

# 大麦黄矮病毒蚜传蛋白中蚜传功能位点分析

吴兴泉, 陈士华, 张苗青, 刘应举  
(河南工业大学 生物工程学院, 河南 郑州 450001)

**摘要:** 为探讨大麦黄矮病毒(BYDV)蚜传(AT)蛋白参与介导蚜虫传播BYDV过程的关键功能区, 利用Clustal等多种序列分析工具对BYDV AT蛋白进行综合分析。结果显示, 在GAV株系、MAV株系和PAV株系的AT蛋白通读蛋白区均具有高度保守序列“VDSS”和“KRFFEY”, 两保守区之间的氨基酸序列在株系间存在特异性变异, AT蛋白的主要糖基化和磷酸化位点亦分布在该区, 表明该区域可能是决定株系间蚜传特性差异的功能区。另外分析表明, AT蛋白具有卷曲螺旋结构, 在动物细胞中主要定位于细胞核内, 其第510位氨基酸残基附近的卷曲螺旋结构不但在序列上高度保守(EYEAA)而且位置上也高度相近, 说明该结构可能在AT蛋白介导蚜虫传播过程中起调控基因表达或分子识别的作用。

**关键词:** 大麦黄矮病毒; 蚜传蛋白; 株系; 分子识别; 功能位点

**中图分类号:** S435.121.4<sup>+</sup>9      **文献标识码:** A      **文章编号:** 1004-3268(2011)08-0134-04

## Analysis of Functional Sites in Aphid Transmission Protein of Barley Yellow Dwarf Virus

WU Xing-quan, CHEN Shi-hua, ZHANG Miao-qing, LIU Ying-ju  
(College of Biotechnology, Henan University of Technology, Zhengzhou 450001, China)

**Abstract:** In order to know the key functional regions in the aphid transmission (AT) protein with which aphids transmit barley yellow dwarf virus (BYDV), BYDV AT protein was analyzed with many sequence analysis tools. The results showed that there were two highly conservative sequences “VDSS” and “KRFFEY” at the read through protein in AT protein of GAV, MAV and PAV strains, the sequence from “VDSS” to “KRFFEY” varied among different strains, and the O-glycosylation and phosphorylation sites of AT protein were mainly distributed in this region, so it was deduced that the sequence from “VDSS” to “KRFFEY” might be the functional region deciding the difference of aphid transmission characteristics between strains. The AT protein was mainly located at cell nucleus. The inserted amino acids were found in the AT proteins of three strains, and the sequences “EYEAA” and the location of the inserted amino acids near the 510 aa were highly conservative. Thus, the inserted amino acid structure of AT protein could play an important role in the process of gene expression regulation and molecular recognition during the aphid transmission of BYDV.

**Key words:** Barley yellow dwarf virus; Aphid transmission (AT) protein; Strain; Molecular recognition; Functional sites

小麦黄矮病是我国北方小麦种植区发生最广、危害最重的病毒病之一, 该病害由大麦黄矮病毒(barley yellow dwarf virus, BYDV)引起。BYDV 不能通过汁液传播, 只能通过蚜虫接种传播。依据蚜虫传播特

性、寄主范围与反应类型、对寄主的毒性等生物学特性可将 BYDV 划分为不同的株系。在我国, BYDV 的株系类型主要有 PAV、GAV、GPV、RMV 等, 其中 PAV 和 GAV 是主流株系, GPV 为我国特有的株系

类型<sup>[1]</sup>, GAV 株系与美国的 MAV 株系可产生强烈的血清学反应, 但两者的有效传播介体不同<sup>[2]</sup>。

BYDV 是正义单链 RNA 病毒, 其基因组包含 6 个开放阅读框(ORF)<sup>[3-4]</sup>, 其中 ORF3 编码外壳蛋白(CP), ORF3 中还含有一个完全重叠、编码分子量为 17kD 运动蛋白(MP)的 ORF4<sup>[5]</sup>, 紧接 ORF3 的 ORF5 可以和前者发生通读, 产生分子量为 72kD 的通读蛋白(read through protein, RTP)。很早人们便猜想, BYDV 血清学和蚜传介体的关系或许反映了病毒外壳蛋白的差异, 在获得多数黄症病毒外壳蛋白的基因序列后, 人们比较了传播介体相同的不同病毒的外壳蛋白氨基酸序列, 发现外壳蛋白不太可能与蚜传特异性有关<sup>[5]</sup>。大量对黄症病毒的研究表明, 通读蛋白与蚜虫传播密切相关, 所以现在也把它称作蚜传(aphid transmission, AT)蛋白<sup>[6-7]</sup>。王锡锋等<sup>[8]</sup>研究证明, 在麦蚜传播 BYDV 的过程中, 病毒的获得和传播需要麦蚜体内的传毒相关蛋白(病毒受体)与 CP、RTP 发生相互识别作用, 由此介导病毒穿越蚜虫体内的传播障碍。并从大麦黄矮病毒介体麦二叉蚜和麦长管蚜体内克隆获得了传毒相关蛋白, 但 AT 蛋白中发挥功能的位点以及参与识别作用的位点尚未明确。目前, 我国已有很多关于寄主与病毒互作的基因和相关蛋白的研究, 但以病毒蛋白为出发点, 通过分析不同株系间蚜传蛋白的特异性分子变异位点, 以明确蚜传蛋白中参与病毒寄主互作位点的研究尚极少报道。鉴此, 综合利用多种序列分析工具对 BYDV AT 蛋白进行蚜传功能位点分析。

1 材料和方法

1.1 供试材料

在 GenBank 数据库中进行检索, 获得已公布的 BYDV 蚜传蛋白(AT)氨基酸序列, 用于分析。

1.2 BYDV 株系间 AT 蛋白序列差异性分析

利用 Clustal、Bi Edit 等软件对 BYDV 不同株系的 AT 蛋白氨基酸序列进行多重序列比对, 分析株系间 AT 蛋白氨基酸序列差异, 明确各株系的分组情况, 找出可代表同组其他分离物的 AT 蛋白氨基酸序列, 以参加下面的株系间 AT 蛋白结构差异分析。

1.3 BYDV 株系间 AT 蛋白结构差异性分析

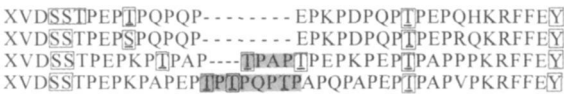
对上述筛选获得的代表不同株系的 AT 蛋白氨基酸序列进行一级结构、二级结构预测, 分析各分离物 AT 蛋白的糖基化位点、磷酸化位点、卷曲螺旋结构、在动植物细胞内的亚细胞定位等, 对来自不同株系 AT 蛋白的结构差异、保守序列区以及可能的功能位点进行比较, 以明确决定株系特异性的 AT 蛋白结构位点。

2 结果与分析

2.1 BYDV 株系间 AT 蛋白重要的分子变异区

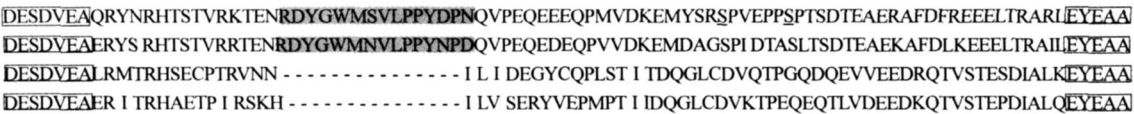
通过对 69 个 BYDV 不同分离物 AT 蛋白氨基酸序列进行多重序列比对分析, 发现在 PAV、GAV、MAV 3 个株系 AT 蛋白的通读蛋白区(CP 蛋白后)均存在高度保守序列“VDSS”和“KRFFEY”, 与 GAV 株系和 MAV 株系相比, PAV 株系在这 2 个保守区之间存在 4 个或 8 个氨基酸残基的插入序列(第 215– 218 位的 TPA(P)P 或第 216– 223 位的 TPTPQPTP, 如图 1 所示)。

另外, GAV、MAV、PAV 3 个株系均存在高度保守序列“DESDVEA”和“EYEAA”, 两保守序列区之间的序列在株系间存在高度变异性(该序列在 GAV AT 蛋白的第 427– 517 位, MAV 株系的第 431– 521 位, PAV 株系的第 435– 509 位)。与 PAV 株系相比, GAV 株系和 MAV 株系存在一段长度为 15 个氨基酸残基的插入序列(GAV 株系为 RDYGWMSVLP-PYDPN, MAV 株系为 RDYGWMNVLPYPNP, 如图 2 所示)。



1. 4 条序列从上至下分别为 GAV (许昌分离物)、MAV (NC\_003680)、PAV (EU332308)、PAV (EU332333) 在该区段的序列; 2. “-”表示缺失位点; 3. 灰色背景的氨基酸为 PAV 株系插入序列; 4. 下划线处的氨基酸为糖基化位点; 5. 带边框的氨基酸为磷酸化位点; 6. 带边框同时又有下划线的氨基酸为既可糖基化又可磷酸化的位点; 7. X 为通读蛋白中终止密码子 UAG 对应的位置

图 1 BYDV 不同株系在两高度保守区“VDSS”和“KRFFEY”之间的氨基酸序列



1. 4 条序列从上至下分别为 GAV (许昌分离物)、MAV (NC\_003680)、PAV (EU332308)、PAV (EU332333) 在该区段的序列; 2. “-”表示缺失位点; 3. 下划线处为糖基化位点; 4. 加边框处为株系间高度保守序列

图 2 BYDV 不同株系在两高度保守区“DESDVEA”和“EYEAA”之间的氨基酸序列

2.2 BYDV 不同株系间 AT 蛋白的结构差异

通过多重序列比对分析证明, GAV 株系不同分离物间 AT 蛋白氨基酸序列相似性均在 97% 以上, 选择 GAV 许昌分离物(氨基酸残基数量 CP/AT 为 195/ 653) 进行结构分析。PAV 株系不同分离物间 AT 蛋白氨基酸序列变异性较高, 依据 AT 蛋白氨基酸序列可将 PAV 分为两大亚组, 因此, 分别从每个亚组中选一个最具代表性的分离物( GenBank 序列号分别为 EU332308 和 EU332333, 氨基酸数量 CP/AT 分别为 199/ 645、200/ 649), 对 AT 蛋白氨基酸序列进行结构分析。MAV 只有一条序列, GenBank 序列号为 NC\_003680(氨基酸数量 CP/AT 为 199/ 656), 用于结构分析。

2.2.1 AT 蛋白的糖基化位点 利用 <http://www.bs.dtu.dk/services/> 网站中提供的神经网络模型, 进行 AT 蛋白糖基化位点分析, 结果见图 1 和图 2。PAV、GAV 株系的 AT 蛋白各存在 4 个糖基化位点, 而 MAV 株系仅有 2 个。PAV、MAV 株系 AT 蛋白的糖基化位点均出现在第 200– 240 位氨基酸残基之间, 而 GAV 株系的 4 个糖基化位点中 2 个出现在该区间。由图 1 可知, 该区域正处于高度保守序列“VDSS”与“KRFFEY”之间, 此区间 PAV 株系具有株系特异性的插入序列“TPA(P)P”或“TPTPQTP”, 在插入序列中的 T 全部为糖基化位点。说明这段在株系间具有明显分子差异的序列对蛋白的糖基化起着重要作用, 也对蛋白的性质产生重要影响。

2.2.2 AT 蛋白的磷酸化位点 利用 <http://www.bs.dtu.dk/services/> 网站中提供的预测模型进行 AT 蛋白磷酸化位点分析。结果表明, GAV 株系 AT 蛋白中磷酸化位点的数量( 64 个) 远多于 PAV 株系( 43~ 52 个), MAV 株系的磷酸化位点数量( 60 个) 与 GAV 株系相近。分析发现, 在不同株系高度保守序

列“VDSS”与“KRFFEY”之间, 对 SS 和 Y 的磷酸化位点高度一致, 而对 T 的磷酸化则出现明显差异( 图 1)。

2.2.3 AT 蛋白的卷曲螺旋结构 利用 <http://www.h.embnet.org/index.html> 网站中的 CIOLS 工具, 进行 AT 蛋白卷曲螺旋结构的分析。结果表明, 3 个株系在第 510 位氨基酸左右均存在一个卷曲螺旋, 但株系间卷曲螺旋的宽度不同, GAV 株系的宽度为 28 个氨基酸残基, MAV 株系为 21 个氨基酸残基, PAV 株系为 14 个氨基酸残基。

采用 <http://pbil.ibp.fr/> 网站中的二级结构分析工具对 AT 蛋白卷曲螺旋结构的二级结构进行分析。结果表明, GAV 株系 AT 蛋白的卷曲螺旋结构位于第 493– 518 位间, 序列为“TEAERAFD-FREEELTRARLEYEAATE”, MAV 株系的卷曲螺旋结构在第 499– 522 位间, 序列为“EAEKAFDL-KEEELTRAILEYEAAT”, PAV 株系 EU332308 分离物的卷曲螺旋结构在第 500– 512 位间和第 620– 644 位间, 序列分别为“DIALKEYEAIAE”和“REQ-LREYTRIRNTLGLTAAKEYKAR”, 而 PAV 株系 EU332333 分离物的卷曲螺旋结构在第 508– 517 位间和第 622– 648 位间, 序列分别为“ALMEYEAATA”和“MTREQLAEYTRIRKSL-GLTAAKEYKAQ”。其中“EYEA”为株系间高度保守序列( 上述序列中下划线部分)。

2.3 BYDV AT 蛋白的亚细胞定位

采用 <http://www.lfports.org/> 网站中的 WLFPSORT 工具对 BYDV AT 蛋白的亚细胞定位进行分析。结果表明, 3 个株系的 AT 蛋白在植物和动物细胞的亚细胞定位非常相似, 在植物细胞中主要在叶绿体和线粒体内, 在动物细胞中主要在细胞核内( 表 1)。

表 1 BYDV AT 蛋白的亚细胞定位

株系	AT 蛋白在植物体内各亚细胞定位的得分	AT 蛋白在动物体内各亚细胞定位的得分
GAV 许昌分离物	叶绿体: 6.5, 叶绿体-线粒体: 5.5, 线粒体: 3.5, 细胞核: 3.0	细胞核: 23.5, 胞浆: 13.5, 线粒体: 6.0
MAV( NC_003680)	叶绿体: 5.5, 叶绿体-线粒体: 7.0, 线粒体: 7.5	细胞核: 22.0, 线粒体: 9.0
PAV( EU332333)	叶绿体: 7.5, 叶绿体-线粒体: 6.0, 线粒体: 3.5, 细胞核: 1.0, 胞浆: 1.0	细胞核: 26.5, 胞浆: 15.9, 线粒体: 3.0

3 结论与讨论

通过上述分析可知, BYDV AT 蛋白存在 2 个重要的作用区, 一是在 CP 蛋白后通读蛋白区的高度保守序列“VDSS”和“KRFFEY”之间, 该区域在株系间存在明显的插入突变, 并是重要的糖基化位点和磷酸化位点聚集区, 可能是决定株系特性差异的关键位点。二是在第 510 位氨基酸残基左右的高度保守序列“EYEA”, 该序列区 3 个株系均存在

卷曲螺旋结构。GAV 株系和 MAV 株系的 AT 蛋白可能为跨膜蛋白, 而 PAV 株系的 AT 蛋白不是跨膜蛋白, 它们在植物细胞内主要位于叶绿体和线粒体, 在动物细胞内主要位于细胞核。大麦黄矮病毒 GAV 株系、PAV 株系和 MAV 株系均可由麦长管蚜有效传播, 但 GAV 株系还可由麦二叉蚜有效传播, PAV 株系还可由禾缢管蚜有效传播, 三者蚜传特性上既存在共性, 又有差异。由 ORF3 通读至 ORF5 形成的通读蛋白 RTP 位于病毒粒体的表面, 现认为

通读蛋白决定着病毒的可传播性和介体专化性,麦蚜获毒过程不存在病毒与传毒蛋白的相互识别,只有在麦蚜的传毒过程,即病毒粒体由血淋巴进入唾液腺的过程,才存在病毒粒子与麦蚜体内传毒蛋白(受体)的相互识别。目前已经明确了麦长管蚜和麦二叉蚜体内的传毒相关蛋白,并且证明这些蛋白主要位于蚜虫头部组织唾液腺膜上。然而,BYDV RTP 中参与寄主识别的功能位点尚未明确<sup>[9]</sup>。

参与识别的生物大分子主要是蛋白质和多糖,本研究通过对 BYDV 3 个株系 RTP 序列差异、结构特征、糖基化位点、磷酸化位点、亚细胞定位等分析,证明 RTP 中存在 2 个可能参与蚜传的功能区。第一是在高度保守序列“VDSS”和“KRFFEY”之间,该区域在氨基酸序列、糖基化位点等方面存在明显的株系特异性差异。PAV 株系 AT 蛋白的 4 个糖基化位点、MAV 株系的 2 个糖基化位点全部分布在此序列区,而 GAV 株系的 4 个糖基化位点有 2 个分布于此。分析证明,3 个株系的 AT 蛋白均是糖蛋白,其中糖基部分对分子识别具有重要的影响,由此说明“VDSS”与“KRFFEY”之间的氨基酸序列区可能在 AT 蛋白参与寄主互作过程起重要的识别作用。另一个可能参与寄主互作过程的功能区是 AT 蛋白通读区的卷曲螺旋结构。3 个株系的 AT 蛋白在第 510 氨基酸附近均具有卷曲螺旋结构,均含有高度保守的“EYEEA”序列。在 GAV、MAV 株系 AT 蛋白中“EYEEA”上游存在一段保守的插入序列“RDYGWMS(N)VLPPYDPN(D)”,经分析该序列无明显的修饰和特征性结构,但由于该序列的插入将卷曲螺旋结构恰好后移至第 510 位氨基酸附近,保证其在位置上的保守性。具有卷曲螺旋结构的蛋白在机体内一般执行着代谢调控、膜通道、分子伴侣等功能。例如跨膜蛋白在细胞与外界环境之间的物质与信息交换中发挥着重要功能,而多种跨膜蛋白的跨膜部分均含有卷曲螺旋结构。另外,这一结构在细胞及分子的相互识别过程中起着确定性的作用。如 SNARE 蛋白,在细胞与细胞相互作用时可由来自 2 个特定细胞的 4 条多肽链形成一个卷曲螺旋四聚体,从而实现细胞之间的特异性识别。在细胞的 DNA 转录过程中,起分子识别作用的多种转录因子也含有卷曲螺旋结构,对特定 DNA 分子的识别及转录过程具有关键作用<sup>[10]</sup>。作为 AT 蛋白中的卷曲螺旋结构,可能也对蚜传识别过程起重要作用。

分析发现,在 BYDV AT 蛋白的卷曲螺旋结构区,3 个株系既具有共性(即在 510 位氨基酸附近均存在卷曲螺旋结构)又存在差异(GAV 株系和 MAV 株系只具有 1 个卷曲螺旋结构,但两者的宽

度不同,PAV 株系具有 2 个卷曲螺旋结构)。另外,3 个株系在 AT 蛋白糖基化位点方面同样既存在共性(即在第 200–240 位氨基酸处均具有糖基化位点)又存在差异(数量和位置不同)。这些均与 3 个株系的有效传毒蚜虫种类相吻合,因此,推测这 2 个区域可能在蚜传过程参与了寄主互作,相关研究尚未见报道。

具有卷曲螺旋结构的蛋白很多可作为转录因子参与基因表达调控,由于 AT 蛋白是一种具有卷曲螺旋的蛋白质,其在动物细胞内主要定位于细胞核内,预示着 AT 蛋白在动物体内可能作为转录因子,参与动物细胞某些基因的转录与表达,相关内容有待于进一步深入研究。

#### 参考文献:

- [1] 张文蔚,成卓敏.大麦黄矮病毒 GPV 株系基因组末端序列的克隆和分析[J].中国农业科技导报,2009,11(1):102-107.
- [2] 周广和,张淑香,钱幼亭.小麦黄矮病毒 4 种株系鉴定与应用[J].中国农业科学,1987,20(4):7-12.
- [3] Miller W A, Dineshkumar S P, Paul C P. Lute virus gene expression[J]. Critical Review in Plant Sciences, 1995, 14(1):179-211.
- [4] May M A, Ziegler G V. Molecular biology of lute viruses[J]. Advances in Virus Research, 1996, 46(1):413-460.
- [5] Vincent J R, Ueng P P, Lister R M, et al. Nucleotide sequence of coat protein genes for three isolates of barley yellow dwarf virus and their relationships to the lute virus coat protein sequences[J]. Journal of General Virology, 1990, 71:2791-2799.
- [6] Chay C A, Gunasinge U B, Dinesh Kumar S P, et al. Aphid transmission and systemic plant infection determinants of barley yellow dwarf lute virus-PAV are contained in the coat protein read-through domain and 17kDa protein, respectively[J]. Virology, 1996, 219(3):57-65.
- [7] Wang M B, Cheng Z, Keese P, et al. Comparison of the coat protein, movement protein and RNA polymerase gene sequences of Australian, Chinese, and American isolates of barley yellow dwarf virus transmitted by *Rhopalosiphum padi*[J]. Archives of Virology, 1998, 143(5):1005-1013.
- [8] 王锡锋,周广和.大麦黄矮病毒介体麦二叉蚜和麦长管蚜体内传毒相关蛋白的确定[J].科学通报,2003,48(15):1671-1675.
- [9] 梁小波,鲁瑞芳,吴云峰,等.植物病毒昆虫介体传播的研究进展[J].生物工程进展,2001,21(4):14-17.
- [10] 魏香,曾宪纲,周海梦.蛋白质结构中卷曲螺旋的研究进展[J].中国生物化学与分子生物学学报,2004,20(5):565-571.