

- invasion. *Acta Universitatis Ouluensis Medica*, 1996, D382 (4): 17~ 36
- 3 Reponen P. The Primary structure and developmental expression of mouse type IV collagenases. *Acta Universitatis Ouluensis, A Scintiae Rerum Naturalium*, 1994, A262 (1): 1~ 33
 - 4 Hulboy D L, Rudolph L A, Matrisian L M. Matrix metalloproteinases as mediators of reproductive function. *Molecular Human Reproduction*, 1997, 3 (1): 27~ 45
 - 5 Giannelli G, Falk-Marzillier J, Schiraldi O, *et al.* Induction of cell migration by matrix metalloproteinase-2 cleavage of laminin-5. *Science*, 1997, 277 (11): 225~ 228
 - 6 Woessner J F. Matrix metalloproteinases and their inhibitors in connective tissue remodeling. *FASEB J*, 1991, 5 (8): 2145~ 2154
 - 7 Murphy G, Docherty J P. The matrix metalloproteinases and their inhibitors. *Am J Resp Cell Mol Biol*, 1992, 7 (2): 120~ 125
 - 8 Baragi V M, Fliszar C J, Conroy M C, *et al.* Contribution of the C-terminal domain of metalloproteinases to binding by tissue inhibitor of metalloproteinases. C-terminal truncated stromelysin and matrilysin exhibit equally compromised binding affinities as compared to full-length stromelysin. *J Biol Chem*, 1994, 269 (17): 12692~ 12697
 - 9 Puente X S, Pendas A M, Liano E, *et al.* Molecular cloning of a novel membrane-type matrix metalloproteinase from a human breast carcinoma. *Cancer Res*, 1996, 56 (5): 944~ 949
 - 10 Zhu C, Woessner J F JR. A tissue inhibitor of metalloproteinases and α -macroglobulins in the ovulating rat ovary: Possible regulators of collagen matrix breakdown. *Biology of Reprod*, 1991, 45 (2): 334~ 342
 - 11 Bjorn S F, Hastrup N, Lund L R, *et al.* Coordinated expression of MMP-2 and its putative activator, MT1-MMP, in human placenta. *Mol Hum Reprod*, 1997, 3 (8): 713~ 723

Extra Cellular Matrix and Matrix Metalloproteinases. ZHAO Yun-Ge, Orbit¹⁾ ANWAR, ZHU Cheng (*The State Key Laboratory of Reproductive*

Biology, Institute of Zoology, The Chinese Academy of Sciences, Beijing 100080, China; ¹⁾ *Department of Biology, Xinjiang Normal University, Ürümqi 830053, China).*

Abstract Extra cellular matrix (ECM) is a dynamic reticular structure, which distributes among cell spaces. The structural proteins of ECM are varied and include collagens, proteoglycans, multidomain glycoproteins and other biomacromolecules. These proteins can combine with the specific receptors which exist on the surface of cells. This will result in different gene expression by connecting with intracellular skeleton frame directly or initiating signal transduction cascades, and lead to cells proliferation and differentiation. Matrix Metalloproteinases (MMPs) and their tissue inhibitors (TIMPs) play an important role in the process of ECM destruction and remodeling. MMPs is also a Zn^{2+} dependant proteolytic enzyme. MMPs expression is regulated strictly by growth factors, cytokines, and hormones at transcription level, by their natural activators and inhibitors at protein level. MMP participate various cells physiological and pathological process by degrading different ECM components.

Key words extra cellular matrix, base membrane, matrix metalloproteinases, tissue inhibitors of matrix metalloproteinases

蛋白质溶液 NMR 结构测定的一些新进展

高广华 王大成

(中国科学院生物物理研究所, 北京 100101)

摘要 新的标记技术的进展和采用稀释的液晶作为溶剂以提供额外的结构信息, 提高了核磁共振技术测定蛋白质溶液三维结构的精度, 扩大了分子质量测定范围. 目前已经利用多维¹⁵N, ¹³C, ²H 标记 NMR 测定了许多分子质量为 30 ku 左右的蛋白质溶液结构, 这一上限可能还会被进一步提高.

关键词 核磁共振, 蛋白质三维结构, 液晶, 偶极耦合, 同位素标记

学科分类号 Q51, Q633

多维核磁共振作为测定生物大分子, 特别是蛋白质溶液三维结构最重要的手段, 近年来取得了飞速的发展. 1997 年 10 月 Nature 结构生物学分册的增刊全面综述了 NMR 用于生物大分子研究的最新进展^[1]. 在这一领域一直存在着两个备受关注的

问题, 即 NMR 测定蛋白质溶液结构的分子质量有无上限^[2]和蛋白质溶液 NMR 结构能否达到与 X 射线晶体结构相比拟的分辨率水平^[3]. 一般认为 NMR 最适于中等分子质量的蛋白质的溶液结构测

定. 表 1 是对截至 1995 年 7 月 PDB 库中蛋白质及

表 1 1995 年 7 月 PDB 库中蛋白质及多肽 NMR 结构统计

残基数目	NMR 结构	所占百分比
小于 40	83	20.2%
40~ 80	186	45.4%
81~ 120	81	19.8%
121~ 160	47	11.5%
大于 160	13	3.2%
总 计	410	100%

多肽 NMR 结构的统计. 由表 1 可见, 氨基酸残基数目为 40 至 120 的蛋白质结构占很大的比例, 而 160 残基以上的蛋白质结构较少. 当时分子质量最大的粒细胞集落刺激因子有 178 个残基, 分子质量为 20 ku 左右.

目前这一上限已被大大提高, 完整测定 NMR 结构的蛋白质 (单体) 分子质量已经突破 35 ku. 表 2 是已测定的高分子质量蛋白质的情况. 本文将就最近在提高 NMR 结构分析的分辨率和增大被解析结构的分子质量方面的主要进展作一简要介绍.

表 2 已测定的高分子质量蛋白质 NMR 结构

	蛋白质名称	PDB 文件	分子质量/u (残基数)
1	色氨酸阻遏物-操纵子 DNA 复合物		37000
2	大肠杆菌磷酸烯醇丙酮酸: 糖磷酸转移酶系统中酶 I 的 N 端结构域 (EIN)	1EZA 1EZB 1EYC	30000 (259aa)
3	rRNA 转甲基酶 (ErmAm)		30000
4	人碳酸酐酶 (HCA II)		29000
5	大肠杆菌色氨酸全、去阻遏物		25000 (216aa)
6	Bcl _{xL} -Bak 复合物	1BXL	196aa+ 16aa
7	Bcl _{xL} 蛋白质	1LXL	196aa
8	She 磷酸化酪氨酸结合结构域		190aa+ 12aa
9	粒细胞集落刺激因子	1GNC	178aa

1 液晶与高分辨率溶液 NMR 结构

核欧佛豪斯效应 (nuclear Overhauser enhancement, NOE) 是核间距的定性指标, 从 NOE 得到的距离约束的定性本质限制了所得到的生物大分子时间平均构象的精度. 原子核间的偶极偶合作用包含许多结构信息, 然而在各向同性的溶液中, 由于蛋白质分子在各个方向上的翻滚, 这种作用平均为零而无法观测. 如果能够从偶极偶合作用中直接测定成键原子对相对于分子其他部分的取向, 就将极大地提高结构测定的精度. 当射频脉冲作用于一个磁性核时, 附近另一个磁性核的存在将会由于偶极偶合作用而在两个共振频率之间产生裂分信号, 信号的强度对该化学键相对于外磁场的取向十分敏感. 人们曾试图用各种方法使蛋白质在溶液中部分取向以观察偶极偶合作用. 当蛋白质分子置于强磁场时会因非零的磁感应各向异性而产生微小角度的部分取向. 这时可以观测相隔一个化学键的¹⁵N-¹H 和¹³C-¹H 的偶极偶合. 由于这些偶极偶合的核间距 (即键长) 是固定的, 因此这些数据就提供了相

应的化学键相对于蛋白质磁感应张量的取向的直接信息. 然而对于单个的蛋白质分子来说这种磁相互作用的能量非常微弱而难以准确测量.

液晶可以使溶剂分子产生定向, 但是这种作用对蛋白质分子又太强, 所有的分子都被牢牢束缚并完全定向. 这时偶极偶合作用就不仅仅出现在相邻原子间而是在整个分子内部, 所以将产生大量信号. 即使对很小的分子, 这些数据也根本无法解析. 1997 年, Tjandra 和 Bax^[3]用简单通用的方法巧妙地解决了这个问题. 他们使用的是一种二己酰卵磷脂 (DHPC) 和二豆蔻酰卵磷脂 (DMPC) 混合物 (1: 2. 9), 室温或略高于室温时这种混合物在水溶液中可以在较低的重量体积比 (3%) 下从凝胶转化为液晶, 这些磷脂分子形成直径为数十纳米厚约 4 nm 的盘状微粒, 而且由于这些微粒具有抗磁性, 它们沿磁场方向取向. 它们之间的空间比通常的蛋白质分子的尺寸大得多, 所以蛋白质分子可以在溶液中自由地扩散, 只是偶尔与这些微粒碰撞. 由于蛋白质分子通常呈椭球状, 这种持续的碰撞会使蛋白质分子在液晶排列的方向上发生部分的

取向。蛋白质取向的程度与蛋白质分子的形状有关, 如图 1 所示。

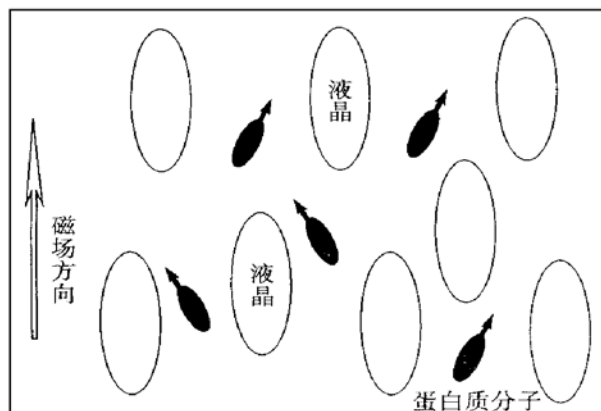


图 1 蛋白质分子在液晶中部分取向示意图

通过改变液晶的浓度可以方便地调节蛋白质分子取向的程度, 这种作用比现在最强的磁场所产生的取向作用强两个数量级。液晶分子与蛋白质之间没有明显的相互作用, 不需要重新再作谱峰的识别。测定这种偶极偶合作用所需的同位素标记的蛋白质样品比测定弱 NOE 峰要少, 一般只需 0.1~0.5 mmol/L。选择适当的液晶浓度可以只使邻近的核给出可观测的偶极偶合作用。这时蛋白质分子所产生的取向程度已经足以精确地测定多种化学键 (^{13}C — ^{13}C , ^{13}C — ^1H , ^{15}N — ^1H , ^{13}C — ^{15}N) 连接的两个核之间的偶极相互作用, 同时又不会使谱图变得极为复杂, 这些准确的结构信息可以大大提高 NMR 测定的蛋白质结构的质量。Tjandra 和 Bax^[3] 对人泛激酶 (76 aa, 8600u) 进行的测定结果与 0.16 nm 高分辨率 X 射线晶体结构非常吻合。此外这种方法还有可能提高 NMR 测定蛋白质结构的分子质量上限, 用高度氘代蛋白质在稀释的液晶溶液中可以快速直接地获得偶极偶合信息, 作为与长程 NOE 相互补充的结构参数, 从而使高分子质量蛋白质溶液结构达到较好的分辨率。虽然目前这种方法还没有用于测定完整的蛋白质结构, 但是在原理上已无太大的问题。一旦用这种方法测定的蛋白质 NMR 结构的精度能一般地达到高分辨率晶体结构的水平, NMR 技术将进入一个新的阶段。

2 氘代与高分子质量蛋白质主链识别

随着蛋白质分子质量增大, 分子中氢原子的数目急剧增加, 谱峰的线宽变宽, 导致谱峰严重重叠, 谱图的解析几乎不可能进行。如果将蛋白质分

子中的 ^1H 用 ^2H 取代将大大简化谱图。目前, 三共振 (triple resonance) 实验结合氘代技术是解决溶液 NMR 分子质量限制的最好办法。1993 年氘代技术用于三共振实验, 并出现了将序列相邻酰胺基团相关的 4D HN (COCA) NH 实验。Yamazaki 等^[4] 设计了一系列用于测定主链 NH, ^{15}N , ^{13}Ca 和侧链 ^{13}CB 化学位移的实验。在高度氘代的蛋白质中, 得到 ^{13}Ca 和 ^{13}CB 的化学位移有助于识别二级结构, 因为这时 NOE 峰的缺失要求尽可能多地测定其他的结构信息。在此之后大量用于主链识别的实验相继出现。最近 Yamazaki 等^[5] 报道了一种获得仅在 ^{13}Ca 位置质子化而其他部分高度氘代的蛋白质的方法。这种标记方法特别有助于因酰胺质子交换速率过快而无法进行的, 基于 NH 的主链识别方法。Yamazaki 等^[4] 还设计了一套用于主链识别的实验来利用通过上述方法制备的蛋白质样品。Farmer 和 Venters^[6] 设计了一系列用于高度氘代的蛋白质侧链识别的实验。在质子化的 ^{15}N , ^{13}C 标记蛋白质中, 侧链化学位移的获得通常是将产生于烷基质子的信号传递到相邻的碳原子然后通过接力传递到主链的 ^{15}N 和 NH 自旋。虽然这种方法易于分析, 但这类实验的灵敏度在蛋白质的分子质量超过 15 ku 时会大大降低。对这类实验的许多改进已经应用于高度氘代的人碳酸酐酶 (HCA II, 259 个残基, 29 ku) 的主链识别, 如使磁化强度从烷基碳开始传递等。与全质子化样品中磁化强度产生于质子的实验相比, 经过改进的实验的灵敏度有明显提高。Farmer 和 Venters^[6] 还设计了许多实验来进行高度氘代样品中侧链 ^{15}N /NH 共振峰的识别。在高度氘代样品中可用于结构测定的 NOE 数目减少, 所以分子中保留下来的所有质子都必须尽可能的识别出来以增加结构信息, 这些质子包括 Glu 和 Asp 的酰胺质子和 Arg 的 ϵ 和 η (如果能够得到) 质子。在对 HCA II 的研究中, 仅对高度氘代样品才能进行主链的全识别, 而质子化的 ^{15}N , ^{13}C 标记样品仅能达到 85% 识别。与此类似, Shan 等^[7] 进行了分子质量达 64 ku 的色氨酸阻遏物, 色氨酸和有 22 个碱基对的 DNA 片段的三聚体的 NH, ^{15}N , ^{13}Ca , 和 ^{13}CB 化学位移的接近完全的识别, 其中蛋白质部分高度氘代 (> 90%), 而氘代率为 70% 时所进行的识别未能成功。Constantine 等^[8] 用高度氘代的样品进行了无底物结合和 NADP⁺ 结合的 UDP-N-乙炔醇丙酮酸葡萄糖胺还原酶 (MurB) 的接近完全的主链识别 (包括 ^{13}CB)。

Garrett 等^[9]利用全氘代¹⁵N, ¹³C 标记的大肠杆菌磷酸烯醇丙酮酸: 糖磷酸转移酶系统中酶 I 的 N 端结构域 (EIN), 长度为 259 个残基的片段的蛋白质样品来识别在全质子化样品中无法识别的 8 个残基. 这些结果表明对于分子质量达 30 ku 或更高的蛋白质进行完全或接近完全的识别必须使氘代达到很高的程度.

目前许多三共振实验用于烷基位置高度氘代的蛋白质样品时灵敏度都相当好, 但是存在的问题是氘代使能够用于结构计算的 NOE 数目减少. 因此 Laue^[10]建议使用约 50% 氘代的样品作为折衷方案, 一方面通过与主链¹⁵N 和 NH 相关进行的侧链¹H 和¹³C 共振峰识别的实验可以获得合适的灵敏度, 另一方面又可以得到较高灵敏度的 NOE 谱.

3 反标记与蛋白质的溶液三维结构

NMR 进行结构测定主要依赖于建立起质子间的距离约束, 这些约束将空间上靠近 (约 0.5 nm 范围之内) 却通常在一级序列上远离的质子联系起来. 由于高度氘代使这些质子的数目大大减少, 从而使高分子质量的蛋白质的结构测定面临极大的困难. 虽然位于可交换位置的质子 (如主链和侧链的 NH) 能够在样品制备之后被重新引入, 在高度氘代样品中仍然可以建立起 NH-NH 间的距离约束, 且仅仅依靠 NH-NH 间的距离约束也可以得到一定分辨率的蛋白质结构, 但是实验表明, 测定精确的蛋白质结构还需要额外的距离信息作为 NH-NH 间距离约束的补充. 为此目的, 许多研究小组致力于研究某些关键基团质子化而其他位置均高度氘代的¹⁵N, ¹³C 标记的样品, 这种方法现在被称为“反标记 (reverse labeling procedures)”. 很明显最合适的选择是 Val, Leu 和 Ile 的甲基. 因为一方面它们是蛋白质疏水核心中最常见的残基, 另一方面它们的堆积是稳定整个分子折叠的重要因素. 此外即使在高分子质量的蛋白质分子中围绕甲基三重轴的旋转也会使¹³C 和¹H 的线宽变窄, 有利于谱图解析. 最后, 相对于谱图的其他部分, ¹³C-¹H 相关谱中的甲基区域较易分析.

现在已有多种不同的方法来制备甲基质子化的高度氘代的蛋白质样品. Metzler 等^[11]和 Laue 等^[12]的方法是将¹⁵N, ¹³C, ¹H-Ile, Leu, Val 引入待研究的蛋白质中以增加可以获得结构信息的位置的数目. 培养基中含有氘代的葡萄糖, 95% D₂O, ¹⁵N, ¹³C, ¹H-Ile, Leu, Val 及一种 ¹⁵N 源如

¹⁵NH₄Cl. 他们都报道在生物合成引入 Leu, Val, Ile 的过程中 C α 位置可以获得高度的氘代, 使用于主链和侧链识别的许多实验的灵敏度大大提高. Rosen 等^[13]报道了另外一种得到甲基质子化高度氘代蛋白质样品的方法. 培养基含有¹³C, ¹H-丙酮酸酯 (CH₃COCOO⁻) 作为碳源, 95% D₂O 以及 NH₄Cl. 质谱和 NMR 的结果表明这个方法得到的蛋白质中大多数侧链位置是氘代的, 而 Ala, Val, Leu, 和 Ile (仅 γ 2 位置) 的甲基是质子化的, 它们都来自丙酮酸酯的质子化的甲基. 甲基化的程度从 Leu 的约 80%, Ile (仅 γ 2 位置) 和 Val 的 60% 到 Ala 的 40%. 另外一些位置, 包括 Asx 的 C β 和 Glx, Arg, Pro 的 C γ 也有一定程度的质子化. 但重要的是位于 Ala, Val, Leu 和 Ile 的质子化的甲基和主链酰胺基团间所有的碳原子均是高度氘代的, 这将有利于用磁化强度沿着侧链碳原子传递到主链酰胺的脉冲序列进行侧链的识别. Kay 等^[14]建立了制备 Val, Leu, Ile (仅 δ 1 位置) 甲基质子化的高度氘代的蛋白质样品的方法. 蛋白质由大肠杆菌高效表达, ¹⁵N, ¹³C, ²H 培养基中含有氘代葡萄糖作为碳源并辅以 [2, 3-²H₂] -¹⁵N, ¹³C-Val 和 [3, 3-²H₂] -¹³C-2-氧代丁酸酯 (CH₃CH₂COCOO⁻).

现在已经建立了综合利用 CH₃-CH₃, CH₃-NH 和 NH-NH 间 NOE 信息来测定蛋白质结构的方法, 与只利用 NH-NH 间 NOE 的方法相比, 结构的正确性和精度都明显提高. PLC SH2 结构域^[15], Grb2 NSH3/SOSE 蛋白质-多肽复合物^[11]和 ras p21 蛋白质 GDP^[12]的结构测定证实了其正确性. 如果要求更高的分辨率, 就要在蛋白质中引入更多的质子. Farmer 等^[16]和 Laue 等^[12]制备了芳香残基质子化的高度氘代的 MurB-NADP⁺ 复合物样品. 此外, 还可以利用二面角的构象数据库或¹³Ca 和¹³C β 的化学位移来精化由有限的 NOE 约束获得的 NMR 结构.

尽管在高度氘代的样品中选择性质子化对于高分子质量蛋白质的结构测定是必需的, 但值得注意的是可以利用多个氘代程度不同的样品进行实验. Fesik 等^[17]采用这种方法测定了分子质量为 29 ku 的 rRNA 转甲基酶的结构, 主链和侧链的识别由部分氘代 (约 75%) 的¹⁵N, ¹³C 标记的样品获得. NH 对间的距离约束来自高度氘代 (> 90%) 的样品, 而另外的 NOE 则来自全质子化的¹⁵N, ¹³C 标记的蛋白质样品. Fesik 领导的研究小组已经利用这种方法测定了 Shc 磷酸化酪氨酸结合结构域

(190 aa)^[18]、Bcl_{x_L} 蛋白质 (209 aa)^[19] 及 Bcl_{x_L} 与 Bak 多肽复合物的结构^[20]。Garrett 等^[9] 采用类似的方法研究了 EIN 的结构。高度氘代的¹⁵N, ¹³C 标记样品用来证实由全质子化样品所建立起来的对 NH, ¹⁵N 和¹³C 的识别, 并得到了在未氘代时不可能进行的识别。全氘代样品则用来以高的灵敏度记录 NH-NH 间的 NOE。对于质子化样品的谱图质量较好的高分子质量蛋白质如 EIN 和 Bcl_{x_L} 来说, 结构测定的最好方法就是首先利用甲基质子化高度氘代的样品获得的 CH₃-CH₃, CH₃-NH, NH-NH 间的 NOE 约束建立初始结构模型, 然后再利用质子化样品获得的大量 NOE 进行结构的精化。

目前 NMR 测定的蛋白质完整溶液三维结构的最大分子质量为 37 ku, 主链完全识别的最大分子质量已达 64 ku, 精度最高可达 0.15 nm 分辨率。与 X 射线衍射方法相比, NMR 还是一种尚未完全成熟的技术, 其各个方面都在经历着不断的变革。从已经显露的技术发展趋势看, 似乎可以明确回答本文开篇所提到的问题, 即 NMR 结构解析的分子质量还有很大提高的余地, 不久将可能扩展到 50 至 60 ku, 结构解析的分辨率可以达到与 X 射线单晶结构分析相当的水平。随着理论知识的日益完善, 实验技术的新的进展以及硬件性能的不断改进, NMR 技术将在结构生物学领域发挥更大的作用。

参 考 文 献

- Wagner G. An account of NMR in structural biology. *Nature Structural Biology*, 1997, **4** (supplement): 841~ 844
- Kay L E, Gardner K H. Solution NMR spectroscopy beyond 25 kDa. *Current opinion in structural biology*, 1997, **7** (5): 722 ~ 731
- Tjandra N, Bax Ad. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science*, 1997, **278** (5340): 1111~ 1114
- Yamazaki T, Lee W, Revington M, *et al.* An HNCA pulse scheme for the backbone assignment of ¹⁵N, ¹³C, ²H labeled proteins: application to a 37 kDa trp repressor-DNA complex. *J Am Chem Soc*, 1994, **116** (14): 6464~ 6465
- Yamazaki T, Tochio H, Furui J, *et al.* Assignment of backbone resonances for larger proteins using the ¹³C-¹H coherence of ¹Ha, ²H, ¹³C and ¹⁵N labeled sample. *J Am Chem Soc*, 1997, **119** (5): 872~ 880
- Farmer B T, Venters R A. Assignment of aliphatic side-chain ¹H/¹⁵N resonances in perdeuterated proteins. *J Biomol NMR*, 1996, **7** (1): 59~ 71
- Shan X, Gardner K H, Muhandiram D R, *et al.* Assignment of ¹⁵N, ¹³Ca, ¹³Cβ and HN resonances in an ¹⁵N, ¹³C, ²H labeled 64 kDa trp repressor-operator complex using triple resonance NMR spectroscopy and ²H-decoupling. *J Am Chem Soc*, 1996, **118** (28): 6570~ 6579
- Constantine K L, Mueller L, Goldfarb V, *et al.* Characterization of NADP+ binding to perdeuterated MurB: backbond atom NMR assignments and chemical-shift changes. *J Mol Biol*, 1997, **267** (5): 1223~ 1246
- Garrett D S, Seok Y J, Liao D I, *et al.* Solution structure of 30 kDa N-terminal domain of enzyme I of the Escherichia coli phosphoenolpyruvate: sugar phosphotransferase system by multidimensional NMR. *Biochemistry*, 1997, **36** (9): 2517 ~ 2530
- Nietlispach D, Clowes R T, Broadhurst R W, *et al.* An approach to the structure determination of larger proteins using triple resonance NMR experiments in conjunction with random fractional deuteration. *J Am Chem Soc*, 1996, **118** (2): 407~ 415
- Metzler W J, Wittekind M, Goldfarb V, *et al.* Incorporation of ¹H/¹³C/¹⁵N- (Ile, Leu, Val) into a perdeuterated, ¹⁵N labeled protein: potential in structure determination of large proteins by NMR. *J Am Chem Soc*, 1996, **118** (28): 6800~ 6801
- Smith B O, Ito Y, Raine A, *et al.* An approach for global fold determination using limited NMR data from larger proteins selectively protonated at specific residues. *J Biomol NMR*, 1996, **8** (3): 360~ 368
- Rosen M K, Gardner K H, Willis R C, *et al.* Selective methyl group protonation of perdeuterated proteins. *J Mol Biol*, 1996, **263** (5): 627~ 636
- Gardner K H, Kay L E. Production and incorporation of ¹⁵N, ¹³C, ²H (¹H-δ1 methyl) isoleucine into proteins for multidimensional NMR studies. *J Am Chem Soc*, 1997, **119** (32): 7599~ 7560
- Gardner K H, Rosen M K, Kay L E. Global folds of highly deuterated, methyl protonated proteins by multidimensional NMR. *Biochemistry*, 1997, **36** (6): 1389~ 1401
- Farmer B T, Constantine K L, Goldfarb V, *et al.* Localizing the NADP+ binding site on the MurB enzyme by NMR. *Nat Struct Biol*, 1996, **3** (12): 995~ 997
- Yu L, Petros A M, Schnuchel A, *et al.* Solution structure of an rRNA methyltransferase (ErmAm) that confers MLS antibiotic resistance. *Nat Struct Biol*, 1997, **4** (6): 483~ 489
- Zhou M M, Ravichandran K S, Olejniczak E F, *et al.* Structure and ligand recognition of the phosphotyrosine binding domain of Shc. *Nature*, 1995, **378** (6577): 584~ 592
- Muchmore S W, Sattler M, Liang H, *et al.* X-ray and NMR structure of human Bcl_{x_L}, an inhibitor of programmed cell death. *Nature*, 1996, **381** (6580): 335~ 341
- Sattler M, Liang H, Nettlesheim D, *et al.* Structure of Bcl_{x_L}-Bak peptide complex: recognition between regulators of apoptosis. *Science*, 1997, **275** (5302): 983~ 986

Progress in the Determination of Protein Solution Structure by NMR. GAO Guang-Hua, WANG Da-Cheng (*Institute of Biophysics, The Chinese Academy of Sciences, Beijing 100101, China*).

Abstract Advances in new label strategies and application of liquid crystalline medium for giving additional structural information have improved the

accuracy of solution protein structures determined by NMR, and extended the molecular weight limit. Several structures of proteins of 30 ku have been solved using ^{15}N , ^{13}C , ^2H labeling multidimensional NMR techniques, and this limit will probably be

surpassed in the near future.

Key words nuclear magnetic resonance, protein three-dimensional structure, liquid crystal, dipolar coupling, isotope labeling

纤维素酶分子结构和功能研究进展*

阎伯旭 齐 飞 张颖舒 高培基¹⁾

(山东大学微生物技术国家重点实验室, 济南 250100)

摘要 概述了近 10 年来利用结构生物学和蛋白质工程技术在纤维素酶分子结构和功能方面研究的进展, 包括: 酶分子结构域的拆分、催化域和纤维素结合结构域的结构和功能的研究, 纤维素酶的分子折叠, 并展望了该领域的研究前景.

关键词 纤维素酶, 催化结构域, 纤维素结合域结构域, 分子折叠, 活性位点, 双置换机制

学科分类号 Q936

纤维素这一巨大可再生资源的有效利用对解决环境污染、食品短缺、能源危机具有重大现实意义, 利用微生物产生的纤维素酶来分解和转化纤维素则是纤维素利用的有效途径, 故对纤维素酶的研究越来越引起人们的重视. 纤维素酶的研究, 自从 1904 年在蜗牛消化液中被发现, 至今已经历三个发展阶段. 第一阶段是 80 年代以前, 主要工作是利用生物化学的方法对纤维素酶进行分离纯化. 但由于纤维素酶来源广泛、组分复杂、纯化甚为困难, 故进展缓慢. 第二阶段是 1980 年至 1988 年, 主要工作是利用基因工程的方法对纤维素酶的基因进行克隆和一级结构的测定. *Trichoderma reesei* 的内切酶 (EG I、EG III) 和外切酶 (CBH I、CBH II)、*Cellulomonas fimi* 的内切酶 (CenA、CenB、CenC、CenD) 和外切酶 (CbhA、CbhB、Cex)、*Clostridium thermocellum* 的内切酶 (CelA、CelB、CelC、CelD) 的基因已被克隆和测序, 并在大肠杆菌、酵母菌等中得到表达. 第三阶段是 1988 年至今, 主要工作是利用结构生物学及蛋白质工程的方法对纤维素酶分子的结构和功能进行研究, 包括纤维素酶结构域的拆分、解析、功能性氨基酸的确定、水解的双置换机制的确立, 分子折叠和催化机制关系的探讨.

1 纤维素酶分子结构域的拆分

1986 年 Tilbeurgh^[1] 用木瓜蛋白酶有限酶切

Trichoderma reesei 的 CBH I 分子得到具有独立活性的两个结构域: 一个是具有催化功能的催化域 (catalytic domain, CD), 另一个是具有结合纤维素功能的纤维素结合 (吸附) 区 (cellulose binding domain, CBD), 之后用类似的方法在多种细菌和真菌的纤维素酶中发现类似的结构. CBD 在纤维素酶中位于氨基端或羧基端, 它通过一段高度糖基化的 linker 与催化区相连.

细菌纤维素酶的 linker 富含 Pho、Thr, 而真菌纤维素酶的 linker 富含 Gly、Ser、Thr; 细菌纤维素酶的 CBD 与 CD 夹角为 135° , 而真菌为 180° ; 细菌纤维素酶有两个酶切位点可将 CBD 与 linker 分别切去, 而真菌纤维素酶一般只有一个酶切位点可将 CBD 与 linker 一同切去.

由于纤维素全酶分子呈蝌蚪状、其 linker 高度糖基化且具有较强柔韧性, 故很难得到结晶. 而结构域的拆分使呈球形催化域的结晶成为可能, 为纤维素酶的结构和功能研究开辟了道路.

2 催化域结构和功能的研究

采用 X 光衍射的方法 1990 年 Rouvinen^[2] 对 *T. reesei* 的 CBH II 的催化域、1992 年 Juy^[3] 对 *C. thermocellum* 的 CelD 的催化域、1993 年 Spezio^[4]

* 国家教委博士点基金资助项目.

¹⁾ 通讯联系人.

收稿日期: 1998-02-19, 修回日期: 1998-05-29