



## 研究报告

### Research Report

# 植物甜蛋白马宾灵(Mabinlin II)的二级结构及其 B 细胞抗原表位预测

姚晶<sup>1</sup>, 顾文亮<sup>1</sup>, 夏启玉<sup>2</sup>, 胡新文<sup>1</sup>, 郭建春<sup>2</sup>

1 海南大学农学院, 海口, 570228;

2 中国热带农业科学院热带生物技术研究所, 海口, 571101

✉ 通讯作者: jianchunguoh@163.com; ✉ 作者

计算分子生物, 2012 年, 第 1 卷, 第 1 篇 doi: 10.5376/cmb.cn.2012.01.0001

收稿日期: 2012 年 03 月 12 日

接受日期: 2012 年 05 月 18 日

发表日期: 2012 年 05 月 25 日

本文首次发表在《基因组学与医学生物学》(2012 年第 31 卷第 2 期 154-159 页)上。现依据版权所有人授权的许可协议, 采用 Creative Commons Attribution License 协议对其进行授权, 再次发表与传播。只要对原作有恰当的引用, 版权所有人允许并同意第三方无条件的使用与传播。

建议最佳引用格式:

引用格式(中文):

姚晶等, 2012, 植物甜蛋白马宾灵(Mabinlin II)的二级结构及其 B 细胞抗原表位预测, 计算分子生物(online) Vol.1 No.1 pp.1-6 (doi: 10.5376/cmb.cn.2012.01.0001)

引用格式(英文):

Yao et al., 2012, Prediction of the Secondary Structure and B Cell Epitopes of the Plant Sweet Protein Mabinlin II, Jisuan Fenzi Shengwu (online) (Computational Molecular Biology) Vol.1 No.1 pp.1-6 (doi: 10.5376/cmb.cn.2012.01.0001)

**摘要** 马宾灵(Mabinlin II)是中国所特有的植物甜蛋白, 在目前已知的七种植物甜蛋白中具有最佳的热稳定性, 将其作为新型甜味剂有着广阔的市场前景。本研究以马宾灵的氨基酸序列为基础, 采用 Garnier-Robson 方法和 Chou-Fasman 方法预测马宾灵的二级结构, 采用 Karplus-Schulz 方法预测马宾灵的蛋白质骨架区柔韧性, 采用 Kyte-Doolittle 方法预测马宾灵的蛋白质疏水性, 采用 Emini 方法预测马宾灵的蛋白质表面可能性, 采用 Jameson-Wolf 方法预测马宾灵的 B 细胞抗原表位。分析结果表明, 马宾灵的 A 链多形成转角和无规则卷曲结构, A 链的 N 端 27~30 区段形成较柔软的蛋白质骨架结构, A 链的 N 端 4~6、7~9、10~11、12~14、20~22、23~26 和 30~33 区段为抗原位点的可能性较大; 马宾灵的 B 链多形成  $\beta$  折叠和转角结构, B 链的 N 端 0~7、9~13、28~29、33~34、40~45、53~54 和 55~57 区段为抗原位点的可能性较大。本研究以生物信息学手段分析预测马宾灵的二级结构及其 B 细胞抗原表位, 为设计马宾灵多克隆抗体以检测马宾灵在不同生物反应器中的表达研究提供参考数据。

**关键词** 植物甜蛋白; 马宾灵; 二级结构; B 细胞抗原表位

## Prediction of the Secondary Structure and B Cell Epitopes of the Plant Sweet Protein Mabinlin II

Yao Jing<sup>1</sup>, Gu Wenliang<sup>1</sup>, Xia Qiyu<sup>2</sup>, Hu Xinwen<sup>1</sup>, Guo Jianchun<sup>2</sup>

1 College of Agriculture, Hainan University, Haikou, 570228;

2 Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou, 571101

✉ Corresponding author: jianchunguoh@163.com; ✉ Authors

**Abstract** Mabinlin II is the unique plant sweet protein in China, and it has the best thermal stability of the seven plant sweet proteins which have been known today. Mabinlin II may have a promising market to be a new food sweetener. Based on the amino acid sequence of Mabinlin II, the secondary structure of it was predicted with the methods of Garnier-Robson and Chou-Fasman, and the flexibility of the backbone structure of it was predicted with the Karplus-Schulz method, and the hydrophobicity of it was predicted with the Kyte-Doolittle method, and the possibility of surface area of it was predicted with the Emini method, the B-cell epitopes of it was predicted with the Jameson-Wolf method. The results show that the structure of the A-chain of Mabinlin II mostly consists of turns and random coils, and the 27~30 amino acid section from the N-terminal become a more flexible skeleton structure, while the 4~6, 7~9, 10~11, 12~14, 20~22, 23~26 and 30~33 amino acid sections from the N-terminal are probably the key antigenic sites. The structure of the B-chain of Mabinlin II mostly consists of  $\beta$ -sheet and turn regions, and the 0~7, 9~13, 28~29, 33~34, 40~45, 53~54 and 55~57 amino acid sections from the N-terminal may be the key antigenic sites. The bioinformatics analysis of this protein may provide a reference for the design of polyclonal antibody against the Mabinlin II which would be used to detect the expression of Mabinlin II in different bioreactors.

**Keywords** Plant sweet protein; Mabinlin II; Secondary structure; B cell epitopes

植物甜蛋白是指自然界中存在的具有甜味或可引起甜味产生的蛋白质, 它们通常来源于热带植物



的果实或种子(施建科, 1998, 化学通报, 8: 22-26)。目前, 世界上已发现的植物甜味蛋白有 7 种, 其中植物甜蛋白马宾灵(Mabinlin II)是马槟榔种子中的贮藏蛋白。马槟榔是我国特有的、仅分布于我国热带、亚热带地区的珍稀濒危野生果树资源(胡新文等, 1998)。马宾灵在马槟榔种子中有 4 种同系物蛋白, 其中仅 Mabinlin II 具有最佳耐酸性和热稳定性因而也研究的最为深入。马宾灵(Mabinlin II)具有高甜度, 无毒、非热源性等特点, 将其作为新型食品甜味剂的应用广泛(范长胜, 2001)。

由于野生的马槟榔资源极其有限且在不断的萎缩, 以基因工程的手段开发应用马宾灵(Mabinlin II)的研究遂成为研究的热点。在 2001 年, 刘敬梅等(2001)人将马宾灵基因转入莴苣中表达; 2004 年, 李晓东等(2004)人将马宾灵基因转入番茄中表达; 2009 年, 顾文亮等(2009a; 2009b)人将重组马宾灵基因转入拟南芥与大肠杆菌中表达; 2011 年, 张菲菲等(2011)人构建了含重组马宾灵基因的植物表达载体, 拟转入人参果中表达。随着技术手段的不断进步, 将马宾灵及其重组基因导入多中生物反应器中表达已经成为可能, 因而制备专一和灵敏马宾灵多克隆抗体用于马宾灵基因工程产物的检测显得尤为必要。

本研究以马宾灵(Mabinlin II)的氨基酸序列为基础, 分别对马宾灵的 A 链与 B 链进行蛋白质的二级结构、骨架区柔韧性、亲水性、表面可能性和 B 细胞抗原表位预测, 经综合分析后得到马宾灵 A 链与 B 链的可能抗原决定位点, 为人工制备特异性的马宾灵多克隆抗体提供一定的参考数据。

## 1 结果与分析

### 1.1 马宾灵(Mabinlin II)的氨基酸序列

成熟的马宾灵(Mabinlin II)蛋白为 A 链与 B 链通过链内与链间二硫键结合的二聚体蛋白结构。马宾灵前体蛋白的信号肽、N-端肽、连接肽以及 C-端肽均在体内翻译后加工过程中被剪接掉, 有甜味活性的马宾灵蛋白是由 A 链含 33 个氨基酸残基与 B 链含 72 个氨基酸残基(图 1)通过非共价作用紧密结合在一起的。

```
A-chain: 1 ELWRCQRQFL QHQLRACQR FIIHRAQFGG QPD  
B-chain: 1 EPRRPA LRQC CNQLRQVDRP CVCPLVRQAA QQVLQRQIIQ  
41 GPQQLRRLFD AARNLPNIGN IPNIGACPF R AW
```

图 1 马宾灵(Mabinlin II)的氨基酸序列

Figure 1 Amino acid sequence of Mabinlin II

### 1.2 马宾灵(Mabinlin II)的二级结构预测

采用 Garnier-Robson 方法进行马宾灵(Mabinlin II)的二级结构预测, 经分析发现其 A 链没有形成  $\alpha$ -螺旋结构, 而在 N-端的 20~22 区段形成  $\beta$ -折叠结构(图 2); 其 B 链没有形成  $\alpha$ -螺旋结构, 而在 N-端的 7~10、12~15、18~19、21~40、42~51、57~62、65~70 多个区段形成  $\beta$ -折叠结构(图 2)。

采用 Chou-Fasman 方法进行马宾灵(Mabinlin II)的二级结构预测, 经分析发现其 A 链在 N-端的 0~27 区段形成  $\alpha$ -螺旋结构但没有形成  $\beta$ -折叠结构(图 3); 其 B 链(图 3)在 N-端的 24~40 和 44~54 区段形成  $\alpha$ -螺旋结构, 在 N-端的 13~27 以及 29~40 区段形成  $\beta$ -折叠结构。

综合 Garnier-Robson 方法和 Chou-Fasman 方法的预测结果显示, 马宾灵(Mabinlin II)的 A 链形成  $\alpha$ -螺旋结构和  $\beta$ -折叠结构的可能性较小; 马宾灵(Mabinlin II)的 B 链形成  $\alpha$ -螺旋结构的可能性较小, 而在其 N-端的 12~15、18~19、21~27 及 29~40 多个区段形成  $\beta$ -折叠结构的可能性较大。

### 1.3 马宾灵(Mabinlin II)的可能抗原位点预测

#### 1.3.1 马宾灵(Mabinlin II)的骨架区柔韧性预测

采用 Karplus-Schulz 方法进行马宾灵(Mabinlin II)的蛋白质骨架区柔韧性预测, 经分析发现其 A 链 N-端的 27~30 区段(图 4)和其 B 链 N-端的 0~7、10~12、15~18、40~42、52~54 和 55~57 区段(图 4)形成蛋白骨架柔韧性区域的可能性较大。

#### 1.3.2 马宾灵(Mabinlin II)的亲水性预测

采用 Kyte-Doolittle 方法进行马宾灵(Mabinlin II)的亲水性预测, 经分析发现在其 A 链 N-端的 0~17 和 19~33 区段形成两个所跨区段较大的亲水性区域结构(图 5); 在其 B 链的 N-端的 0~19、27~28、30~34、37~51、52~54 和 55~57 区段形成多个亲水性区域结构(图 5)。

#### 1.3.3 马宾灵(Mabinlin II)的表面可能性预测

采用 Emini 方法进行马宾灵(Mabinlin II)的蛋白质表面可能性预测, 经分析发现在其 A 链的 N-端 4~6、7~9、10~14、20~22、23~26 和 30~33 区段位于蛋白质表面的可能性较大(图 6); 在其 B 链的 N-端 0~6、14~18、28~30、32~34、41~45 和 54~55 区段位于蛋白质表面的可能性较大(图 6)。

#### 1.3.4 马宾灵(Mabinlin II)的 B 细胞抗原表位预测

采用 Jamesorr Wolf 方法进行马宾灵(Mabinlin II)



的 B 细胞抗原表位预测, 经分析发现在其 A 链的 N-端 0~11、12~19 和 20~33 区段为抗原指数较

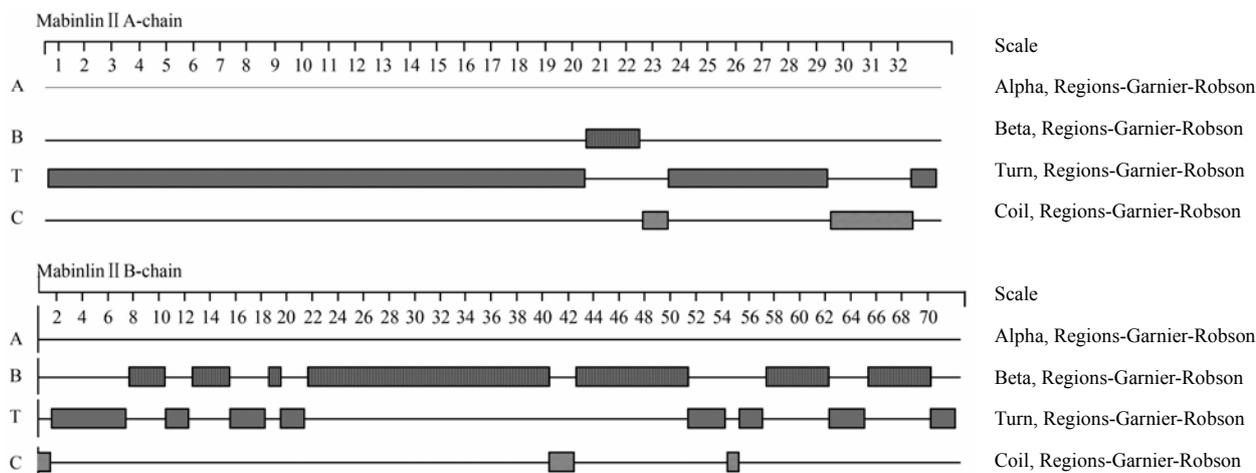


图 2 按 Garnier-Robson 方法预测的马宾灵(Mabinlin II) A 链及 B 链的二级结构

注: A:  $\alpha$ -螺旋区段; B:  $\beta$ -折叠区段; T: 转角区域; C: 无规则卷曲区域

Figure 2 Secondary structure of Mabinlin II A-chain and B-chain predicted with Garnier-Robson method

Note: A:  $\alpha$ -helix segments; B:  $\beta$ -sheet segments; T: Corner region; C: Random coil region

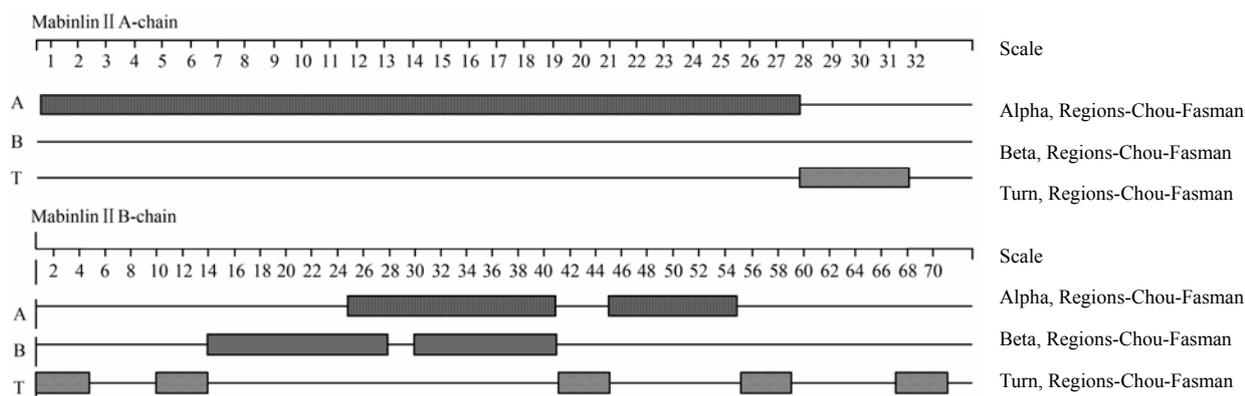


图 3 按 Chou-Fasman 方法预测的马宾灵(Mabinlin II)A 链及 B 链的二级结构

注: A:  $\alpha$ -螺旋区段; B:  $\beta$ -折叠区段; T: 转角区域

Figure 3 Secondary structure of Mabinlin II A-chain and B-chain predicted with Chou-Fasman method

Note: A:  $\alpha$ -helix segments; B:  $\beta$ -sheet segments; T: Corner region

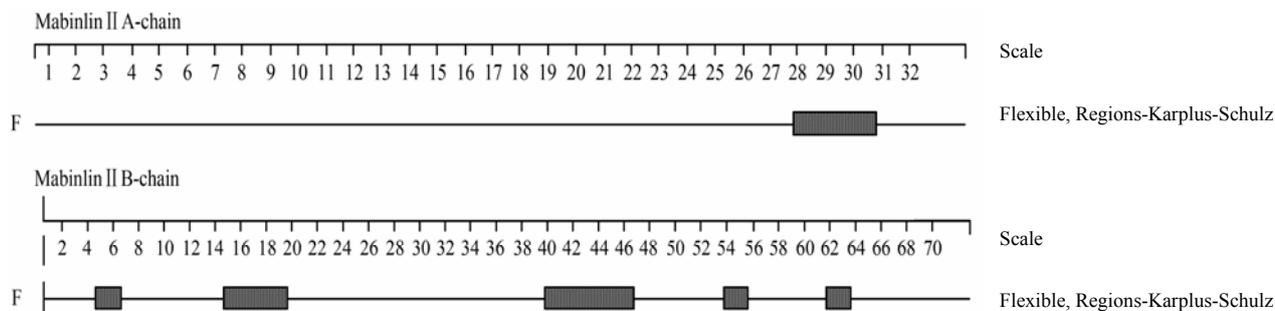


图 4 按 Karplus-Schulz 方法预测的马宾灵(Mabinlin II) A 链及 B 链的柔韧性区域

注: F: 柔韧性区域

Figure 4 The flexible region of Mabinlin II A-chain and B-chain predicted with Karplus-Schulz method

Note: F: Flexible region

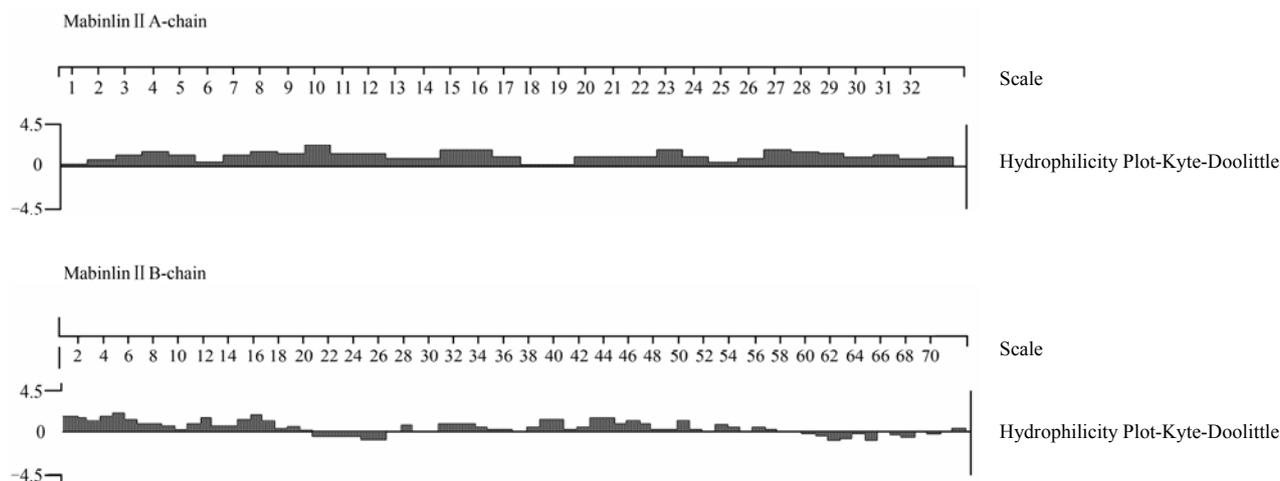


图 5 按 Kyte-Doolittle 方法预测的马宾灵(Mabinlin II)A 链及 B 链的亲水性分析  
Figure 5 The hydrophilic analysis of Mabinlin II A-chain and B-chain predicted with Kyte-Doolittle method

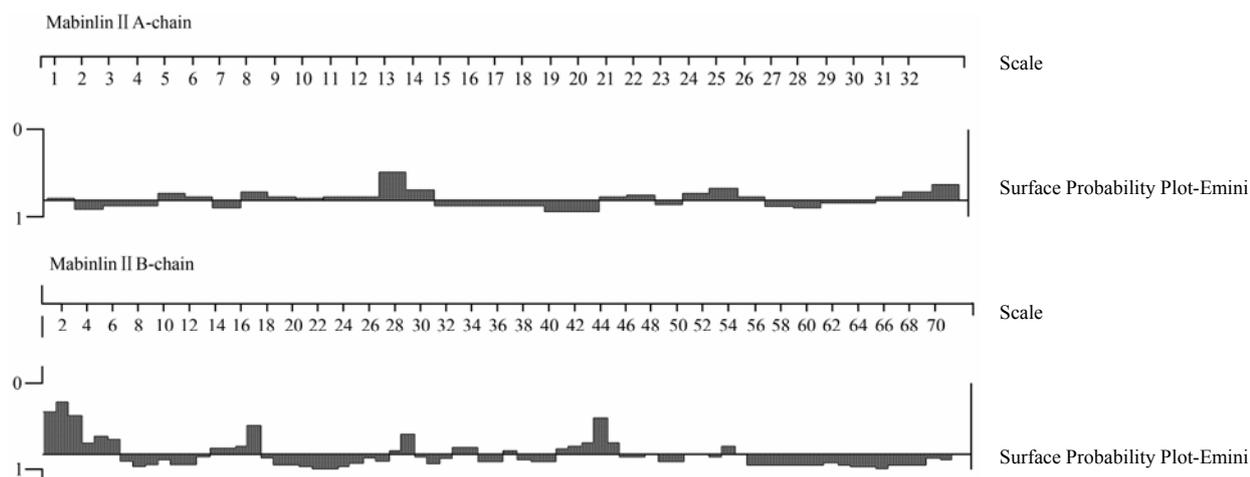


图 6 按 Emini 方法预测的马宾灵(Mabinlin II)A 链及 B 链的表面可能性区域  
Figure 6 The possibility of surface area of Mabinlin II A-chain and B-chain predicted with Emini method

高的区段(图 7); 在其 B 链的 N-端 0~7、9~21、28~29、33~34、40~57、62~63 和 66~68 区段为抗原指数较高的区段(图 7)。

综合 Karplus-Schulz 方法的蛋白质骨架区柔韧性预测、Kyte-Doolittle 方法的蛋白质亲水性预测、Emini 方法的蛋白质表面可能性预测和 Jamesorr Wolf 方法的 B 细胞抗原表位预测, 经分析发现马宾灵(Mabinlin II)A 链的 N-端 4~6、7~9、10~11、12~14、20~22、23~26 和 30~33 区段为抗原位点的可能性较大; 马宾灵(Mabinlin II)B 链的 N-端 0~7、9~13、28~29、33~34、40~45、53~54 和 55~57 区段为抗原位点的可能性较大。

## 2 讨论

随着生物信息学技术的不断发展, 应用生物信息学对蛋白质的结构与功能进行分析预测的方法已经日益完善。蛋白质的二级结构是影响蛋白质的抗原表位的主要因素之一, 相比于  $\alpha$ -螺旋与  $\beta$ -折叠等二级结构的较高化学键键能下所维持的较牢固的蛋白质高级结构, 蛋白质的转角及无规则卷曲等二级结构通常比较松散, 易于发生扭曲或盘绕并展示在蛋白质的表面, 因此在该氨基酸残基区域通常含有优势的蛋白质抗原表位。另一方面, 由于在多肽折叠为成熟的蛋白质的过程中通常会发生构象变化, 某些具有较高抗原指数的区段会被外层的氨基酸残基或糖基所屏蔽, 故在采用 Jameson-Wolf

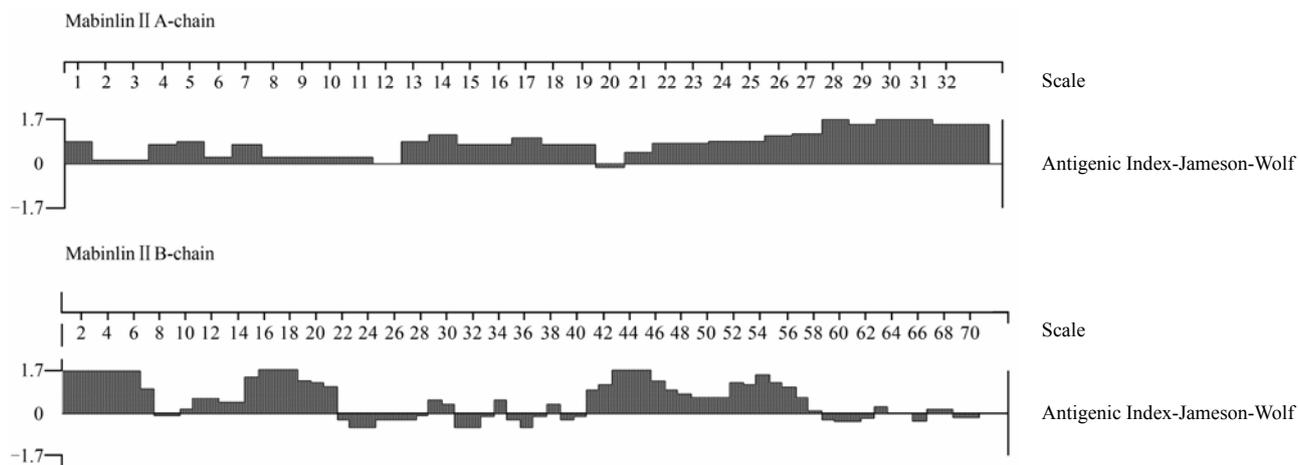


图 7 按 Jamesorr Wolf 方法预测的马宾灵(Mabinlin II) A 链及 B 链的抗原指数  
Figure 7 The antigen index of Mabinlin II protein A-chain and B-chain predicted with Jamesorr Wolf method

方法预测蛋白质的抗原指数时, 抗原指数较高的区段并不一定就是优势抗原表位区域。根据本研究分析预测结果, 马宾灵 A 链的近 C-端和中间区段多形成较多的无规则卷曲结构, 同时该区段的抗原指数较高且其亲水性指数和蛋白表面可能性也比较大, 故该区段为优势抗原表位的可能性较大; 虽然 A 链的其它一些区段如 0~1、4~6、7~9 和 10~11 区段的抗原指数也较高, 但是其亲水性和蛋白质表面可能性指数偏低, 综合分析这些区段作为抗原表位的可能性不大。马宾灵 B 链的 N-端和中间区段的抗原指数较高, 且亲水性指数和蛋白质表面可能性也较大, 故该区段为优势抗原表位的可能性较大。

### 3 材料与方法

#### 3.1 马宾灵(Mabinlin II)的氨基酸序列

马宾灵(Mabinlin II)的 A 链与 B 链的氨基酸序列分别通过 GeneBank 序列编号为 D83997.1 的马宾灵(Mabinlin II) mRNA 序列推导, 其 A 链含有 33 个氨基酸残基, 其 B 链含有 72 个氨基酸残基。

#### 3.2 马宾灵(Mabinlin II)的二级结构的预测

应用 DNA Star Protean 软件提供的模块分别对马宾灵(Mabinlin II)的 A 链和 B 链的氨基酸序列进行分析, 采用 Garnier-Robson 方法和 Chou-Fasman 方法预测马宾灵(Mabinlin II)的二级结构。

#### 3.3 马宾灵(Mabinlin II)的可能抗原位点的预测

应用 DNA Star Protean 软件提供的模块分别对马宾灵(Mabinlin II)的 A 链和 B 链进行可能抗原位点分析, 采用 Karplus-Schulz 方法预测马宾灵

(Mabinlin II) 的蛋白质骨架区柔韧性, 采用 Kyte-Doolittle 方法预测马宾灵(Mabinlin II)的蛋白质疏水性, 采用 Emini 方法预测马宾灵(Mabinlin II)的蛋白质表面可能性, 采用 Jameson-Wolf 方法预测马宾灵(Mabinlin II)的 B 细胞抗原表位, 最后综合分析马宾灵(Mabinlin II)的可能抗原位点。

#### 作者贡献

姚晶负责该论文的设计、操作及文章的撰写; 顾文亮和夏启玉协助实验设计及数据分析; 胡新文老师负责实验指导; 郭建春老师是本论文的通讯作者, 负责整个论文数据分析及文章撰写的整个过程。

#### 致谢

本研究由国家自然科学基金(31060123)和 2010 年海南省研究生创新科研课题资助

#### 参考文献

- Fan C.S., 2001, Research on development and application of sweet-tasting proteins, *Shipin Yu Fajiao Gongye (Food and Fermentation Industries)*, (12): 50-54 (范长胜, 2001, 甜蛋白的开发与应用研究, *食品与发酵工业*, (12): 50-54)
- Gu W.L., Hu X.W., and Guo J.C., 2009a, Reorganization and structure analysis of MBL II gene, *Zhiwu Shenglixue Tongxun (Plant Physiology Communications)*, 45(4): 333-339 (顾文亮, 胡新文, 郭建春, 2009a, 马槟榔甜蛋白基因(MBL II)的剪切重组和结构分析, *植物生理学通讯*, 45(4): 333-339)
- Gu W.L., Xia Q.Y., Hu X.W., and Guo J.C., 2009b, Expression of recombinant mabinlin II in *Escherichia coli* induced by IPTG, *Jiyinzuxue Yu Yingyong Shengwuxue (Genomics)*



- and Applied Biology), 28(3): 439-444 (顾文亮, 夏启玉, 胡新文, 郭建春, 2009b, 重组马槟榔甜蛋白 MabinlinII 在大肠杆菌中的表达, 基因组学与应用生物学, 28(3): 439-444)
- Hu X.W., Guo J.C., and Zheng X.Q., 1998, Cloning and sequencing of cDNA encoding sweet tasting plant protein Mabinlin II, Shengming Kexue Yanjiu (Life Science Research), 2(3): 189-193 (胡新文, 郭建春, 郑学勤, 1998, 植物甜蛋白 Mabinlin II cDNA 的克隆与序列分析, 生命科学研究, 2(3): 189-193)
- Li X.D., Liu J.M., Chen H., Cheng Z.H., and Chai M., 2004, Genetic transformation of sweet protein MBL II on tomato, Xibei Zhiwu Xuebao (Acta Botanica Boreale—Occidentalia Sinica), 24(5): 808-811 (李晓东, 刘敬梅, 陈杭, 程智慧, 柴敏, 2004, 甜蛋白基因 MBL II 对番茄的遗传转化, 西北植物学报, 24(5): 808-811)
- Liu J.M., Chen D.M., and Chen H., 2001, Genetic transformation and plant regeneration of lettuce with sweet protein gene MBL II, Yuanyi Xuebao (Acta Horticulturae Sinica), (3): 246-250 (刘敬梅, 陈大明, 陈杭, 2001, 甜蛋白基因 MBL II 对莴苣的遗传转化, 园艺学报, (3): 246-250)
- Zhang F.F., Peng Z.Y., and Zhang J.W., 2011, Recombination of MBL gene and the construction of plant fruit-specific expressed vector, Gansu Nongye Daxue Xuebao (Journal of Gansu Agricultural University), (1): 49-54, 62(张菲菲, 彭治云, 张金文, 2011, 马槟榔 MBLII 基因重组及果实特异表达载体构建, 甘肃农业大学学报, (1): 49-54, 62)