

锚蛋白重复序列介导的蛋白质-蛋白质相互作用*

杜海宁 胡红雨**

(中国科学院上海生命科学研究院, 生物化学与细胞生物学研究所, 上海 200031)

摘要 锚蛋白重复序列 (ANK) 是生物体中广泛利用的一种序列模体. ANK 模体在 ANK 结构域中折叠成 $\beta_2\alpha_2$ 结构, 在空间上则形成 L 型结构. 数目不等的 ANK 串联起来, 依靠氢键和疏水相互作用, 组成紧密、稳定的结构域, 并且形成了种类众多但功能各异的 ANK 蛋白质分子. ANK 结构域介导蛋白质与蛋白质的相互作用, 它能够和多种配体结合, 实现纷繁复杂的生物功能. 着重介绍几类结构已知的 ANK 家族蛋白质分子及复合物的结构特征、生理功能及与疾病的关系.

关键词 锚蛋白重复序列, ANK 结构域, 蛋白质相互作用, 疾病

学科分类号 Q71

锚蛋白重复序列 (ankyrin repeat, ANK) 是普遍存在于生物体中的一种蛋白质序列模体 (motif). 在已被确定的 400 多种含有该模体的蛋白质中, 就发现了约 3 000 个 ANK 重复序列 (参阅 <http://smart.embl-heidelberg.de>). 由此可见, 研究 ANK 模体对于认识这类蛋白质家族成员的结构和功能具有十分重要的作用.

1987 年, Breeden 和 Nasmyth 在两个酵母细胞周期调控蛋白 Swi6 和 Cdc10 中首次发现了 ANK 模体. 随后, 在细胞骨架锚蛋白 (ankyrin) 中发现了含有 24 个此重复序列模体的结构, 因此人们将其命名为锚蛋白重复序列模体 (ANK repeat motif)^[1].

含有 ANK 的蛋白质分布十分广泛, 从病毒、原核生物、真核生物到人体; 从细胞核、细胞质到细胞外介质都能找到. 这类蛋白质数量众多, 形成了一个大的家族, 其中主要包括 CDK (cyclin-dependent kinase) 抑制剂, 转录调控因子, 细胞骨架形成体 (cytoskeletal organizers), 发育调控子, 膜蛋白和毒素等. 在这些蛋白质分子中, 2 到 25 个数目不等的 ANK 单元串联起来, 形成许多大的 ANK 结构域 (ANK domain), 行使生物功能. 已报道的最小的 ANK 蛋白是果蝇中的钚蛋白 (plutonium protein), 只有两个相邻的 ANK 重复序列, 在细胞发育中可抑制 DNA 的早期复制, 但它不具有独立完整的空间结构. 由 p16^{INK4c} 蛋白 C 端两个 ANK 组成的 p16c 是目前发现的能够独立折叠成高级结构的最小单位.

ANK 结构域的主要生物功能是介导蛋白质与

蛋白质之间的相互作用. 由于该蛋白质家族成员中 ANK 在数目、一级序列以及空间结构上都存在差异, 使得 ANK 模体能够与多种配体结合, 实现纷繁复杂的生物功能.

1 ANK 的结构特征

典型的 ANK 重复序列一般含有 33 个残基, 保守残基的一级序列和二级结构可由图 1 表示.

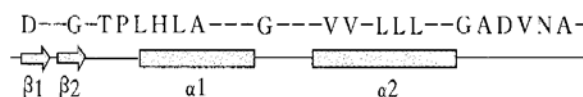


Fig. 1 The ANK repeat consensus

图 1 保守的 ANK 重复序列

同源性比较显示, 该序列中的每个位置都可以被特定类型的氨基酸残基所代替, 并且蛋白质的功能越类似, 其保守性也越强. 从 ANK 家族蛋白质分子整体上看, 具有保守性的氨基酸残基在结构上的作用大致可分为两类: 一类参与每个独立重复单位的形成和构建; 另一类形成各重复单位之间相互作用的表面残基, 对整个 ANK 结构域起到稳定的作用^[2-4].

晶体结构显示, ANK 模体在 ANK 结构域中折叠成 $\beta_2\alpha_2$ 结构 (图 2a), 在空间上则形成 L 型结构. β 发夹 (字母 “L” 的短臂) 延伸在反平行的

* 国家自然科学基金资助项目 (NSFC30070165).

** 通讯联系人.

Tel: 021-64374430-5121, E-mail: hyhu@sunm.shenc.ac.cn

收稿日期: 2001-06-01, 接受日期: 2001-08-25

两个 α 螺旋(字母“L”的长臂)之外,与 α 螺旋构成的平面几乎垂直.彼此相邻的ANK模体之间通过 β 发夹形成一个连续的反平行 β 片层,而几对 α 螺旋肩并肩排列,依靠疏水作用相互堆积,形成一个紧密而舒展的螺旋平面,一侧疏水,另一侧亲

水.从空间上看,ANK结构域就象一只弯曲的手(图2b): β 片层形成手指,而螺旋平面形成手掌部分.手指与手掌相互垂直,从而形成手心内部的一个疏水性深沟^[1~4].

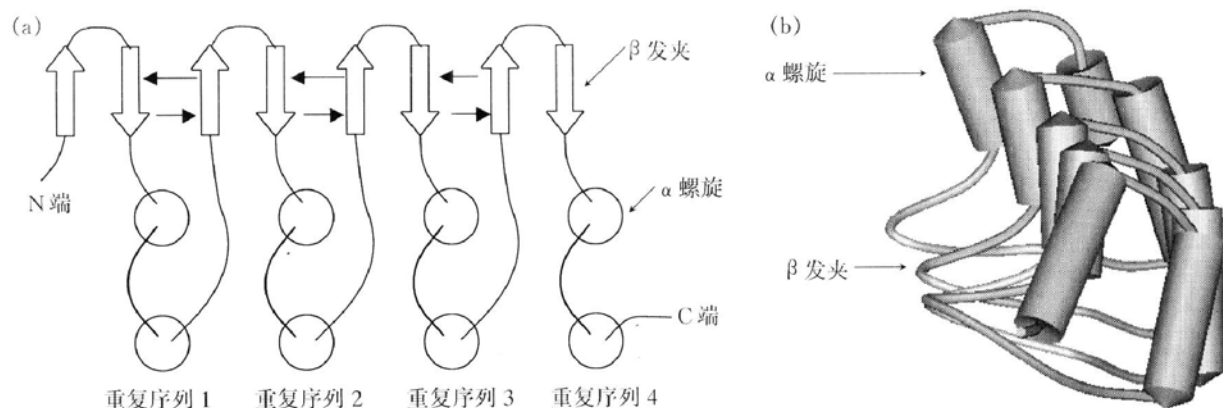


Fig. 2 Typical ANK repeat structure

图2 典型的ANK重复序列结构

(a) ANK重复序列二级结构拓扑示意图.空心圆圈代表 α 螺旋,其螺旋轴轴心与图示平面垂直;空心箭头代表 β 发夹结构;实心箭头代表 β 片层之间的氢键作用.(b) GABPB蛋白(PDB编号1AWC)ANK结构域的X晶体衍射结构图.

β 发夹结构通常由4个亲水性氨基酸组成,它一方面通过形成连续的 β 片层,另一方面通过主链与Asp1(ANK序列上的第一个氨基酸)侧链之间的氢键作用来稳定结构. α 螺旋中的疏水性氨基酸残基侧链形成非极性表面,临近的两个ANK螺旋对则在形状上完全互补,从而结成一个紧密的堆积结构.另外,彼此相互临近的两个ANK模体的主链原子之间及His9、His16和Asn31之间的氢键作用也进一步稳定了堆积结构.Gly4、Gly15和Gly27非常保守,Gly4终止了 β 2,Gly15和Gly27则分别终止了 α 1和 α 2.ANK模体中另一特征是由四肽序列TPLH形成的一个紧密转角(tight turn),它是连接 β 2和 α 1的纽带.Thr6和His9之间的互补氢键保证了转角的刚性结构.

同时需要指出的是,许多ANK家族蛋白质分子的ANK结构域并不完全保持典型的结构特征.例如DNA结合蛋白GABPB(GA-binding protein β)含有4.5个ANK重复序列.而INK4家族成员中的第二个重复单元(repeat2, R2)序列长度相对于其他单元较短.另外,有些ANK结构域不具有连续性,一些氨基酸序列会插入到临近的两个ANK重复单元之间,有一些更是位于ANK模体的内部.比如I κ B蛋白质家族中的I κ B α 和I κ B β ,R1和R2以及R3和R4之间的连接部分都各插有

一个螺旋片段.而在Swi6蛋白中,R2单元内部连接螺旋的环形结构上以及R2和R3之间各含有一个较短的螺旋片段.更特别的是,在R3和R4之间,有一段由40个残基组成的小结构域,包含两个螺旋,一个短的 β 片层和一段无序的环形结构.一方面,这些结构上的“变异”,使得ANK的空间结构多样化,能够与毫无关联的蛋白质配体结合,实现了功能上的多样化;另一方面,由此也必将造成其结构稳定性的下降,这就需要增加重复序列数目,以及形成独特的空间结构来增强其稳定性.因此, α 螺旋束整体趋向于形成左手扭曲,使得整个螺旋平面向ANK深沟弯曲;保守的非极性侧链与内部的螺旋相互作用,使得ANK这个非球形结构域形成更小的结构,为增强其稳定性提供了保证.

2 ANK结构域的生物功能

由于ANK结构域作用广泛,从发现伊始就引起了广泛关注.目前通过X射线和NMR方法已经解析出了几类ANK蛋白质家族及其配体复合物的结构,可以使我们从分子水平上了解它们的生物功能.

53BP2-p53复合物是第一个被解析出的含有ANK模体的蛋白质结构,53BP2上的ANK结构

域与 p53 相互结合, 共同承担肿瘤抑制作用。

GABPB 是另一类含有 ANK 的蛋白质。该结构域可以结合 α 亚基, 形成异源二聚体, 作用于 DNA 的特定碱基上, 从而在发育调控、癌症发生及病毒基因表达方面发挥重要作用。

INK4 家族是 ANK 结构研究得最清楚的一类蛋白质。四个成员 p16^{INK4a}、p15^{INK4b}、p18^{INK4c} 和 p19^{INK4d} 含有 4 或 5 个 ANK 序列。ANK 结构域是重要的功能区, 它们可以特异性地抑制 CDK4 和 CDK6 的激酶活性, 对哺乳动物细胞周期 G1 后期的调控产生影响; 另外与人的癌症也紧密相关。

I κ B 蛋白质家族成员都含有 6 个 ANK 序列, 其主要功能是与 NF κ B 上的 RHR 结构域 (Rel homology domain) 发生作用, 抑制 NF κ B 转入细胞核或者阻止 NF κ B 与 DNA 的结合。从 I κ B α -NF κ B 的晶体结构中看出, 在发生相互作用时, 每个 ANK 单元都参与了结合, ANK 结构域得到了充分利用。

Swi6/Cdc10 家族蛋白质成员 Swi4/6、Mbp1、Res1/2、Cdc10 等都具有由 5 个 ANK 序列组成的核心结构域, 其参与整个细胞周期中的信号传导事件, 抑制 CDK 的活性, 正向或负向地影响转录激活。值得注意的是, Swi6 蛋白是迄今为止发现的唯一 ANK 结构域通过分子内作用发挥功能的例子。多个插入到 ANK 结构域中的结构单元和 ANK 重复序列一起参与了转录激活功能。

ANK 结构域不仅参与和细胞内信号传导、转录调控和发育调控等有关的相互作用, 它还直接或间接地导致人体疾病。Notch 受体家族中的 ANK 结构域能够诱导 T-细胞的白血病。Swi6 蛋白家族同源的脑内蛋白质 V-1, 可以影响儿茶酚胺的生物合成, 与神经性和精神性疾病紧密相关。

肌营养蛋白 (myotrophin) 是含有 2.5 个 ANK 重复序列的蛋白质, 它可以通过蛋白质-蛋白质相互作用, 诱导与心肌肥大有关的特异性基因的表达, 从而导致多种心脏病的发生。从 NMR 结构上看出, ANK 组成了 myotrophin 的主体结构, 其突出的两个“手指”没有形成特征的 β 发夹结构, 相互之间也没有氢键作用, 而处于近乎无序的状态。该无序的“手指”在序列上具有较大的可变性, 它很可能是发生蛋白质-蛋白质相互作用的关键位点, 进一步研究清楚 myotrophin 的分子作用机制对于治疗与心肌肥大有关的心脏病具有重要意义。

肝癌在亚洲和非洲是一种常见的疾病, 最近日本科学家从肝细胞癌中找到一种含有 6 个 ANK 序列的致癌蛋白 gankyrin^[5], 它能够与细胞癌基因 RB1 结合, 增强 RB1 的磷酸化, 加速 RB1 的降解。另外, 通过酵母双杂交技术发现一个新的结合蛋白 synphilin 1, 它由 920 个残基组成, 在其中部含有 4 个 ANK 序列。该蛋白质可以在神经元中与 α -synuclein 特异性结合。 α -synuclein 的积聚被认为是造成帕金森病的主要原因, 而它的积聚与该蛋白质突变、过量表达或与其他蛋白质结合有关。对于 gankyrin 和 synphilin 1, 其 ANK 的空间结构和功能尚不明了, 但已有实验证明它们与蛋白质之间的相互作用有关, 解析其 ANK 结构是解开致病机制的必要手段。

3 结 语

生物体内数以万计的蛋白质的产生往往来源于编码特定序列的外显子的复制和“穿梭”, 这些特定的序列有的能够独立折叠成有功能的结构域 (如 SH2/SH3 结构域, 免疫球蛋白结构域); 而另一些为只含有 20~40 个残基的序列, 单一的模体没有一定的结构和功能, 但是它们是构成蛋白质骨架的基本元素, 几个重复序列串联起来就形成了特定的发挥功能的结构 (如 HEAT 模体, 富亮氨酸模体)。

事实上, 复杂的生物反应大部分正是通过有限的蛋白质结构模体完成的。蛋白质的进化既需要关键氨基酸的保守性来保持结构上的完整性, 又要通过序列上的变化来实现功能上的特异性和生物多样性。正如 ANK 模体, 螺旋堆积内保守的疏水性残基既保证了整个 ANK 结构域的形成和稳定, 同时又作为发生相互作用的结合位点; β 发夹指尖和其他非保守元件的插入为其功能多样化提供了结构基础。进一步研究 ANK 结构域对于找出蛋白质-蛋白质相互作用的规律具有重要意义。

参 考 文 献

- 1 Sedgwick S G, Smerdon S J. The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem Sci*, 1999, **24** (8): 311~316
- 2 Russo A, Tong L, Pavletich N P, *et al.* Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16^{INK4a}. *Nature*, 1998, **395**: 237~243
- 3 Jacobs M D, Harrison S C. Structure of an I κ B α /NF- κ B complex. *Cell*, 1998, **95** (6): 749~758
- 4 Foord R, Taylor I A, Smerdon S J, *et al.* X-ray structural analysis of the yeast cell cycle regulator Swi6 reveals variations of

- the ankyrin fold and has implications for Swi6 function. *Nat Struct Biol*, 1999, **6** (2): 157~ 165
- 5 Higashitsuji H, Itoh K, Fujita J, *et al.* Reduced stability of retinoblastoma protein by gankyrin, an oncogenic ankyrin repeat protein overexpressed in hepatomas. *Nat Med*, 2000, **6** (1): 96 ~ 99

Ankyrin Repeat Mediated Protein-protein Interaction^{*}

DU Hai-Ning, HU Hong-Yu^{**}

(*Institute of Biochemistry and Cell Biology, Shanghai Institute for Biological Sciences,
The Chinese Academy of Sciences, Shanghai 200031, China*)

Abstract Ankyrin repeat (ANK) is a widely used sequence motif in organisms. An individual ankyrin repeat has an L-shaped structure consisting of a β hairpin followed by two α helices, namely $\beta_2\alpha_2$ pattern. These ANK repeats form a compact domain stabilized through hydrogen-bonding interaction and hydrophobic stacking. Massive of structurally similar but functionally diverse proteins containing ANK repeats can be formed by virtue of assembly of the tandem array motifs. The central role of ANK repeat is to mediate protein-protein interaction, so that proteins can make sophisticated functions by interacting with a variety of ligands. This review focuses on structure, function and related diseases of the ANK repeats and their complexes from several members of the ANK family.

Key words ankyrin repeat, ANK domain, protein interaction, disease

^{*} This work was supported by a grant from the National Natural Sciences Foundation of China (NSFC30070165).

^{**} Corresponding author. Tel: 86-21-64374430-5121, E-mail: hyhu@sunm.shnc.ac.cn

Received: June 1, 2001 Accepted: August 25, 2001