

# 细粒棘球绦虫 AgB1 抗原表位的生物信息学预测

马秀敏<sup>1,2</sup>, 胡晓安<sup>1,2</sup>, 阿尔孜古丽·吐尔逊<sup>1</sup>, 李艳华<sup>2</sup>, 刘学磊<sup>2</sup>, 朱明<sup>2</sup>, 丁剑冰<sup>2</sup>

1. 新疆医科大学第一附属医院, 省部共建国家重点实验室培育基地——新疆重大疾病医学重点实验室, 乌鲁木齐 830054
2. 新疆医科大学基础医学院, 乌鲁木齐 830011

**摘要** 以细粒棘球绦虫 AgB1 基因序列为基础, 采用 PredictProtein 软件预测其编码蛋白的二级结构; 应用在线预测软件 Bcepred、Abcpred、IEDB 及 SYFPEITHI 等对细粒棘球绦虫 AgB1 的 B 细胞表位和 T 细胞表位进行预测。结果提示, AgB1 抗原蛋白存在可以构成抗原表位的区域, 经软件分析, 分值高的 B 细胞表位区域: 2~9、15~20、22~35 和 41~52 氨基酸序列; T 细胞表位区域: 3~12、26~33、34~44 和 52~61 氨基酸序列。研究运用生物信息学确定 AgB1 抗原的 4 个 B 细胞优势表位和 4 个 T 细胞优势表位, 对进一步研究 AgB1 的抗原性和研发更有价值的免疫诊断方法具有重要意义。

**关键词** AgB1; B 细胞表位; T 细胞表位; 生物信息学

**中图分类号** R392

**文献标志码** A

**doi** 10.3981/j.issn.1000-7857.2013.27.003

## Bioinformatics Prediction of AgB1 Antigen Epitopes of *Echinococcus granulosus*

MA Xiumin<sup>1,2</sup>, HU Xiaolan<sup>1,2</sup>, TUERXUN Arzgul<sup>1</sup>, LI Yanhua<sup>2</sup>, LIU Xuelei<sup>2</sup>, ZHU Ming<sup>2</sup>, DING Jianbing<sup>2</sup>

1. State Key Laboratory Incubation Base of Xinjiang Major Diseases Research, the First Affiliated Hospital of Xinjiang Medical University, Urumqi 830054, China
2. Department of Basic Medicine, Xinjiang Medical University, Urumqi 830011, China

**Abstract** The secondary structure, B cell and T cell epitopes for the AgB1 antigen protein of *Echinococcus granulosus* were predicted and analyzed. According to the gene sequences of AgB1 antigen, the protein secondary structure was analyzed by PredictProtein, and B cell and T cell epitopes of AgB1 were predicted by Internet website Bcepred, Abcpred, IEDB and SYFPEITHI. The results show that many distinct antigenic epitopes of AgB1 were identified by computation, and the high scored B cell epitopes in the regions of 2-9, 15-20, 22-35 and 41-52. The possible T cell epitopes is in the regions of 3-12, 26-33, 34-44 and 52-61. In this study it is determined that there are 4 B cell dominant epitopes and 4 T cell ones in AgB1 antigen by bioinformatics methods, which are significant in the further research of AgB1 antigenic and development of the immune diagnosis method.

**Keywords** AgB1; B cell epitopes; T cell epitopes; bioinformatics

### 0 引言

细粒棘球蚴病 (echinococcosis) 又称囊型包虫病 (Cystic Echinococcosis, C), 是由细粒棘球绦虫 (*Echinococcus granulosus*, Eg) 的幼虫寄生于人或动物体内引起的一类具有地方流行性的人兽共患慢性寄生虫病, 呈全球性分布。中国

是细粒棘球蚴病发病最高的国家之一, 主要在西部的新疆、青海、甘肃、宁夏等牧区和半牧区流行<sup>[1]</sup>。此病患者大多早期无明显症状, 往往使患者错过最佳治疗时期。因此, 寻找敏感性高、特异性强的抗原成为细粒棘球蚴病免疫学诊断的重要任务。

收稿日期: 2012-12-21; 修回日期: 2013-08-20

基金项目: 新疆维吾尔自治区自然科学基金项目 (2012211A034); 国家自然科学基金项目 (31160194, 81260253, 30960358)

作者简介: 马秀敏, 教授, 研究方向为感染免疫, 电子信箱: maxiumin1210@163.com; 丁剑冰 (通信作者), 教授, 研究方向为感染免疫, 电子信箱: djbing002@163.com

表位(epitope)是抗原分子中决定抗原特异性的特殊化学基团,又称抗原决定簇(antigenic determinant),它是T细胞抗原受体(T Cell Receptor,TCR)和B细胞抗原受体(B Cell Receptor,BCR)或抗体特异性结合的基本结构单位。其中,B细胞表位位于抗原分子表面,能诱导B细胞产生特异性抗体,从而诱导机体产生体液免疫应答。而T细胞表位是指在特异性免疫应答中抗原经APC处理后由MHC分子向TCR递呈的线性肽段,与机体的细胞免疫有关。这些细胞表位可诱导体液免疫或细胞免疫,对表位筛选意义重大。本研究综合运用多种生物信息学方法,对AgB1亚单位基因的二级结构以及B细胞、T细胞抗原表位进行分析预测,以为构建多表位的重组抗原提供依据,为包虫病诊断和表位疫苗研制奠定基础。

### 1 材料与方法

#### 1.1 AgB1 的氨基酸序列

从GeneBank提供的AgB1基因序列中,选取进行分析的基因序列,用DNAMAN软件推导出相应抗原氨基酸序列。

#### 1.2 蛋白质二级结构预测软件

使用在线预测网站PredictProtein(<https://www.predictprotein.org>)预测蛋白质二级结构。

#### 1.3 B细胞表位的预测

使用在线预测软件BCEpred(<http://www.imtech.res.in/>

[http://www.imtech.res.in/raghava/bcepred/bcepred\\_submission.html](http://www.imtech.res.in/raghava/bcepred/bcepred_submission.html)),其中的亲水性参数、柔韧性参数、抗原性指数等,以及在线网络软件ABCpred(<http://www.imtech.res.in/raghava/abcpred/>)对细粒棘球蚴AgB1的B细胞表位进行综合预测。

#### 1.4 T细胞表位的预测

应用在线预测软件IEDB([http://tools.immuneepitope.org/main/html/tcell\\_tools.html](http://tools.immuneepitope.org/main/html/tcell_tools.html))提供的模块预测MHC I类HLA-A0201限制性T细胞表位和SYFPEITHI(<http://www.syfpeithi.de/>)预测MHC I类HLA A0201限制性T细胞表位。

## 2 结果

### 2.1 AgB1 基因编码的氨基酸序列

根据GeneBank的AgB1基因序列,经过多重序列比对,确定最终进行分析的AgB1基因序列(Genebank登录号:HM357141),用DNAMAN软件推导出其翻译的相应氨基酸序列,共编码65个氨基酸残基:ddgltstsrsvmkmfgevkyfferdplgk kvvdllleeleevefqlrkkklrmlrshlrgliaege

### 2.2 蛋白质二级结构预测结果

使用在线预测网站PredictProtein预测AgB1抗原蛋白的二级结构,结果显示,其中的螺旋结构helix占氨基酸总数的75.38%,环状结构loop占24.62%。各种结构在AgB1抗原蛋白中的分布情况见图1。

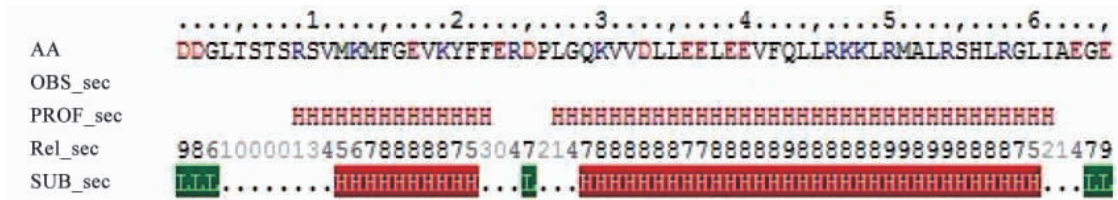


图1 PredictProtein分析的AgB1二级结构  
Fig. 1 Secondary structure of AgB1 using PredictProtein analysis

### 2.3 B细胞表位预测结果

使用在线网络软件BCEpred([http://www.imtech.res.in/raghava/bcepred/bcepred\\_submission.html](http://www.imtech.res.in/raghava/bcepred/bcepred_submission.html))分析AgB1的亲水性参数、柔韧性参数、抗原性指数、表面暴露区域和表面可及性参数,根据分析结果,确定出各参数分值较高的氨基酸序列,结果如下。

亲水性分析结果:1~7氨基酸序列(图2);柔韧性分析结果:1~9氨基酸序列(图3);抗原性分析结果:26~47氨基酸序列(图4);表面暴露区域分析结果:43~52氨基酸序列(图5);表面可及性分析结果:19~30氨基酸序列和41~52氨基酸序列(图6)。

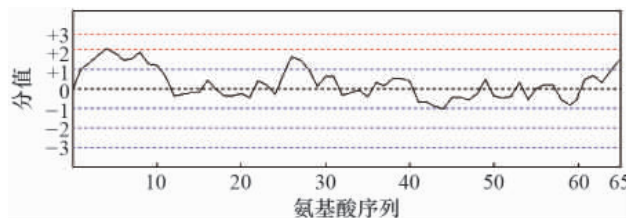


图2 亲水性参数预测结果  
Fig. 2 Prediction of hydrophilicity

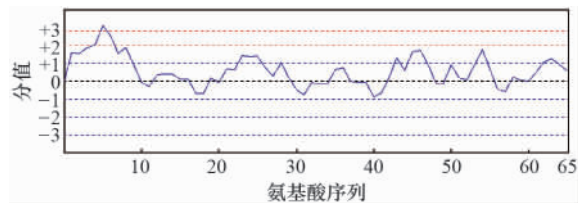


图3 柔韧性区域预测结果  
Fig. 3 Prediction of flexibility

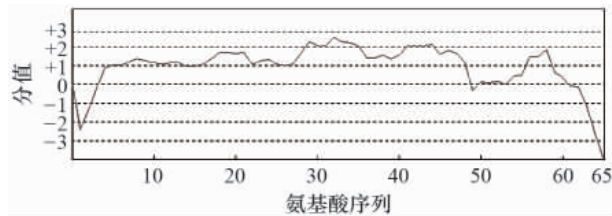


图4 抗原性预测结果  
Fig. 4 Prediction of antigenic propensity

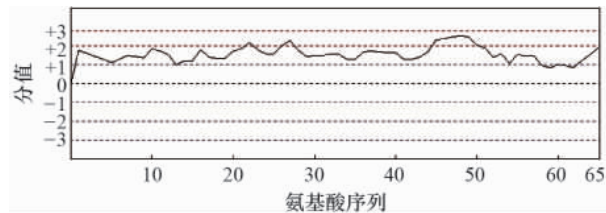


图5 表面暴露区域预测结果  
Fig. 5 Prediction of exposed surface

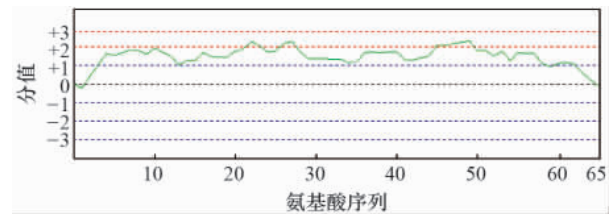


图6 表面可及性参数预测结果  
Fig. 6 Prediction of accessibility

使用 ABCpred 预测服务器对 AgB1 的 B 细胞表位进行预测,得到可能构成表位的分值较高的氨基酸序列为 15~30、2~17、28~43 和 49~64(表 1)。

表 1 ABCpred 预测 B 细胞表位预测结果  
Table 1 Prediction of B epitope using ABCpred

氨基酸序列	开始位点	分值
FGEVKYFFERDPLGQK	15	0.83
DGLTSTSRVSMKMFGE	2	0.74
GQKVVDLLEELVEVFQ	28	0.62
LRMALRSHIRGLIAEG	49	0.55

综合两种预测 B 细胞表位的方法,选出 4 个潜在的优势表位区域,分别是 2~9、15~20、22~35 和 41~52 氨基酸序列。

#### 2.4 T 细胞表位预测结果

分别选取 10 个由在线预测软件 IEDB 和 SYFPEITHI 预测的分值较高的 AgB1 蛋白 MHC I 类 HLA-A 0201 限制性 T 细胞表位,综合分析、对比展示见表 2。

表 2 IEDB 与 SYFPEITHI 预测的 MHC I 类限制性 T 细胞表位  
Table 2 MHC I nonamers T cell epitope using IEDB and SYFPEITHI

序号	IEDB 在线软件			SYFPEITHI 软件		
	起始位点	氨基酸构成	分值	起始位点	氨基酸构成	分值
1	35	LEELVEVFQ	93.10	33	DLLEELVEV	29
2	25	DPLGQKVVD	92.10	52	AI RSHLRGL	27
3	1	DDGLTSTR	91.10	26	PLGQKVVDL	24
4	39	EEVFQLLRK	88.40	30	KVVDLLEEL	23
5	8	SRSVMKMFGE	87.50	3	GLTSTSRV	22
6	45	LRKKLRMAL	86.80	44	LIRKKLRMA	21
7	46	RKKLRMALR	85.00	56	HIRGLIAEG	21
8	17	EVKYFFERD	85.00	10	SVMKMFGEV	19
9	38	LEEVFQLLR	83.00	37	ELEEVFQLL	18
10	22	FERDPLGQK	82.70	43	QLRKKLRM	18

综合两个在线软件的预测结果,两种方法得分都比较高的区域作为可能的 T 细胞表位,分别是 3~12、26~33、34~44、52~61 氨基酸序列。

### 3 讨论

中国是包虫病高发国家之一,该病严重影响中国西部农、牧民身体健康,并给中国西部经济发展带来沉重的负担。

目前对于包虫病的治疗主要以手术治疗为主,药物治疗为辅,但是效果欠佳。因此有必要通过免疫防治途径来控制 and 清除包虫病。

抗原 B 被公认为是细粒棘球绦虫中具有高度免疫原性和较好特异性的囊液抗原<sup>[2]</sup>,是囊液中含有最丰富且稳定,并与诊断相关的一个主要抗原成份。目前已经发现了 5 个抗原 B 亚单位,包括 EgAgB8/1<sup>[3]</sup>,EgAgB8/2<sup>[4]</sup>,EgAgB8/3<sup>[5]</sup>,EgAgB8/4<sup>[6]</sup>和 EgAgB8/5<sup>[7-8]</sup>。Gonzalez 等<sup>[9]</sup>发现 AgB8/1 分子 N 末端的 38 个氨基酸长度合成肽 P176,诊断 CE 患者的敏感性和特异性分别为 80%和 94%。张海涛等<sup>[10]</sup>发现 EgAgB8/1 在棘球绦虫生发层大量表达,表达量为 68.874,而人体感染细粒棘球绦虫的主要阶段为中绦期,此期包含生发层和原头蚴,因此通过利用 EgAgB8/1 重组蛋白检测血清中特异性抗体可达到早期诊断的目的,以便早期治疗。

生物信息学在预测蛋白质结构和功能,以及确定蛋白质生物学特性等方面具有重要作用,已广泛应用于蛋白质抗原表位的分析预测<sup>[11]</sup>。随着生物信息学的迅速发展,生物学家设计了许多分子生物学软件,用于蛋白质的分析,目前已被认可的并具有较好预测效果的方法主要有二级结构预测、亲水性方案、抗原性方案、可塑性方案、可及性方案和电荷分布方案;上述方案中单参数预测的正确率不高,通常需将多种方案综合考虑,其中亲水性、表面可及性、可塑性、二级结构和抗原性预测较为重要;基于人工神经网络算法的 ABCpred 方案,预测准确率最高可达 65.9%<sup>[12]</sup>。为提高预测结果的准确性,本文综合这些方法对细粒棘球绦虫 AgB1 蛋白潜在的 B 细胞抗原表位进行预测和分析,最终得到 4 个优势 B 细胞表位。

T 细胞表位预测软件 SYFPEITHI 是由 Rammensee 等开发的,根据 MHC 配体和结合肽模板方案设计的算法,其中包括了人、小鼠及大鼠的 MHC I 和 II 类分子表位的预测<sup>[13]</sup>。Immune Epitope Database (IEDB)也是目前国际公认并且被广泛应用的一种 T 细胞表位预测分析软件。为了避免单一软件预测的局限性,本研究同时运用了上述两种 T 细胞表位预测软件,找出共同的 MHC I 类 HLA-A 0201 限制性 T 细胞表位的氨基酸序列区域,这无疑提高了预测结果的准确性。

#### 4 结论

细粒棘球绦虫 AgB1 蛋白具有形成 T、B 细胞表位的潜能。经软件 Bcepred 和 Abepred 分析,得到可能存在 B 细胞表位的氨基酸序列为:2~9、15~20、22~35 和 41~52;经软件 IEDB 和 SYFPEITHI 分析,得到可能存在 T 细胞表位的氨基酸序列为:3~12、26~33、34~44 和 52~61;该结果为确定和筛选优势表位,研制更加安全、高效的诊断方法和表位疫苗提供了参

考资料,为更好地防治囊型包虫病提供了依据。

#### 参考文献 (References)

- [1] 刘春燕,马秀敏,丁剑冰,等.我国棘球绦虫感染的不同宿主状况[J].中国人兽共患病学报,2009,25(6):586-588.  
Liu Chunyan, Ma Xiumin, Ding Jianbing, et al. Chinese Journal of Zoonoses, 2009, 25(6): 586-588.
- [2] Williams J F, Pérez Esandi M V, Oriol R. Evaluation of purified lipoprotein antigens of *Echinococcus granulosus* in the immunodiagnosis of human infection[J]. The American Society of Tropical Medicine and Hygiene, 1971, 20(4): 575-579.
- [3] Frosch P, Hartmann M, Mühlischlegel F, et al. Sequence heterogeneity of the echinococcal antigen B[J]. Molecular and Biochemical Parasitology, 1994, 64(1): 171-175.
- [4] Fernández V, Ferreira H B, Fernández C, et al. Molecular characterization of a novel 8kDa subunit of *Echinococcus granulosus* antigen B[J]. Molecular and Biochemical Parasitology, 1996, 77(2): 247-250.
- [5] Chemale G, Haag K L, Ferreira H B, et al. *Echinococcus granulosus* antigen B is encoded by a gene family[J]. Molecular and Biochemical Parasitology, 2001, 116(2): 233-237.
- [6] Arend A C, Zaha A, Ayala F J, et al. The *Echinococcus granulosus* antigen B shows a high degree of genetic variability [J]. Experimental Parasitology, 2004, 18(1-2): 76-80.
- [7] Haag K L, Alves-Junior L, Zaha A, et al. Contingent, non-neutral evolution in a multicellular parasite: Natural selection and gene conversion in the *Echinococcus granulosus* antigen B gene family[J]. Gene, 2004, 26(333): 157-167.
- [8] Mamuti W, Sako Y, Bat J M, et al. Molecular characterization of a novel gene encoding an 8-kDa-subunit of antigen B from *Echinococcus granulosus* genotypes I and 6[J]. Parasitology International, 2007, 56(4): 313-316.
- [9] Gonzalez S G, Lorenzo C, Nieto A. Improved immunodiagnosis of cystic hydatid disease by using a synthetic peptide with higher diagnostic value than that of its parent protein *Echinococcus granulosus* antigen B[J]. Journal of Clinical Microbiology, 2000, 38(11): 3979-3983.
- [10] 张海涛,马秀敏,吾拉木·马木提,等.细粒棘球绦虫抗原 EgAgB8/1 和 EgAgB8/3 亚单位基因在虫体发育过程中的阶段性表达[J].中国病原生物学杂志,2009,4(5): 355-358.  
Zhang Haitao, Ma Xiumin, Mamuti Wulamu, et al. Journal of Pathogen Biology, 2009, 4(5): 355-358.
- [11] Korber B, LaBute M, Yusim K. Immunoinformatics comes of age[J]. PLOS Computational Biology, 2006, 2(6): 484-492.
- [12] Saha S, Raghava G P. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network[J]. Proteins, 2006, 65(1): 40-48.
- [13] Rammensee H, Bachmann J, Emmeifch N P, et al. SYFPEITHI: Database for MHC ligands and peptide motifs [J]. Immunogenetics, 1999, 50(3-4): 213-219.

(责任编辑 吴晓丽)