

家蚕蛹糖水解酶 18 家族几丁质酶的抗原表位预测与结构分析

赵祥杰^{1,2}, 李琳¹, 杨荣玲², 毕艳红², 王朝宇², 赵立², 邝哲师³, 穆丽霞³, 李冰¹

(1. 华南理工大学轻工与食品学院, 广东广州 510640) (2. 淮阴工学院生命科学与食品工程学院, 江苏淮安 223003) (3. 广东省农业科学院蚕业与农产品加工研究所, 广东广州 510610)

摘要: 家蚕蛹是蚕桑产业中的重要副产物资源, 因营养丰富、蛋白质含量高、富含不饱和脂肪酸等优点, 常用作人们的休闲食品, 近年来, 随着以蚕蛹为原料的多元化产品开发研究的不断发展, 蚕蛹资源的利用越来越成为增加蚕桑产业效益的重要途径, 但因食用或饲用家蚕蛹而引发的过敏症状较为严重, 制约了其在食品和饲料中的应用。已有报道家蚕蛹中的糖水解酶 18 家族几丁质酶 (BmChi) 是一种重要变应原物质, 但至今尚无其三维结构数据, 本文通过生物信息学软件分析, 对家蚕蛹糖水解酶 18 家族几丁质酶的同源性、二级结构、亲疏水性、可塑性、表面可及性和抗原性等进行了综合分析, 确定其抗原表位位点, 为深入开展该变应原的免疫学特性研究打下基础, 并为家蚕蛹蛋白的脱敏研究提供理论基础。

关键词: 家蚕蛹; 糖水解酶 18 家族几丁质酶; 食物变应原; 抗原表位

文章编号: 1673-9078(2016)11-85-90

DOI: 10.13982/j.mfst.1673-9078.2016.11.013

Antigenic Epitope Prediction and Structural Analysis of Glycoside Hydrolase Family 18 Chitinase Allergen from Domestic Silkworm (*Bombyx mori*) Pupa

ZHAO Xiang-Jie^{1,2}, LI Lin¹, YANG Rong-Ling², BI Yan-Hong², WANG Zhao-Yu², ZHAO Li², KUANG Zhe-Shi³, MU Li-Xia³, LI Bing¹

(1.College of Light Industry and Food Sciences, South China University of Technology, Guangzhou 510640, China)

(2.School of Life Science and Food Engineering, Huaiyin Institute of Technology, Huai'an 223003, China) (3.Sericultural & Agri-food Research Institute Guangdong Academy of Agricultural Sciences, Guangzhou 510610, China)

Abstract: Domestic silkworm (*Bombyx mori*) pupa is an important byproduct of the sericulture industry and is commonly consumed as a snack owing to its nutritional properties and high content of proteins and unsaturated fatty acids. Recent progress in the study and development of diversified products prepared from silkworm pupa has boosted the revenue obtained from the sericulture industry. However, the severe allergy symptoms resulting from consuming silkworm pupa have limited its applications in the food and feed industries. It has been reported that the glycoside hydrolase family 18 chitinase (BmChi) from domestic silkworm (*Bombyx mori*) pupa is an important food allergen; however, no 3-dimensional structure studies have been conducted on the allergen BmChi. In this study, several bioinformatics software tools were used to comprehensively analyze the homology, secondary structure, hydrophilicity and hydrophobicity, molecular flexibility, antigenicity, and surface accessibility of BmChi, and subsequently, the B-cell epitope region was predicted. The results of this study can provide the foundation for an in-depth study on the immunological characteristics of BmChi and help developing methods for desensitization to silkworm pupa.

Key words: domestic silkworm (*Bombyx mori*) pupa; glycoside hydrolase family 18 chitinase; food allergen; antigenic epitope

收稿日期: 2016-04-19

基金项目: 国家自然科学基金项目(31501421); 国家公益性农业行业科技专项(201403049); 广东省教育部产学研专项资金项目(2013B090600060)

作者简介: 赵祥杰(1982-), 男, 副研究员, 博士研究生, 研究方向: 食品安全、糖生物技术与安全

通讯作者: 李冰(1976-), 女, 教授, 博士生导师, 研究方向: 食品安全、糖生物技术与安全

近年来,随着人们消费方式与消费观念的巨大变化,以及世界范围内货物贸易的便捷化,人们越来越容易接触到新的和不常见的食物品种,从而使得食物过敏疾病的发病率日益增高,例如在美国 1990~2006 年间食物过敏的发生率增加了接近 4 倍;同时,在澳大利亚 1995~2006 年间食物过敏的发生率惊人的增加了约 12 倍^[1]。在食物过敏发病人群中,成人的发病率可达约为 4%,而婴幼儿阶段的发病率可达约 8%,有部分婴幼儿会在 5~6 岁阶段时对牛乳、鸡蛋、花生、小麦和大豆等致敏性食物产生一定的免疫耐受性^[2],从而严重影响过敏患者的身体健康。在已经报道的引起食物过敏反应的 170 多种常见食物中,90%以上致敏患者均由包括牛奶、鱼、蛋类、花生、大豆、小麦、坚果和水生贝壳类动物等在内的八大类易诱发过敏的食物所引起^[3]。除此之外,来自不同地域和生态环境的人们受遗传体质、饮食习惯、主要食物种类及风俗传统等因素影响,其各自对不同食物的过敏性各不相同,有的也会存在显著差异,如欧美等西方国家的人群对芹菜、坚果等食物容易过敏,而东亚和东南亚区域的人群则对牛乳和水产类食品的过敏性程度更高。

家蚕蛹是蚕桑产业中的重要副产物资源,因其营养丰富,在我国和全球各主要蚕桑产地均存在以蚕蛹为食物和饲料的应用,食用或饲用蚕蛹常导致食物过敏现象的发生。刘硕^[4]和 Zhao^[5]等均通过蛋白质组学和免疫印记实验确定了糖水解酶 18 家族几丁质酶(BmChi)为引起家蚕蛹过敏的一种变应原物质。研究表明,引发患者严重致敏反应的并非是整个变应原蛋白分子,因为食物经过动物胃肠道的消化与吸收,变应原蛋白分子发生降解或断裂,小分子的肽段是致敏反应发生的物质基础,而小分子肽段上的抗原表位(Antigenic epitopes)是起到与抗体特异结合,触发机体的一系列免疫反应的关键因素^[6],对抗原表位的鉴定将有助于理解变应原致敏机理、过敏疾病的诊断与治疗、变应原蛋白分子的定点改造脱敏,因此确定变应原的抗原表位位点在食物过敏研究中至关重要。

但由于尚未有该变应原蛋白在分子水平上的结构数据信息,对其抗原表位位点的研究仍非常缺乏,而借助于多种结构与性质预测和模拟工具能有效地获得变应原的抗原表位信息,为该变应原的定向改造提供重要信息。因此本文主要结合 DNASTar、Bioedit、ClustralW 及 BepiPred 等蛋白质分析程序,对 BmChi 的二级结构、部分理化性质、抗原性质与抗原表位等进行综合分析,预测并分析出 BmChi 的抗原表位位点,为后续相关脱敏研究提供帮助。

1 材料与方法

1.1 材料

1.1.1 家蚕蛹 BmChi 序列

利用 NCBI (National Center for Biotechnology Information) 的 Uniprot 蛋白质数据库 (<http://www.uniprot.org>) 检索 BmChi 的全部氨基酸序列(登录号: Q869E2), 得到其全部氨基酸序列。

1.1.2 不同来源几丁质酶变应原的序列

利用 Allergen(<http://www.allergen.org>) 和 Allergome(<http://www.allergome.org>) 在线变应原数据库检索已发现并命名的几丁质酶变应原,采用 Bioedit 和 ClustralW 软件对几丁质酶变应原氨基酸序列(包括 BmChi) 进行多序列比对。

1.2 试验方法

1.2.1 不同几丁质酶变应原的序列比对

将 BmChi 和不同的几丁质酶变应原序列采用 Bioedit 和 ClustralW 软件进行多序列比对。

1.2.2 BmChi 蛋白二级结构一致性预测与分析

通过利用法国 Network Protein Sequence Analysis (NPSA, <https://prabi.ibcp.fr/>) 在线系统的 GOR4 (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html)、PHD (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_phd.html) 和 SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) 等蛋白质二级结构预测工具,对 BmChi 的二级结构进行预测与分析。

1.2.3 BmChi 蛋白的亲疏水性、可塑性、抗原性以及表面可及性预测与分析

BmChi 蛋白的亲疏水性、可塑性、抗原性以及表面可及性等分别利用 DNASTar 软件中的 Kyte-Doolittle 法和 Hopp-woods 法、Karplus-Schulz 法、Jameson-Wolf 法以及 Plot-Emini 法来进行预测与分析^[7]。

1.2.4 BmChi 蛋白 B 细胞抗原表位的预测与分析

应用 BepiPred 1.0 Server (<http://www.cbs.dtu.dk/services/BepiPred/>) 在线分析工具,预测 BmChi 蛋白序列中的 B 细胞抗原表位的位点。

1.2.5 BmChi 抗原表位的综合分析

综合以上分析,选择亲水性高、可塑性好、可及性高和抗原指数大的片段作为候选表位区,比对家蚕蛹变应原几丁质酶 chitinase 的结构预测结果和其亲水性、可塑性、可及性和抗原指数,取各预测结果的重

合区域作为该变应原蛋白潜在的 B 细胞抗原表位。

2 结果与讨论

2.1 BmChi 与不同几丁质酶变应原的同源分析

我们在 Allergen 和 Allergome 变应原数据库检索并选择了14种几丁质酶变应原进行同源分析。图1是15种不同来源几丁质酶变应原蛋白的氨基酸序列多重比对的结果。从图1可以看出,15种几丁质酶变应原的氨基酸序列具有一定的相似度,其中相似度最高的是来自于巴西橡胶(Hevea brasiliensis)的 Hev b 11.0101 和 Hev b 11.0102, 经 Bioedit 软件计算得出的两序列相似度为97.2%, 但该两几丁质酶变应原与同样来自巴西橡胶(Hevea brasiliensis)的 Hev b 14的相似度仅分别为5.5%和4.8%。不同来源几丁质酶变应原中序列

相似度最高的为来源于香蕉(Musa acuminata)的 Mus a 2和来源于鳄梨(Persea americana)的 Pers a 1, 相似度达到76.0%。不同来源中序列相似度最低的为来源于台湾青枣(Ziziphus mauritiana)的 Ziz m 1和来源于美洲屋尘螨(Dermatophagoides farinae)的 Der f 15, 两者相似度仅为2.3%。与 BmChi 蛋白相似度较高的为来自欧洲屋尘螨(Dermatophagoides pteronyssinus)的 Der p 18、美洲屋尘螨(Dermatophagoides farinae)的 Der p 18、欧洲屋尘螨(Dermatophagoides pteronyssinus)的 Der p 15和美洲屋尘螨(Dermatophagoides farinae)的 Der f 15, 其与 BmChi 的相似度分别为8.8%、8.8%、7.4%和6.9%, 其余几丁质变应原与 BmChi 的相似度均低于5.0%。与 BmChi 序列相似度高的既有来自于昆虫类的几丁质酶, 也有来自植物类的几丁质酶, 说明这些变应原虽然来源不同, 但亲缘关系却较近, 根据已有报道这些相似度很高的变应原之间存在交叉反应的可能性更高[8]。

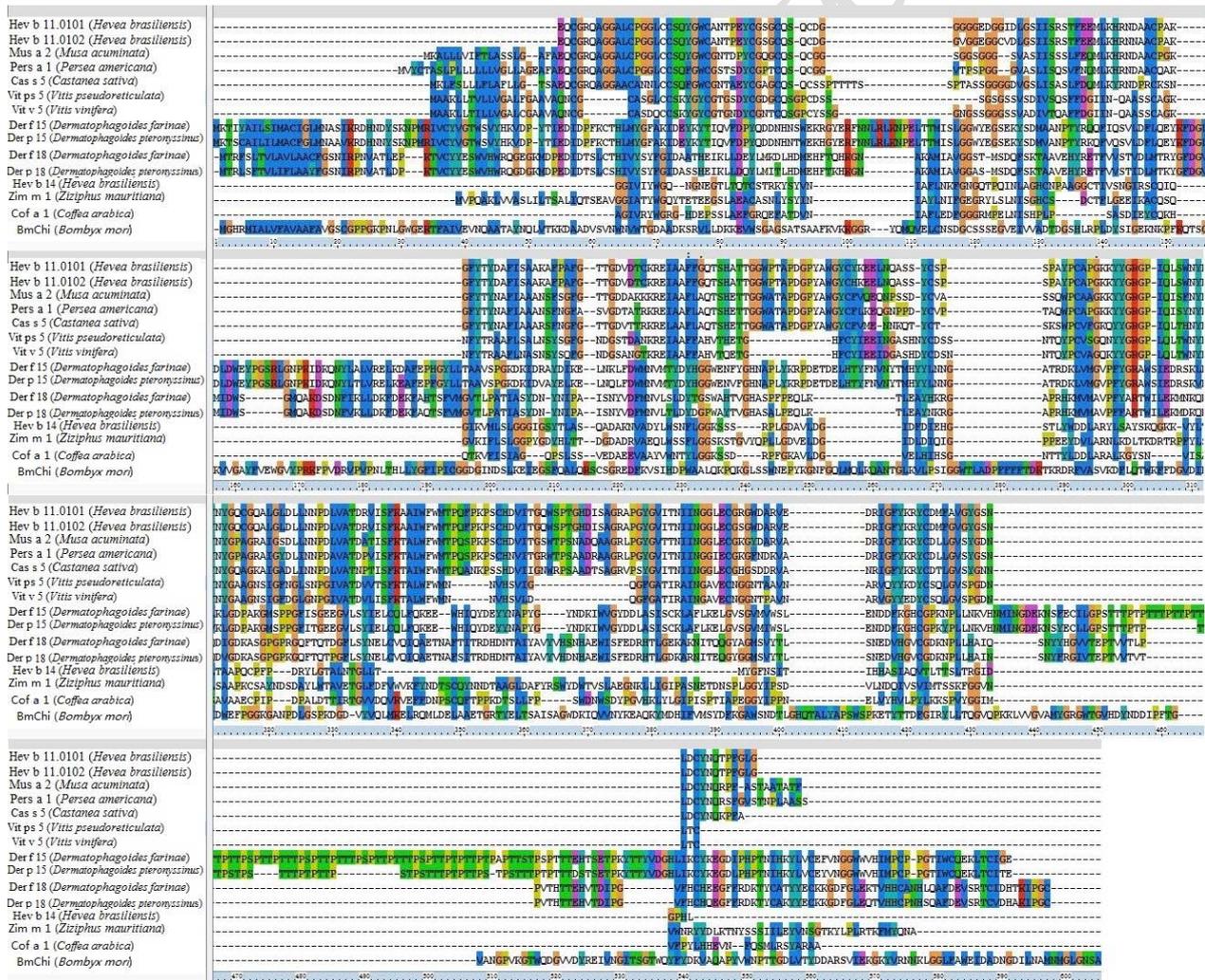


图 1 15 种不同几丁质变应原的氨基酸序列多重比对结果

Fig.1 Multiple alignments of amino acid sequences of 15 chitinase allergens obtained from different sources

2.2 BmChi 蛋白的二级结构预测

表 1 采用 GOR4、PHD 和 SOPMA 预测的家蚕几丁质酶二级结构组成情况

Table 1 Secondary structures of BmChi predicted by GOR4, PHD, and SOPMA

结构预测方法	α -螺旋/%	延伸链/%	β -转角/%	无规卷曲/%
GOR4	22.52	24.86	0.01	52.61
PHD	26.13	24.50	0.00	49.37
SOPMA	23.96	25.23	10.99	39.82

表 1 是分别采用目前蛋白质二级结构预测中准确度较高和使用较广的 GOR4、PHD 和 SOPMA 三种方法后的结果。从该表中可以看出，三种方法虽然对二级结构的预测不完全一致，但两两完全相同的比率均在 90%左右，因此总体预测的一致率是非常高的。同时三种预测方法中无规卷曲结构所占的比例均最高，优于无规卷曲因受侧链的影响较大，这类有序的非重复结构易于构成酶活性部位或功能部位，因此是抗原表位的重要形成区域。而 α -螺旋和延伸链均为比较稳

定的二级结构，不容易成为抗原表位的位点。

2.3 BmChi 蛋白的亲疏水性预测与分析

BmChi 分子亲疏水性分布区域的预测与分析主要采用 Kyte-Doolittle 方法和 Hopp-Woods 方法。由图 2 可知，两种方法得出的亲水性区域均多于疏水性区域，特别 Kyte-Doolittle 方法预测的结果亲水性区域接近 80%，而高的亲水性质说明该蛋白质亲水性氨基酸含量较高，且易于位于蛋白质分子的表面，因此也更有可能是抗原抗体特异结合的位点。由图 2 还可以看出，虽然 Kyte-Doolittle 和 Hopp-Woods 两种方法预测的亲疏水性区域并未完全重合，但两者分别在亲水性和疏水性预测区域的序列上是高度相似的，其中两种方法预测的重叠亲水区域有 21~27、50~60、70~85、112~120、128~132、139~153、168~170、193~205、193~204、210~220、232~238、282~291、315~329、336~353、370~378、394~398、408~414、470~479、510~530 和 536~542。

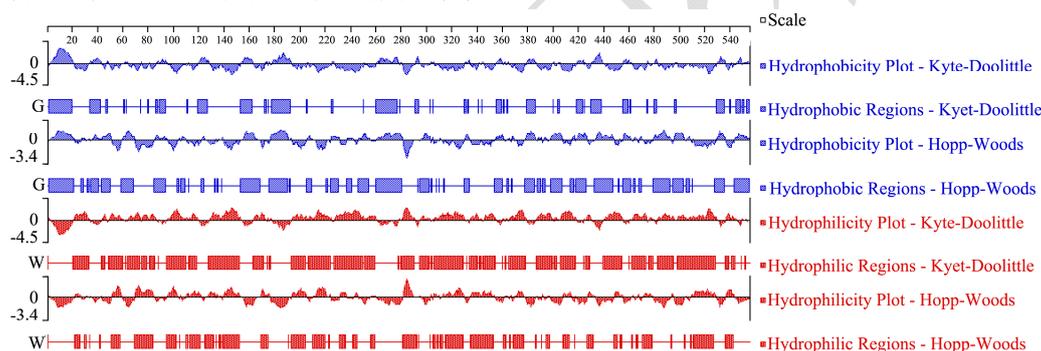


图 2 家蚕蛹几丁质酶分子亲疏水性区域分析

Fig.2 Hydrophilicity and hydrophobicity analysis of silkworm pupa chitinase with DNASTar

2.4 BmChi 蛋白的可塑性预测与分析

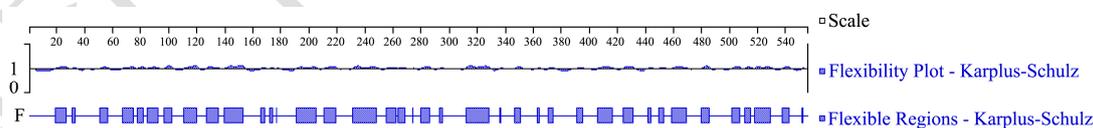


图 3 家蚕蛹几丁质酶分子可塑性区域分析

Table 3 Molecular flexibility analysis of silkworm pupa chitinase with DNASTar

BmChi 分子可塑性区域的预测与分析主要利用 DNASTar 软件的 Karplus-Schulz 法进行，序列可塑性代表了预测蛋白质多肽链骨架的柔韧性，高柔韧性的区域更可能成为抗原表位位点^[9]。经软件计算得出 BmChi 可塑性高的区域为：19~27、31~33、51~56、67~75、78~82、85~92、97~102、111~120、127~135、140~153、166~168、172~174、191~205、211~219、

231~248、255~262、264~268、280~286、293~295、312~328、336~337、347~351、363~364、371~374、391~395、406~416、424~431、442~444、450~453、459~469、480~485、502~507、511~515、518~529、538~542 和 552。

2.5 BmChi 蛋白的抗原性预测与分析

采用Jameson-Wolf方法结合现有已知蛋白质结构的抗原性来预测该蛋白抗原指数的高低,从而了解可能的抗原表位点。由图4可以看出,该蛋白多肽序列上存在有多个潜在抗原表位点,其中抗原指数较高的区域主要在氨基酸残基的第19~32、51~58、69~84、

95~103、111~120、125~152、165~175、191~204、211~220、230~248、280~293、303~329、335~354、369~377、383~397、406~413、426~432、511~530和536~543。

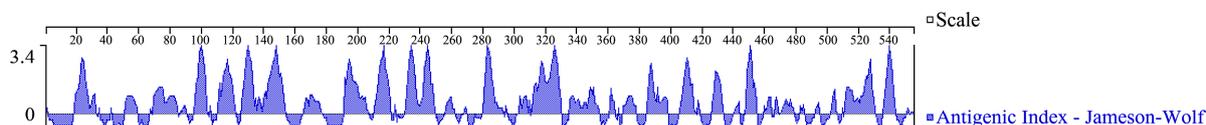


图4 家蚕蛹几丁质酶分子抗原性区域分布分析

Fig.4 Antigenicity analysis of silkworm pupa chitinase with DNASTar

2.6 BmChi 蛋白的表面可及性预测与分

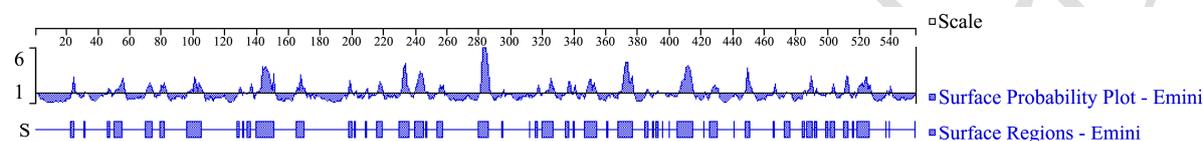


图5 家蚕蛹几丁质酶表面可及性区域分布分析

Fig.5 Surface accessibility analysis of silkworm pupa chitinase with DNASTar

BmChi 的表面可及性预测与分析主要采用 Plot-Emini 法。表面可及性是指蛋白质分子中氨基酸残基被溶剂分子接近或接触的可能性,该指数反映了蛋白质分子中各氨基酸残基在溶剂中分布情况。由图5可知,该蛋白的表面可及性区域分布较广,说明这些区域氨基酸残基很可能处于蛋白分子表面,易于接触溶剂而发生抗原抗体特异性结合。图中所示,表面可及性指数>1的区域有:23~25、31~32、46~47、50~55、70~74、79~82、96~105、128~129、13~132、134~136、140~151、165~170、198~200、216~219、230~236、240~245、254~257、280~286、320~327、335~337、347~354、368~377、385~387、405~415、426~430、448~451、473~476、484~485、487~490、492~493、499~500、502~504、510~513、519~526和537~539。

2.7 BmChi 蛋白 B 细胞抗原表位的预测与分析

BepiPred 方法是综合了隐马尔可夫模型 (hidden Markov model) 和倾向量表法 (propensity scale method) 来预测线性 B 细胞抗原表位的方法,该方法在氨基酸残基构象倾向、序列边缘效应、残基比率、二级结构反馈效应等方面均具有独到之处^[10],从而对抗原表位点预测具有较好表现。经 BepiPred 1.0 Server 预测的抗原表位区域得分见图6,该结果的阈值

为0.350,平均值为0.243,最高值为2.313,超过阈值的部分(黄色填充区域)可看做是潜在抗原表位点^[11],因此经该方法预测的区域为:18~31、42~45、54~59、66~73、83~92、100~102、112~120、126~133、139~154、167~176、192~197、199~201、213~218、228~247、270~271、283~285、308~328、346~353、359~360、372~374、389~397、402~415、427~429、443~472、482~484、494~495、497~510、512~517、522~523、537~542和554~555。

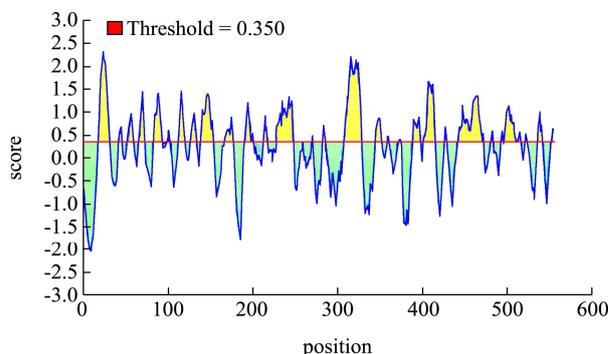


图6 采用 BepiPred 1.0 Server 方法预测的 BmChi 的 B 细胞抗原表位区域

Fig.6 Predicted B-cell epitope regions of BmChi by BepiPred 1.0 Server

2.8 BmChi 抗原表位的综合分析

将以上由来自 DNASTar 的 Protean、Bioedit、

ClustralW 和 BepiPred 等多种方法获得的 BmChi 亲水性、可塑性、表面可及性、抗原指数以及抗原表位等的预测结果进行综合,且当亲水性>0、可塑性>1、表面可及性>1以及抗原指数>0的区段形成抗原表位的可能性最高^[12],综合以上研究选出最有可能的抗原表位区域,然后与 BepiPred 的方法进行比较。从各区段的集中情况,抗原表位的位点比较分散,可能的位点也较多,但主要集中在相似度 BmChi 蛋白序列的 15~80、110~235、320~430和490~550四个比较大的分布区域,从图1中可以看出,在这些区域的变应原,特别是与 BmChi 蛋白序列相似度较高的 Der p 18、Der f 18、Der p 15和 Der f 15变应原都是有存在保持了较高的保守性的区域,由于变应原蛋白结构对其免疫原性形成起着非常重要的作用^[13],因此,这些与已知变应原序列上较为保守区域是潜在的抗原表位的位点,而且抗原表位的位点要求亲水性和表面可及性要高,这就有利于所在区段的氨基酸残基从蛋白中外露或更易与抗体发生特异性结合,而且一般这些区段内氨基酸残基长度在8~20之间,且带电荷的极性氨基酸残基出现概率也比较高,精氨酸、酪氨酸和色氨酸出现概率较高,但缬氨酸出现在表位的概率极低,其二级结构一般为较柔性的无规卷曲等居多^[14]。综合以上各方面信息最终得出的抗原表位位点有: 23~37、66~74、70~82、107~120、139~153、192~201、228~236、270~279、308~327、402~415、497~510和537~539。

3 结论

采用计算机模拟的方法进行抗原表位的预测主要依据蛋白上氨基酸序列的理化性质及其二级结构构成,通过数据库比对与分析,得出理论上能成为抗原表位的最佳区域,在很大程度上提高了工作效率。虽然限于篇幅关系没有展示免疫学反应的结果,但本文与很多单纯进行计算机模拟获取抗原表位位点的研究是有很大区别的,即还从变应原数据库中选取与 BmChi 有一定同源性的几丁质酶变应原,通过分析其保守序列和已报道抗原表位位点序列,获得准确性更高的潜在抗原表位位点,为开展该变应原的免疫反应性研究及针对性的致敏失活研究提供了重要参考。目前主要的抗原表位预测以 B 抗体细胞的线性表位为主,实际上蛋白在空间上的表位位点非常复杂,受到多种外界环境的影响,抗原表位也有可能是非线性的位点,这需要我们借助一些三维结构的模拟与对接方法来丰富和完善抗原表位位点的认识,以上从 BmChi 蛋白二级结构及相关抗原性的预测结果已经在抗原表位预测方面具有较高的准确性,为进一步开展相关研

究打下基础,同时计算机模拟的结果仍需要通过进一步实验来验证,从而为 BmChi 变应原的抗原性研究提供技术支撑。

参考文献

- [1] Schernhammer E, Vutuc C, Waldhör T. Time trends of the prevalence of asthma and allergic disease in Austrian children [J]. *Pediatric Allergy and Immunology*, 2008, 19(2): 125-131
- [2] Davis C M. Food allergies: clinical manifestations, diagnosis, and management [J]. *Current Problems in Pediatric and Adolescent Health Care*, 2009, 39(10): 236-254
- [3] Sampson H A. Update on food allergy [J]. *Journal of Allergy and Clinical Immunology*, 2004, 113(5): 805-819
- [4] 刘硕.家蚕蚕蛹过敏原蛋白质组学和分子生物学研究[D].南昌大学,2010
LIU Shuo. Proteomic and molecular biology analysis of silkworm (*Bombyx mori* L.) pupa allergens [D]. Nanchang University, 2010
- [5] Zhao X J, Li L, Kuang Z S, et al. Proteomic and immunological identification of two new allergens from silkworm (*Bombyx mori* L.) pupae [J]. *Central European Journal of Immunology*, 2015, 40(1): 30-34
- [6] 张海雷,白换力,薛慧亮,等.犬瘟热病毒 N 蛋白的 B 细胞抗原表位预测[J].*中国兽医学报*,2014,34(1):10-16
ZHANG Hai-lei, BAI Huan-li, XUE Hui-liang, et al. Prediction of antigenic epitopes of canine distemper virus N protein [J]. *Chinese Journal of Veterinary Science*, 2014, 34(1): 10-16
- [7] Sun X L, Shan X H, Yan Z H, et al. Prediction and characterization of the linear IgE epitopes for the major soybean allergen β -conglycinin using immunoinformatics tools [J]. *Food and Chemical Toxicology*, 2013, 56: 254-260
- [8] Lee M F, Chen Y H, Lan J L, et al. IgE cross-reactivity between the major peanut allergen Ara h 2 and the nonhomologous allergens Ara h 1 and Ara h 3 [J]. *Journal of Allergy and Clinical Immunology*, 2013, 132(1): 118-124
- [9] Nyárady Z, Czömpöly T, Bősze Sz, et al. Validation of in silico prediction by *in vitro* immunoserological results of fine epitope mapping on citrate synthase specific autoantibodies [J]. *Molecular Immunology*, 2006, 43(7): 830-838
- [10] Zheng L N, Lin H, Pawar R, et al. Mapping IgE binding epitopes of major shrimp (*Penaeus monodon*) allergen with immunoinformatics tools [J]. *Food and Chemical Toxicology*, 2011, 49: 2954-2960
- [11] Song J, He Q F. Bioinformatics analysis of the structure and

- linear B-cell epitopes of aquaporin-3 from *Schistosoma japonicum* [J]. *Asian Pacific Journal of Tropical Medicine*, 2012, 5(2): 107-109
- [12] Zhang Z W, Zhang Y G, Wang Y L, et al. Screening and identification of B cell epitopes of structural proteins of foot-and-mouth disease virus serotype Asia1 [J]. *Veterinary Microbiology*, 2010, 140(1-2): 25-33
- [13] Mueller G A, Edwards L L, Aloor J J, et al. The structure of the dust mite allergen Der p 7 reveals similarities to innate immune proteins [J]. *Journal of Allergy and Clinical Immunology*, 2010, 125(4): 909-917
- [14] 朱倩倩,吴序栋,肖杰,等.腰果过敏原 Ana o 2 结构及抗原表位预测[J].*食品与生物技术学报*,2012,32(2):141-145
- ZHU Qian-qian, WU Xu-li, XIAO Jie, et al. Prediction of the structure and antigenic epitopes of cashew allergen Ana o 2 [J]. *Journal of Food Science and Biotechnology*, 2012, 32(2): 141-145