

WW 结构域 —— 蛋白质-蛋白质相互作用的一种模式*

李琦 胡红雨**

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

摘要 WW 结构域是由 38~40 个氨基酸残基严密组织形成一个连贯、紧凑的结构域; 它以包含两个色氨酸残基为主要特征, 能专一地与含有 XPPXY 保守序列的蛋白质相互作用。这种相互作用涉及许多细胞内事件, 如非受体信号传导、转录调节、蛋白质降解等等, 并且这种相互作用的变化会直接或间接影响到人体的正常生理代谢功能而引起疾病。

关键词 WW 结构域, 蛋白质-蛋白质相互作用, 疾病

学科分类号 Q51

Tom Blundell 定义蛋白质结构域为: a. 它是一个连贯的三维结构; b. 可互换并且半独立的功能单位; c. 在高等真核细胞中, 对应于一个穿梭外显子; d. 它由至少 35~40 个残基构成, 最小、最紧密、也最稳定; e. 作为一个结构和功能单位, 在同一个蛋白质中有时会重复出现^[1]。

WW 结构域由 38~40 个半保守的氨基酸残基组成, 介导蛋白质-蛋白质的相互作用, 在功能上它与 SH3 结构域很相似。它具有以下两个突出特点: a. 含有两个色氨酸, 彼此间隔 20~22 个残基, 也因此被叫做 WW 结构域; b. 作为细胞内结构单元, 它的分布十分广泛, 存在于许多种蛋白质中 (可参阅 <http://www.embl-heidelberg.de/bork/www1.html>)。这些具有 WW 结构域的蛋白质功能多样, 包括结构功能、调节功能、信号传导功能等。本文拟重点介绍 WW 结构域的结构、功能研究的主要成果和最近进展。

1 WW 结构域的结构特征

引出 WW 结构域源于对 SH2 结构研究的深入。80 年代晚期, 在 Rous 肉瘤病毒的研究中, 发现 SRC 家族激酶的 N 端在介导 SRC、YES 和其他非受体激酶的信号传导中具有重要作用。那时位于 N 端的 SH2 结构域首先被发现, 并作为介导激酶和其他蛋白质相互作用的重要候选结构。很快地, 在 SRC 的 SH2 结构域上游又发现 SH3 结构域。又以 YES 的 SH3 结构域为‘饵’, ‘钓出’了 YAP (YES-associated protein), 该蛋白质含有 PXXP (P 为脯氨酸残基, X 为任意残基) 的保守序列。正是

这一结果, 直接为 WW 结构域的登场奠定了坚实的基础。

一方面, YAP 不仅可通过 PXXP 保守序列与 YES 相互作用, 而且可与其他蛋白质发生相互作用, 可以假设有另外的蛋白质作用位点; 另一方面, 通过酵母、鼠、人、鸡等生物 YAP 的同源比较, 发现有一保守区域, 这就是 WW 结构域^[2,3]。

1.1 WW 结构域的一级结构及其配体

WW 结构域一般含有 38 个氨基酸残基, 在一级序列上以含有两个相隔一定数目氨基酸残基的色氨酸残基为主要特征 (且这两个色氨酸相隔 20~22 个氨基酸残基)。在少数情况下, 下游的色氨酸可被酪氨酸或苯丙氨酸所代替。其他位置的氨基酸残基至少有 5 个位置是保守或半保守的。其一级序列为: W-X (9, 11) - [VFY] - [FYW] - (6, 7) - [GSTNE] - [GSTQCR] - [FYW] - X (2) P

WW 结构域的配体最早证明是脯氨酸丰富区域, 这样就产生一个问题: 已知 SH3 结构域的配体也是脯氨酸丰富区, 那么两者的配体是否相同, 是否会因此产生竞争结合呢? 体外实验显示, 序列 XPPXY (称为 PY 模体; P 为脯氨酸, X 为任意氨基酸) 为配体结合 WW 结构域所必需, 相比之下, SH3 的配体必须有 PXXP 保守序列^[4]。最近的一

* 中国科学院基金资助项目 (STZ98-2-02)。

** 通讯联系人。

Tel: 021-64374430-5245, E-mail: hyhu@sunm.shenc.ac.cn

收稿日期: 2000-06-26, 接受日期: 2000-08-23

些实验结果表明, WW 的配体保守序列还有不同于 XPPXY 的其他多聚脯氨酸形式^[5].

在对含有 WW 结构域的蛋白质作系统分析后, 还得到一些有趣的结果. 例如, WW 结构域邻近区域通常有一个组氨酸或者半胱氨酸丰富区, 这些区域在一些蛋白质中可结合金属离子, 如肌营养不良蛋白^[2]; 又如在 WW 结构域附近作同源分析, 又找到了一个经常与 WW 结构域相伴出现的结构域——FF 结构域^[6].

1.2 WW 结构域的高级结构及与配体的作用

发现 WW 结构域时, 就推测它和 SH3 结构域一样是一个 β 片层结构. Macias 等^[7]用 NMR 首先解出 WW 结构域和它的配体复合物的溶液结构 (图 1), 证实了这个猜想. 实验所用 WW 结构域来自 YAP 蛋白, 配体选用一段含 PPXY 保守序列的 10 肽. 从整体上来看, WW 结构域被认为是已知的、存在于自然界中最紧密的球状结构. 它主要表现为一个扭转并稍弯曲的三个反平行 β 折叠片状结构. 这三个 β 折叠片分别包括 16~22, 26~32 和 35~39 位置的残基, 形成的 β 片层结构十分紧密, 在不需要二硫键和辅因子能单独存在的球状折叠形式中, WW 结构域是最小的之一. 值得注意的是位于片层底部的残基 Tyr28 和 Trp39, 两者的侧链形成了一个平坦的疏水面, 这个疏水面就是与配体相互作用的区域. 边上两个疏水氨基酸残基 Leu30 和 His32 也部分参与了这种相互作用. 随后的 X 光衍射结果^[8]及突变研究证明了 Trp39、Tyr28、Leu30 和 His32 这几个位点确实直接参与了与配体的相互作用.

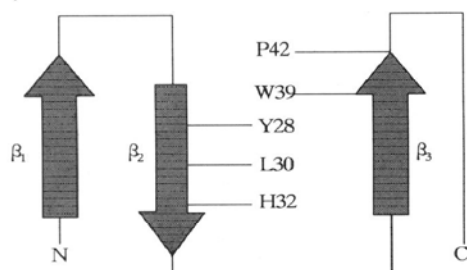


Fig. 1 Schematic representation of the secondary structure of WW domain

图 1 WW 结构域二级结构示意图

那么, 这种相互作用是如何实现的呢? 首先配体的两个脯氨酸和 WW 结构域的第二个色氨酸形成范德华相互作用力; 其次是配体的酪氨酸插入由相对保守的 Leu30 和 His32 形成的疏水口袋中, 与 His32 建立起氢键相互作用 (图 2).

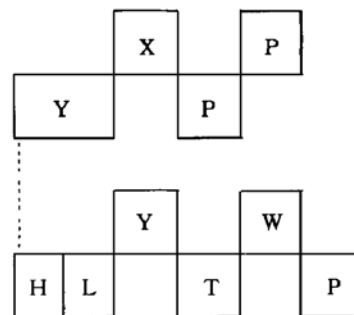


Fig. 2 Schematic view of the interaction of WW domain with a proline-rich ligand

图 2 WW 结构域与配体复合体的相互作用

知道这种相互作用的形式后, 一个重要的问题在于搞清楚这种相互作用是如何得到有效地调控, 是不是最常见的磷酸化调控机制. 目前的实验结果显示, WW-配体复合物的形成受许多因素的影响, 它们的 K_d 值 (解离常数) 高至纳摩尔水平, 低至微摩尔水平. 在不同的结构域-配体对、是否磷酸化、不同的磷酸化位置、不同的缓冲液条件以及分析方法下, 得到的 K_d 值不尽相同^[8~10].

在 Chen 等^[11]的实验中, 他们发现配体保守序列 PPXY 中的酪氨酸一旦被磷酸化, YAP 的 WW 结构域就不再能和它结合. 也就是说, 这种相互作用是受酪氨酸磷酸化的负调控, 只有当配体蛋白去磷酸化时, WW 结构域才可介导供体蛋白与配体蛋白的相互作用. 然而, Pin1 蛋白的 WW 结构域可以和磷酸化的配体蛋白结合, 但是磷酸化的位点是丝氨酸或苏氨酸^[10]. 因此, 可以推测, 在不同的 WW 结构域-多聚脯氨酸蛋白对中, 它们作用的调控方式也是不同的.

2 WW 结构域-配体复合物的生理功能

自 WW 结构域发现伊始, 就引起研究者的广泛关注. 这是因为它和配体复合物直接或间接地参与多种人体疾病的发生与发展 (表 1).

Table 1 The WW ligand complexes and their related physiological functions

表 1 WW 结构域-配体复合体和它参与的生理功能

含 WW 结构域的蛋白质	作用蛋白	作用模式	相关疾病, 相关生理过程	参考文献
NEDD4	ENaC	PPPXY	Liddle's 综合症	[12], [13]
Dystrophin	β -dystroglycan	PPXY	肌营养不良症	[2], [14]
FE65	β APP	PPPPPPL/R	阿尔茨海默症	[15]
HYP A, HYP B, HYP C	Huntingtin	PPXY	亨廷顿氏病	[16]
FBPs	formin	PPPPPPL/R	四肢、肾脏发育	[17]
PQP B-1	Brr-2	未确定	抑制转录过程	[18]

2.1 WW 结构域-配体复合物与 Liddle's 综合症

1963 年, Liddle 与其同事们描述了一种罕见的盐敏感型高血压综合症. 该病患者有严重的高血压和代谢碱中毒, 这两种症状可能都和肾上腺醛固酮的过量产生有关. Liddle 等当时就证明不是由于甾体激素的过量产生而影响 Na^+ 离子通道正常功能, 导致这种罕见的高血压症状^[19]. 于是必须从肾表皮 Na^+ 通道本身来寻找原因.

研究者运用突变的方法深入研究 ENaC (肾表皮钠离子通道, epithelial Na^+ channel) 的结构与功能的关系, 发现在 β 、 γ 亚基 C 端引入突变或提前引入终止密码子会使得远端肾上皮细胞对钠离子的重吸收, 导致 Liddle's 综合症症状. 于是把注意力集中到 ENaC 三个亚基的 C 端, 找到一个在三个亚基中都很保守的脯氨酸丰富区: PPPXY. 实验中只需对其中的脯氨酸突变, 就会造成截去 C 端的同样后果, 出现病变^[12, 13]. 在临床中, 也先后发现了 β 亚基保守序列变成 PPLNY 或 PPPNH 而导致 Liddle's 综合症的病例. 此后又有大量的体内实验证明了 PPPXY 序列中突变与 Liddle's 综合症存在着因果关系.

通过酵母双杂交系统, 研究者找到与 ENaC 脯氨酸丰富区相作用的蛋白质 NEDD4, 它包含三个 WW 结构域和一个泛肽连接酶结构域. 在推测的模型中, 三个 WW 结构域分别通过脯氨酸丰富区结合 ENaC 的三个亚基. 随后, 泛肽连接酶通过修饰离子通道而调节它的降解过程, 当三个作用位点, 有一个出现问题都会妨碍复合物的形成, 影响离子通道的泛肽化. 在其他的模型中, 有研究者认为这种相互作用可以直接调控离子通道的开关状态, 而不仅仅是它的降解. 虽然目前的假设尚不能真正反映实际的过程, 公认的一点是 WW 结构域-PPPXY (ENaC) 的相互作用是与 Liddle's 综合症相联系的一个关键步骤^[12].

2.2 WW 结构域和肌营养不良症

Bork 等在肌营养不良蛋白的 C 端, 第四个脯氨酸丰富区和半胱氨酸丰富区之间, 鉴别出一个 WW 结构域, 由于已经知道了 C 端和其他蛋白质有相互作用, 于是他们又把重点放在和 C 端相互作用的两个膜蛋白上: β -dystroglycan (43 ku) 和 syntrophin (59 ku). 通过一系列实验证实, β -dystroglycan 正是通过细胞质侧的一个脯氨酸丰富区与肌营养不良蛋白相互作用的. 那么是否 β -dystroglycan 受体的脯氨酸丰富区的突变就会导致肌营养不良症, 这需要更多的实验证据的支持^[2, 14].

2.3 WW 结构域和 Alzheimer's 病

Bork 等在 FE65 接头蛋白上发现了 WW 结构域. FE65 接头蛋白在信号传导中起着“二传手”的作用, 一方面它可以通过一个 PTB 结构域和膜上的 β APP (Alzheimer's β -amyloid precursor protein) 发生相互作用; 另一方面, 以 WW 结构域和另一个 PTB 结构域与其他蛋白质相互作用并引起一系列的细胞内反应^[15]. 有关的研究小组已找到富含脯氨酸的蛋白质, 并深入研究该蛋白质与 FE65 的相互作用与 Alzheimer 病^[20]的相互关系.

2.4 WW 结构域与 Huntingtin 病

Peter 等以酵母双杂交系统对人胎脑 cDNA 表达文库进行筛选, 结果用 Huntingtin 的 N 端筛到 7 个新蛋白质 (而用中间及 C 端片段未筛到). 从筛到的蛋白质中鉴定到一个主要的作用家族, 该家族蛋白质都含有 WW 结构域. 这些蛋白质都通过 WW 结构域与 Huntingtin N 端的脯氨酸丰富区相作用. 当同样位于 N 端的多聚谷氨酰胺重复次数增加时, 会大大加强这种相互作用. 研究者推测, WW 结构域-多聚脯氨酸介导的过程, 如非受体的信号传导, 蛋白质降解或者 mRNA 的剪切可能参与了 Huntingtin 病^[21]的病理发生^[16].

2.5 其他可能参与的生理过程

WW 结构域与胚胎时期人体四肢和肾脏的发育有关^[21]。有一类叫做形成素 (formin) 的蛋白质, 它们包含一个脯氨酸丰富区, 该区域内 100 个氨基酸中有 66 个是脯氨酸。它的配体蛋白 FBP (formin binding protein) 含有一个 WW 结构域。奇怪的是, 它们还同时含有一个 SH3 结构域。WW 和 SH3 结构域结合的都是多聚脯氨酸区, 那么体内为什么存在一种蛋白质 (FBP) 可以两个不同的结构域与其他蛋白质 (formin) 相互作用呢? 通过体外实验, 推测可能由 FBP 的 WW 结构域通过竞争 SH3 结构域的结合位点来调节 SH3 结构域与 formin 的相互作用。也可能这种调节还涉及磷酸化过程或者其他结构域的参与如 SH2 等。

WW 结构域还与逆转录病毒的芽生^[22]和致癌性转移^[23]有关。逆转录病毒颗粒组装和出芽过程都依赖一个名为 GAG 的蛋白质。该蛋白质含有一个保守的脯氨酸丰富区。如果突变该脯氨酸区, 就能阻断病毒的出芽。研究者认为 GAG 蛋白正是通过脯氨酸丰富区与宿主细胞中含 WW 结构域的蛋白质相互作用, 以实现出芽过程的。同样人体内编码 GAG 的致癌基因, 如果缺失该部分, 就无法发生致癌性转移, 只是其中的细节还不明了。

近来, 又在 PQBP-1^[18]、C-jun, AP2, NF-E^[9], C/EBP α , PEBP2/CBF 等转录调控因子中发现有 PPXY 序列, 含有 WW 结构域的蛋白质可以通过识别这个序列与相应蛋白质作用。这种作用对它们的转录活性有重要影响, 而这些含有 WW 结构域的蛋白质就充当了转录辅因子的作用。

蛋白质-蛋白质相互作用是许多细胞事件的基础。然而, 细胞内蛋白质的数量众多、种类各异, 它们之间的相互作用是各具特异性还是遵循一定规律, 亦或两者皆有。目前的研究进展越来越支持后一种观点。已经发现的 SH2、SH3、PH 以及 PTH 等结构域广泛参与了蛋白质间的相互作用, 涉及各种细胞内事件。WW 结构域作为一个新发现的成员, 同样具有分布及功能上的广泛性。它既见于单细胞生物体内, 又存在于多细胞生物体中, 它与多聚脯氨酸形成的复合物是蛋白质间相互作用的模式之一。

参 考 文 献

- Sudol M. Structure and function of the WW domain. *Prog Biophys Mol Biol*, 1996, **65** (1~2): 113~132
- Bork P, Sudol M. The WW domain: a signaling site in dystrophin?. *Trends Biochem Sci*, 1994, **19** (12): 531~533
- Sudol M. YAP-65 is a proline rich phosphoprotein that binds to the SH3 domain of the Yes proto-oncogene product. *Oncogene*, 1994, **9** (8): 2145~2152
- Chen H I, Sudol M. The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc Natl Acad Sci USA*, 1995, **92** (17): 7819~7823
- Kay B K, Williamson M P, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *The FASEB Journal*, 2000, **14** (2): 231~241
- Bedford M T, Philip Leder. The FF domain: a novel motif that often accompanies WW domains. *Trends Biochem Sci*, 1999, **24** (6): 261~263
- Macias M J, Hyvonen M, Schultz J, *et al.* The structure of the WW domain in complex with a proline-rich peptide. *Nature*, 1996, **382** (6592): 646~649
- Ranganathan R, Lu K P, Hunter T, *et al.* Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell*, 1997, **89** (6): 875~886
- Mosser E A, Kasanov J D, Forsberg E C, *et al.* Physical and functional interactions between the transactivation domain of the hematopoietic transcription factor NF-E2 and WW domains. *Biochemistry*, 1998, **37** (39): 13686~13695
- Lu P J, Zhou X Z, Shen M, *et al.* Function of WW domains as phosphoserine or phosphothreonine-binding modules. *Science*, 1999, **283** (5406): 1325~1328
- Chen H I, Einbond A, Kwak S J, *et al.* Characterization of the WW domain of human yes-associated protein and its polyproline-containing ligands. *J Biol Chem*, 1997, **272** (27): 17070~17077
- Schild L, Canessa C M, Shimkets R A, *et al.* A mutation in the epithelial sodium channel causing Liddle disease increase channel activity in the *Xenopus laevis* oocyte expression system. *Proc Natl Acad Sci USA*, 1995, **92** (12): 5699~5703
- Snyder P M, Price M P, McDonald F J, *et al.* Mechanism by which Liddle's Syndrome mutations increase activity of a human epithelial Na⁺ channel. *Cell*, 1995, **83** (6): 969~978
- Campbell K P. Three muscular dystrophies: loss of cytoskeleton-extracellular matrix linkage. *Cell*, 1995, **80** (5): 675~679
- Russo T, Faraario R, Minopoli G, *et al.* Fe65 and the protein network centered around the cytosolic domain of the Alzheimer's β -amyloid precursor protein. *FEBS Lett*, 1998, **434** (1~2): 1~7
- Faber P W, Barnes G T, Srinidhe J, *et al.* Huntingtin interacts with a family of WW domain proteins. *Hum Mol Genet*, 1998, **7** (9): 1463~1474
- Sudol M. The WW module competes with the SH3 domain?. *Trends Biochem Sci*, 1996, **21** (2): 161~163
- Waragai M, Lammers C H, Takeuchi S, *et al.* PQBP-1, a novel polyglutamine tract-binding protein, inhibits transcription activation by Brr-2 and affects cell survival. *Hum Mol Genet*, 1999, **8** (6): 977~987
- Lifton R P. Genetic determinants of human hypertension. *Proc Natl Acad Sci USA*, 1995, **92** (19): 8545~8551
- Schellenberg G C. Genetic dissection of Alzheimer disease, a heterogeneous disorder. *Proc Natl Acad Sci USA*, 1995, **92** (19): 8552~8559
- Huntingtin's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*, 1993, **72** (6): 971~983

- 22 Garnier L, Wills J W, Verderame M F, *et al.* WW domain and retrovirus budding. *Nature*, 1996, **381** (6584): 744~ 745
- 23 Mayer B J, Hanafusa H. Mutagenic analysis of the v-crk oncogene: requirement for SH2 and SH3 domains and correlation between increased cellular phosphotyrosine and transformation. *J Virol*, 1990, **64** (8): 3581~ 3589

WW domain: A Module for Protein-protein Interaction^{*}

LI Qi, HU Hong-Yu^{**}

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

Abstract WW domain is a coherent and compact domain generally composed of 38~ 40 amino acid residues. It features two tryptophan residues, and interacts specifically with proteins containing XPPXY conserved sequence. The interactions are involved in many intracellular affairs, such as non-receptor signaling, transcriptional regulation and protein degradation. Changes in these interactions will directly or indirectly interfere in normal metabolic processes and cause diseases.

Key words WW domain, protein-protein interaction, disease

^{*} This work was supported by a grant from The Chinese Academy of Sciences (STZ98-2-02).

^{**} Corresponding author. Tel: 86-21-64374430-5245, E-mail: hyhu@sunm.shnc.ac.cn

Received: June 26, 2000 Accepted: August 23, 2000