinc binding module[J]. *Biochemistry*, 2005, 44: 8801-8809.
- [15] Wu Fangming, Zhang Jiahai, Sun Jianping, et al. Solution structure of human DESR1, a CSL zinc-binding protein[J]. *Proteins*, 2008, 71(1): 514-518.
- [16] Zhou Heyue, Xu Yingqi, Yang Yuedong, et al.

- Solution structure of AF-6 PDZ domain and its interaction with the C-terminal peptides from neuexin and Bcr[J]. *J Biol Chem*, 2005,280:13 841-13 847.
- [17] Chen Quan, Niu Xiaogang, Xu Yingqi, et al. Solution structure and backbone dynamics of the AF-6 PDZ domain/Bcr peptide complex [J]. *Protein Science*, 2007,16:1 053-1 062.
- [18] Wu Jiawen, Yang Yinshan, Zhang Jiahai, et al. Domain-swapped dimerization of the second PDZ domain of ZO2 may provide a structural basis for the polymerization of claudins[J]. *J Biol Chem*, 2007,282 (49):35 988-35 999.
- [19] Li Xiang, Zhang Jiahai, Cao Zanzia, et al. Solution structure of GOPC PDZ domain and its interaction with the C-terminal motif of Neuroligin[J]. *Protein Science*, 2006,15(9):2 149-2 158.
- [20] Mittermaier A, Kay L E. New tools provide new insights in NMR studies of protein dynamics [J]. *Science*, 2006,312:224-228.
- [21] Palmer A G 3rd. NMR characterization of the dynamics of biomacromolecules[J]. *Chem Rev*, 2004, 104:3 623-3 640.
- [22] Palmer A G, 3rd, Massi F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy [J]. *Chem Rev*, 2006,106:1 700-1 719.
- [23] Niu Xiaogang, Chen Quan, Zhang Jiahai, et al. Interesting structural and dynamical behaviors exhibited by the AF-6 PDZ domain upon Bcr peptide binding [J]. *Biochemistry*, 2007, 46 (51): 15 042-15 053.
- [24] Liu Jiangxin, Zhang Jiahai, Yang Yinshan, et al. Conformational change upon ligand binding and dynamics of the PDZ domain from leukemia-associated Rho guanine nucleotide exchange factor[J]. *Protein Sci*, 2008,17(6):1 003-1 014.
- [25] Mittag T, Forman-Kay J D. Atomic-level characterization of disordered protein ensembles [J]. *Curr Opin Struct Biol*, 2007,17: 3-14.
- [26] Xu Yingqi, Wu Jihui, Pei Jiming, et al. Solution structure of BmP02, a new potassium channel blocker form the venom of the Chinese Scorpion *Buthus martensi* Karsch [J]. *Biochemistry*, 2000, 39: 1 369-1 375.
- [27] Yang Wulin, Xu Yinqi, Wu Jihui, et al. Solution structure and DNA-binding property of the fifth HMG box domain in comparison with the first HMG box domain in human upstream binding factor [J]. *Biochemistry*, 2003,42:1 930-1 938.
- [28] Zhang Xuecheng, Zhang Jiahai, Li Xuan, et al. Compact molten globule-like state of hUBF HMG Box1 at extremely low pH[J]. *Biochim Biophys Acta*, 2005, 1 748(1):66-73.
- [29] Zhang Xuecheng, Xu Yingqi, Zhang Jiahai, et al. Structural and dynamic properties of the acid-unfolded state of hUBF HMG Box 1 provide clues for the early events in protein folding[J]. *Biochemistry*, 2005, 44: 8 117-8 125.
- [30] Englander S W. Protein folding intermediates and pathways studied by hydrogen exchange[J]. *Annu Rev Biophys Biomol Struct*, 2000,29: 213-238.
- [31] Wang Dandan, Zhang Jiahai, Jin Xianju, et al. Investigation of the Structural Stability of hUBF HMG Box 5 by native-state hydrogen exchange [J]. *Biochemistry*, 2007,46(5):1 293-1 302.
- [32] Jin Xianju, Zhang Jiahai, Dai Haiming, et al. Investigation of the four cooperative unfolding units existing in the MICAL-1 CH domain[J]. *Biophysical Chemistry*, 2007,129:269-278.
- [33] Shi Zhe, Sartaj Tabassum, Jiang Wei, et al. Identification of a potent inhibitor of human dual-specific phosphatase, VHR, from computer-aided and NMR-based screening to cellular effects [J]. *Chembiochem*, 2007,8(17):2 092-2 099.
- [34] Shi Yunyu, Wu Jihui. Structural basis of protein-protein interaction studied by NMR [J]. *Journal of Structural and Functional Genomics*, 2007,8:67-72.
- [35] Structural Genomics Consortium, Architecture et Fonction des Macromolécules Biologiques, Berkeley Structural Genomics, China structural Genomics Consortium, et al. Protein Production and purification [J]. *Nature Methods*, 2008,5(2):135-146.