

综述与专论

蛋白质卷曲研究进展* (上)

石 颖 许根俊 鲁子贤

(中国科学院上海生物化学研究所, 上海 200031)

提 要

对于简单的球状蛋白来说, 它形成天然态有活力的空间结构的信息都包含在氨基酸序列之中。理论上讲, 从一级结构预测空间结构进而推测出蛋白质生物功能是可行的。但到目前为止, 这条路尚未走通。原因之一是用一维信息编码三维结构的过程十分复杂。这个过程就称为蛋白质卷曲。

文章介绍了蛋白质的体外和体内卷曲以及卷曲的起始等方面近年的研究状况。

关键词 蛋白质卷曲, 卷曲密码, 卷曲起始点, 体外卷曲与体内卷曲。

三十多年前, 两项实验结果导致了蛋白质卷曲问题的提出。一是 Perutz 与 Kendrew 等人向人们展示了(对比 DNA 分子来说)血红蛋白具有高度复杂性的空间结构。二是 Anfinsen 通过对核糖核酸酶的研究向人们提出了蛋白质的一级结构决定空间结构的假说。从那时起, 致力于攻克一级结构如何决定空间结构这一难题(有人称为破译生命的另一半密码)的工作一直没有停止过。

近几年来, 蛋白质卷曲研究获得长足进展。除了它的纯学术价值仍具有强大吸引力而外, 新的驱动力还来自以下几方面: 一是蛋白质分子设计与蛋白质工程的需求。分子生物学技术已能方便地制造一个全新的蛋白质的多肽链, 但还需要使它能进行正确的卷曲以生成有生物活性的空间结构。二是越来越多的又各具特点的基因工程产物的复性、复活需要蛋白质卷曲的理论及技术的指导。三是蛋白质卷曲研究方法的进步。例如二维核磁共振技术及高分辨率核磁共振仪的完善, 使解析蛋白质溶液构象的能力大大增强。阐明卷曲过程和捕捉卷曲中间态成为可能; 重组 DNA 技术的应用为蛋白质

卷曲的研究提供了丰富的研究对象, 这些变异蛋白质可以按人们的设计去获得一级结构和高级结构的大量信息; 此外, 随着计算机的更新换代, 实验数据的增加, 各种类型的卷曲理论研究(例如能量优化、分子动力学等)正取得较大进展。

本文将围绕卷曲密码, 体外卷曲与体内卷曲、卷曲起始点等问题的实验研究进展进行综述。

1 蛋白质卷曲的密码

蛋白质功能是由其空间结构决定的, 而蛋白质生物合成却只能给出氨基酸顺序。为了使蛋白质分子具有生物活性, 一级结构与空间结构之间必定存在某种关系。有人称其为蛋白质卷曲的密码(code of protein folding), 立体化学码(stereochemical code)^[1,2], 用以强调空间结构的信息蕴藏在一级结构之中。

* 谨以此文祝贺邹承鲁教授七十寿辰。

本文由国家自然科学基金重大项目资助, 编号 9388006。

收稿日期: 1992-05-04 修回日期: 1992-10-07。

在核糖核酸酶(RNase)可逆变性研究的基础上, Anfinsen 提出了“一级结构决定高级结构”的观点. 越来越多的体外蛋白质卷曲的实验结果(特别是单肽链的小蛋白的卷曲)都证明了 Anfinsen 命题的正确性.

另一方面, 如果一级结构被破坏了, 卷曲就可能出现异常. 人们在实践中发现绝大多数小蛋白都能进行伸展与再卷曲的可逆过程. 但是枯草杆菌蛋白酶在体外完全伸展后, 却不能正确卷曲回到天然状态. 进一步研究发现该酶在酶原阶段可以正确卷曲. 而酶原 N 端的自我剪切发生在卷曲之后, 这样, 成熟的酶由于丢失了一段序列, 就不能正确卷曲了.

Karplus 等人曾指出: 对一个长度为 100 个残基的多肽链来说, 若从可能的构象中随机筛选出一个特定的构象, 需要时间为 10^{50} 年. 事实上, 即便在体外实践中, 完全松散的肽链恢复其天然态结构的时间一般在秒和分钟范围内, 至多以小时计, 卷曲即可完成. 因此蛋白质卷曲不可能是一个完全随机的过程.

对于有 6 个半胱氨酸的牛胰蛋白酶抑制剂(BPTI)来说, 它可能的二硫键配对数为 15, 但从还原产物卷曲到天然产物的过程中, 由捕捉的中间产物分析, 有些二硫键配对显然呈绝对优势, 而另外一些可能从未出现过. 这可能意味着, 一级结构不仅为空间结构编码, 而且以某种方式规定指导着达到这种空间结构所经历的途径或过程.

尽管空间结构信息包含在一级结构中的观点已为人们接受, 但关于立体结构密码的形式却一无所知. 已知蛋白质空间结构相比于一级结构具有较强的保守性. 也就是说, 在一级结构的许多位置上进行残基置换, 对空间结构的影响并不大, 如不同种属的细胞色素 c (cyt. c). 对于 KVLDA 这样一个五肽来说, 在前白蛋白中呈 β 折叠结构, 而在碳酸酐酶中, 却呈 α 螺旋结构. 这种一级结构相同, 二级结构不同的肽段广泛存在于蛋白质中, 被称为 α/β 构象两可肽. 一方面, 序列相近的蛋白质取相同的空间结构, 另一方面, 序列相同的肽段可以取

不同的二级结构. 这反映了卷曲密码的复杂性, 它不可能象 mRNA 中三联码那样简单明了.

2 蛋白质的体外卷曲与体内卷曲

目前尚不能准确指出体外卷曲能在多大程度上代表了体内卷曲.

2.1 一种观点认为体外卷曲可以完全模拟体内卷曲, 他们的依据是: 有些小蛋白, 例如核糖核酸酶和葡萄球菌核酸酶(SNase), 在切除 C 端一段或 C 端本身时, 伸展后即不能回到天然态的结构. 在用合成的不同长度的 C 端片段与其互补时, 其构象和活力均得以恢复. 因此, 卷曲应发生在肽链合成完成后, 才能得到有活力的正确构象. 多结构域的章鱼碱脱氢酶要求 C 端完整的情况下, 才能恢复天然态结构.

另外一些工作表明, 对于无二硫键的小蛋白来说, 体外卷曲速度非常快: 葡萄球菌核酸酶和醛缩酶都在 1s 内完成, 这样的速度比肽链生物合成速度至少快 10 倍以上. 因此, 合成后卷曲是完全可能的. 蛋白质合成效率不会因为合成后卷曲而降低, 也就是说, 合成后卷曲不会成为产生活性蛋白质的限速步骤.

最近, 用脉冲标记二维核磁共振技术研究细胞色素 c 的再卷曲时, 发现位于 N 端和 C 端的两个 α 螺旋是最早形成的结构, 并且通过相互作用而稳定. 这可能暗示着再卷曲一开始就要求远程作用, 并且, 由于 C 端螺旋参与早期构象形成, 卷曲应在肽链全部合成后才启动.

2.2 另一种观点主张体外卷曲不能完全模拟体内卷曲, 其依据为: 肽链合成尚未结束的 β -半乳糖苷酶(未脱离核糖体)已具有酶的活性, 并且可与抗 β -半乳糖苷酶的抗体起反应. 这表明 C 端不完整的 β -半乳糖苷酶具有天然态的空间结构. 还有, 在体外合成 IgG 分子时, 至少在(四条链中的)一条链合成尚未结束时, 抗体分子已具有了活性. 这些都说明卷曲随肽链合成同步进行着. 随着肽链延长, 分子具有越来越多的天然态结构.

另外, 缺少 C 端二肽的小蛋白卵清溶菌酶

的卷曲及活力均不受影响,说明 C 端并不是必不可少的。近年来,独立的片段可以卷曲成其在天然分子中具有的例子越来越多。似乎表明卷曲并不依赖一级结构的完整性。

关于边合成边卷曲的一个较有说服力的例子是 IgG 的重链在小鼠骨髓瘤细胞中合成。作者能清楚展示重链内二硫键的生成、甚至重链与轻链间二硫键的形成都发生在肽链合成完成之前。

2.3 邹承鲁提出了一个新生肽链卷曲的模型^[3]。他认为新生肽链的卷曲很可能在转译过程中即已开始,随着肽链延伸,其空间构象不断调整,这一调整过程在合成终结之后才完成并达到天然态构象。这是一个既包括 co-translational,也包括 post-translational 的过程。据此,他还提出用不同长度的含 N 端的肽段作为研究对象可能更接近于新生肽链的卷曲。

2.4 以往认为,蛋白质卷曲不需要任何辅助因子,这一观念已不能适用于体内卷曲。近来发现了一些与体内卷曲有关的蛋白质,被称为助卷曲蛋白(folding helper)。第一类是能加速卷曲过程中的慢反应步骤的蛋白质,例如蛋白质二硫键异构酶(PDI)和肽酰脯氨酰顺反异构酶(PPIase)。它们可以帮助蛋白质克服卷曲限速步骤,尽快完成卷曲,大大减少新生肽链水解以及卷曲中间态发生凝聚的机会。

第二类是参与体内伸展与再卷曲,保护卷曲中间态以利于新生肽链跨膜运送的一类蛋白质。如分子伴侣,总的来说,它们保护了卷曲中间态,也稳定了中间态,这在体外卷曲中是很难实现的。

3 蛋白质卷曲起始点研究进展

蛋白质卷曲是如何发动的?这方面的研究主要集中在对以下几种假说的检验上。

3.1 卷曲由疏水塌缩(hydrophobic collapse)开始,然后进入一个卷曲中间态——熔球态(molten globule state)。这个熔球态的性质决定了以后的卷曲步骤,有可能是通过限定

肽链进一步卷曲的构象数目来影响以后的卷曲步骤的。

疏水塌缩一词是在比较了蛋白质的结构与肥皂微泡的结构之后推测得到的。一个球状蛋白与一个肥皂微泡的结构很相似:极性基团朝外,疏水部分朝内。二者差别只是在球蛋白中疏水部分紧密堆积在一起形成一个疏水核,而肥皂微泡的疏水部分相比之下较为松散。因此推测蛋白质在卷曲开始时可能经历过一个非特异的疏水塌缩过程。

这方面的证据来自葡萄球菌核酸酶的点突变对其卷曲行为的影响。该突变不影响天然态酶的稳定性,但却大大降低了伸展态的稳定性。由于突变引入的是疏水残基,因此推论它加强疏水塌缩的程度。

近来,小肽中的疏水作用研究也有进展^[4]。例如,在 VQWL 这样的一个四肽中,Val 和 Trp 的疏水作用使这个肽表现出结构有序性。还有个例子是,来自菜豆铜蓝蛋白的一个片段(26—33)KIVFKNNA 中的 Ile, Val 和 Phe 在水溶液中也形成了疏水簇(hydrophobic clusters),从而影响了肽链摆动的自由度。

所以,疏水作用可以在近距离发生,在小肽中发生,在没有形成二级结构之前发生,这些结果提示在分析卷曲起始点时不能忽略疏水作用的贡献。

3.2 二级结构生成作为卷曲起始点。按照框架模型的假设,二级结构在卷曲中起决定性作用。蛋白质卷曲开始于独立存在的,忽隐忽现的二级结构。这些二级结构接触靠近时互起稳定作用,形成稳定结构单元。三级结构的形成是将这些二级结构单元组装起来,并将水从疏水核中排挤出去。

近来关于小肽在水溶液中可以形成稳定的二级结构的报导引起人们极大兴趣。以往人们一直认为少于 20 个残基的肽段在水溶液中不能形成 α 螺旋。这样的观念现在已被更新。一个突出的例子是 RNase A 的 C 肽(N 端 13 肽)在水溶液中仍能形成 α 螺旋结构。进一步研究发现, Glu2 和 Arg10 之间形成的盐键对螺旋起了

稳定作用,这一盐键在完整的 RNase A 分子中也存在.另外 Phe8 与 His12 的侧链也有相互作用.这说明,小肽可以通过一些特殊作用形成稳定的二级结构.

利用核磁共振技术, Dyson 等还在短肽中找到一个不稳定的新生螺旋结构 (nascent helix). 这是在具有免疫原性的蚯蚓肌红蛋白的 C 螺旋片段 (19 肽) 观察到的. 在水中, 该片段的 C 末端有一个转角样的结构与伸展态快速互变. 在水与三氟乙醇的混合溶剂中, 这结构被稳定成为 α 螺旋.

除了 α 螺旋而外, Dyson 发现 YPXDV (X 代表除 Trp 和 Pro 之外的 L 型氨基酸) 能够形成 β 转角结构, 特别是 YPGDV, 在 5°C 时, 50% 的反式异构体以 β 转角结构存在. 用其它氨基酸取代第 3 位的 Gly 时导致这个结构的稳定性下降. 看来, 第 3 位的氨基酸的性质是影响稳定性的重要因素. 一个有意义的发现是用不同的残基置换 YPXDV 中的 X, 得到的 β 转角的含量与用 Chou 和 Fasman 的方法得到的转角形成几率存在着相关性. 这可能暗示着天然态蛋白质中的 β 转角只是由几个残基参与的近程作用决定的. 形成转角的肽段极有可能作为蛋白质卷曲的起始点^[4], 它在卷曲早期非常快地限定了肽链可能采取的构象.

二级结构作为卷曲起始点的证据不仅来自小肽在水溶液中能形成稳定的二级结构的实验, 而且还表现在卷曲动力学实验中, 蛋白质的二级结构先于三级结构形成这一事实. Kuwajima 在用 stopped flow-CD 研究 cyt c 和 β -乳球蛋白 (β -LG) 的再卷曲时发现, 这两个蛋白质的二级结构形成时间 (由远紫外 CD 监测) 少于 18ms, 而三级结构恢复时间 (由近紫外 CD 监测) 为 100—500s. 所以, 二级结构生成明显快于三级结构. 用类似的实验方法对 RNase S, α -乳清蛋白, 溶菌酶, 碳酸酐酶等进行了研究, 结果基本相同. 因此, 这是小蛋白卷曲中的普遍规律.

最近, Udgaonkar 等人与 Roder 等人分别建立了类似的动态研究卷曲途径、捕捉中间态

结构的新方法, 称作脉冲标记二维核磁共振法 (pulse-labeling-2D-NMR). 用上述方法研究 cyt c 的再卷曲, 发现它的 N 端螺旋和 C 端螺旋是最早形成的结构, 并且两个螺旋通过相互作用而稳定. 在研究 RNase A 时发现, 卷曲起始, β 折叠迅速形成并具有很好的协同性. 这是一个 β 折叠在卷曲的早期就已经形成的例证. 对去辅基肌红蛋白 (apo-Mb) 的研究发现, 它分子中的 A, G, H 螺旋是最先生成的结构, 此时 B, C, D, E 肽段尚处在松散态^[5].

3.3 特殊作用启动蛋白质卷曲. 如果一个天然蛋白质的空间结构能帮助特异的化学键生成, 那么这个化学键将会对稳定蛋白质空间结构做出贡献. 因此, 特异的成键作用能够发动蛋白质卷曲.

对于 BPTI 的卷曲, 30-51 位的二硫键起到了稳定在卷曲过程中很重要的一个中间态构象的作用. 通常认为, 二硫键通过降低构象熵 (conformational entropy) 而减小伸展态肽链的稳定性, 这个构象熵之差值可用来评估二硫键对导致它形成的那个结构的稳定性的贡献. 得到的结果是单 S-S 键中间态 (30-51) 比不含 S-S 键那个结构在稳定性上增加了至少 100 倍. 已知这个中间态对 BPTI 的进一步卷曲是非常重要的, 它的 (30-51) S-S 键一直保留到终态.

Oas 和 Kim^[6]进一步用实验来证明二硫键在卷曲中的作用. 他们分别合成了 BPTI 的 C 端 16 肽 ($P\alpha$, 43—58 肽段, 天然态分子中为 α 螺旋) 和中间 14 肽 ($P\beta$, 20—33 肽段, 天然分子中为反平行 β 折叠). 还原态的 $P\alpha+P\beta$ 的 CD 信号表明两个肽都是伸展的. 当两个肽由 (30-51) S-S 键连接起来时, 各自表现出了在天然态分子中的构象.

3.4 伸展态中的有序结构启动卷曲. 随着检测蛋白质分子构象的仪器的灵敏度提高, 小肽在水溶液中有序结构的发现, 人们对所谓的完全伸展态的认识不断加深^[7]. 现在看来, 用各种手段变性的蛋白质的肽链并不象以前想的那样松散, 而是不同程度存在一些有序结构^[8].

这样的一些有序结构是否扮演蛋白质卷曲

起始点的角色这个话题又重新提起。从前面介绍的实验结果看,小肽能形成 α 螺旋, β 折叠, β 转角,也能形成疏水簇。近程、中程作用力足以维持它们稳定存在,二硫键还原及许多变性条件都不能破坏它们的有序结构。所以,Creighton认为,尽管现在没有证据表明伸展态中的有序结构对卷曲有动力学上的重要性(如RNase A中N端稳定螺旋仍未显示动力学上的重要性),它们的作用是不能忽视的。就象BPTI中能偏向30-51二硫键生成的那个有序结构一样,一旦这些有序结构与其它的和卷曲起始相关的构象或特殊作用偶联起来,它们对于卷曲起始来说就会显得非常重要。

早些时候,Ptitsyn经过热力学及动力学的计算指出, α 螺旋可以预先存在于松散的肽链中作为结构胚胎启动蛋白质卷曲。但他认为 β 折叠与 β 转角由于没有预先存在于伸展态结构中而不参与启动卷曲。现在看来,这后一观念也被实验结果更新^[4]。

参 考 文 献

- 1 Jaenicke R. Folding and association of proteins. *Prog Biophys Mol Biol*, 1987; **49**: 117
- 2 Kim P S, Baldwin R L. Intermediate in the folding reactions of small proteins. *Annual Rev Biochem*, 1990; **59**: 631
- 3 Tsou C L. Folding of the nascent peptide chain into a biologically active protein. *Biochemistry*, 1988; **27**: 1809
- 4 Wright P E, Dyson H J, Lerner R A. Conformation of peptide fragments of proteins in aqueous solution; Implications for initiation of protein folding. *Biochemistry*, 1988; **27**: 7167
- 5 Hughson F M, Wright P E, Baldwin R L. Structural characterization of a partly folded apomyoglobin intermediate. *Science*, 1990; **249**: 1544
- 6 Oas T G, Kim P S. A peptide model of a protein folding. *Nature*, 1988; **336**: 42
- 7 Dill K A, Shortle D. Denatured states of proteins. *Annu Rev Biochem*, 1991; **60**: 795
- 8 Richards F M. The protein folding problem. *Scientific American*, 1991; **264** (1): 34

真核核基因 RNA 剪接研究的新进展

徐 军 陈润生

(中国科学院生物物理研究所, 北京 100101)

提 要

简述了真核生物核基因 RNA 剪接研究的一些新进展, 主要包括剪接体的结构组成, 剪接体的逐步组装, 剪接反应的影响因素等方面。

关键词 剪接体, snRNA, PRP, 预组装

现已发现的绝大多数真核蛋白质基因是不连续的, 编码蛋白的序列被不编码的序列所间隔。前者称作外显子(exon), 后者称作内含子(intron), 转录时先生成一条包括有内含子的前 mRNA(pre-mRNA), 然后经过剪接得到成熟的 mRNA, 可以作为模板翻译蛋白质。

剪接由剪接体完成。剪接体由特定的分子 RNA 和蛋白质构成。

剪接是基因表达过程中一个非常重要的环节, 对于它的逐步了解大大丰富了我们在分子这一层次上对生命活动的认识, 同时内含子的起源和存在的生物学意义是一个极其诱人的课题, 对它的研究将推进我们对于整个基因组结构规律的认识。