Compare the Difference of B-cell Epitopes of EgAgB1 and EgAgB3 Proteins Selected through Bioinformatic Analysis

Mengting An, Fengbo Zhang, Yuejie Zhu, Xiao Zhao and Jianbing Ding*
XinJiang Medical University, 830011 XinJiang, Urumqi, China
Email:1601379937@qq.com

Abstract. Cystic echinococcosis, as a zoonosis, seriously endangers humans and animals, so early diagnosis of this disease is particularly important. Therefore, this study is to predict B-cell epitopes of EgAgB1 and EgAgB3 proteins by bioinformatic software. B-cell epitopes of EgAgB1 and EgAgB3 proteins are predicted using DNATstar and IEDB software. The results suggest that there are two potential B-cell epitopes in EgAgB1, which located in the 8-15 and 31-37 amino acid residue segments. And two potential B-cell epitopes in EgAgB2, located in the 20-27 and 47-53 amino acid residue segments. This study predicted the B-cell epitopes of EgAgB1 and EgAgB3 proteins, which laid the foundation for the early diagnosis of Cystic echinococcosis.

1. Introduction

Cystic Echinococcosis (CE) is also known as cystic hydatid disease, a class of local epidemic zoonotic chronic parasitic disease in human or animal caused by Echinococcus granulosus (Eg) Larvae infection. The adults of Eg mainly parasitic on canine carnivorous, Larvae of Eg mainly parasitic in humans and a variety of herbivorous livestock and other animals [1]. CE is a global distribution; China is one of the countries with the highest incidence of CE, Mainly prevalent in the west of Xinjiang, Qinghai, Gansu, Ningxia and other pastoral or semi-pastoral areas. CE has a long incubation period, Patients often no symptoms, most patients have been found to be late stages of chronic infection [4], so that patients miss the best treatment period. It is urgent to find specific and sensitive candidate antigen molecules for the diagnosis and prevention of CE.

EgAgB is the main antigen rich in its cyst fluid. It has high immunogenicity and good specificity. Therefore, it is the focus of research at home and abroad. Studies have shown that EgAgB is encoded by a polymorphic multigene family, consisting of 5 subunits of EgAgB1 - EgAgB5 [2], the serum reactivity of the five subunits of the EgAgB antigen was analyzed and it was shown that EgAgB1 is the major reactive subunit for the identification of serum antibodies [3]. EgAgB3 was expressed in all stages of the Eg, and the expression level increased gradually in the developmental stage [4]. And the expression of EgAgB3 in the adult stage of Eg was the most significant [5]. But, a single source or single component antigen is used to diagnose the disease; the sensitivity and specificity are unstable and even larger. Therefore, if the specific epitope of each subunit antigen is combined in an antigenic component,

It is possible to make up for the shortcomings that the difference in the expression of subunits caused by natural antigen instability, also can improve the diagnostic efficiency of EgAgB. So, this study aims to provide further theoretical support for the immunological diagnosis of CE by analyzing and comparing the dominant epitopes of EgAgB1 and EgAgB3.
2. Materials Methods
Released from GenBank of AgB1 protein of Eg. (GenBankno. AAX36085.1; http://www.ncbi.nih.gov/genbank/). The EgAgB1 protein is composed of 72 amino acid residues, the amino acid sequence is as the following: FVVVTQADDGLTSTSRSMKMKFGEVKYFFERDPLGQKVVDLKECLEEVFQLLRRKLKRMLAR SHLGLVAEGK

Released from GenBank of AgB3 protein of Eg. (GenBankno. ACO24476.1; http://www.ncbi.nih.gov/genbank/). The EgAgB3 protein is composed of 79 amino acid residues, the amino acid sequence is as the following: MKFCMLLALALVSFVVVARADDDEVTKTKKGV-MKAISEIKHFFQSDPLGKKLVEVMKDVASCEMVRR-KKARMALKEY

2.1. Prediction of the Secondary Structure of the EgAgB1 and EgAgB3 Proteins
Using Chou-Fasman and Garnier-Robson methods of DNAStar Protean software analyze the areas of secondary structure elements such as Alpha helix, beta fold, beta-turn and random coil.

2.2. Prediction of the B-cell Epitopes for the EgAgB1 and EgAgB3 Antigen
The software which predict B-cell epitopes are DNAStar (V5.0) (http://www.dnastar.com) and the online prediction software IEDB (http://tools.immuneepitope.org/bcell).

By using DNAStar software, Emini method predict the EgAgB1 and EgAgB3 proteins molecular surface accessibility, according to the specific composition of amino acid residues; Kyte-Doolittle[6] method based on the EgAgB1,EgAgB3 amino acid molecules composition, predicting their hydrophobic region and hydrophilic region; Karplus-Schulz [7] method can predict the EgAgB1, EgAgB3 polypeptide skeleton flexibility; Jameson-Wolf method[8] can finish the prediction of antigenic determinant of EgAgB1, EgAgB3 protein.

IEDB software is used for prediction and analysis of EgAgB1 and EgAgB3 proteins, B-cell linear epitopes, beta corner, and surface accessibility, flexibility of skeleton, antigen index, and hydrophobicity.

EgAgB1 and EgAgB3 hydrophilicity, Surface Accessibility, antigenicity, flexibility, linear epitope and other parameters were analyzed by DNA Star and IEDB software. Comprehensive analysis of the results, predict the B-cell epitopes.

3. Results
3.1. Prediction of the Secondary Structure of EgAgB1 and EgAgB3 Proteins
Using DNAStar software to predicted alpha helices of EgAgB1 protein, we found that alpha helices are mainly distributed in 1~2, 4~7, 13~32 and 35~72. The beta fold are located at position 1~7. In alpha helices and beta fold regions, the distribution of different three turn regions and one coil region, turn regions are located at positions: 8~12, 13~16 and 31~35. coil regions are mainly distributed in:9~13.(Figure1A)

Using DNAStar software to predicted alpha helices of EgAgB3 protein, we found that alpha helices are mainly distributed in 1~47 and 51~79. The beta fold are located at position 8~20. In alpha helices and beta fold regions, the distribution of different three turn regions and one coil region, turn regions are located at positions: 20~24, 29~33 and 47~51. coil regions are mainly distributed in:48~49.(Figure1B)

3.2. B-cell epitopes prediction of EgAgB1 and EgAgB3 proteins
DNAStar (Figure2A and 2B) and IEDB software (Figure3 and Figure 4) are used for comprehensive prediction and analysis of EgAgB1 and EgAgB3 protein and dominant B-cell epitopes. The contents of the analysis include hydrophilicity, SurfaceAccessibility, antigenicity, flexibility and other parameters. IEDB software is based on its various aspects of the score to filter out the higher score of 20 amino acid region (Table 1 and Table 2).

Comprehensive analysis of the results, predict the B-cell epitopes of EgAgB1 are located at positions: 8~15 and 31~37. Which Sequences are DDGLTSTS and RDPLGQK. Predict the B-cell
epitopes of EgAgB3 are located at positions: 20–27 and 47–53. Which Sequences are ADDDDDEV and SDPLGKK.

Figure 1. DNAStar Server software to predict the secondary structure of EgAgB1 protein(A),DNAStar Server software to predict the secondary structure of EgAgB3 protein.(B)

Figure 2. DNAStar Server software to predict B-cell epitope of the EgAgB1 protein(A),DNAStar Server software to predict B-cell epitope of the EgAgB3 protein(B)
Figure 3. IEDB online software was used to predict the B-cell epitopes of EgAgB1. hydrophilicity (A), flexibility (B), Antigencity (C), surface accessibility (D), Beta-turn (E)
Figure 4. IEDB online software was used to predict the B-cell epitopes of EgAgB4. Hydrophilicity (A), flexibility (B), Antigenicity (C), surface accessibility (D), Beta-turn (E).

Table 1. DNASTAR and IEDB software were used to calculate the amino acid region of EgAgB1 with higher parameter scores.

| EgAgB1   | Predict B-cell epitopes software |
|----------|---------------------------------|
|          | DNASTAR                         |
|          | IEDB                            |
| Hydrophilicity | 9~13,15~16, 4~10,8~14,12~18,7~13, |
|              | 26~35,53~55, 9~15,5~11,6~12,30~36, |
|              | 57~59,61~62, 3~9,11~17,31~37,10~16, |
|              | 13~19,14~20,32~38,66~72, 65~71,2~8,29~35,15~21 |
| Flexibility | 6~17,30~37, 11~17,12~18,10~16,9~15, |
|             | 42~47,52~55, 32~38,13~19,33~39,31~37,5~11,8~14,3 |
|             | 0~36,4~10, 7~13,6~12,29~35,34~40, |
|             | 51~57,3~9,41~47,14~20 |
| Antigenicity | 6~18,23~24, 1~7,36~42,2~8,37~43, |
|              | 26~39,40~47, 38~44,33~39,63~69,48~54,39~45,62~6 |
|              | 52~61,62~64,67~72, 8,47~53,34~40,35~41,45~51,46~52,64~ |
|              | 70,50~56,3~9,59~65,32~38 |
| Surface     | 7~8,13~14, 53~58,52~57,12~17,50~55,26~31,32~3 |
| Accessibility | 27~34, 7,30~35,31~36,42~47,5~10,27~32,28~3 |
|             | 44~46,51~52,54~57, 3, |
|             | 54~59,43~48,29~34,11~16,61~66,4~9,1 |
|             | 5~20,13~18 |
| Beta-Turn   | 8~12,13~16, 8~14,9~15,7~13,31~37, |
|             | 31~35, 10~16,30~36,11~17,5~11, |
|             | 6~12,12~18,32~38,4~10, |
|             | 29~35,13~19,60~66,61~67,27~33,14~2 |
|             | 0,33~39,66~72 |
Table 2. DNAstar and IEDB software were used to calculate the amino acid region of EgAgB3 with higher parameter scores.

| EgAgB3   | predict B-cell epitopes software |
|----------|----------------------------------|
| DNASTAR  | IEDB                             |
| Hydrophilicity | 20–34,42–52,6  
8–76 | 20–26,19–25,21–27,22–28,18–24,23–29,24–30,25–31,17  
5,30–36,69–75,59–65 |
| Flexibility | 19–33,39–41,4  
5–53,68–71 | 27–33,28–34,26–32,46–52,45–51,25–31,47–53,29–35,44  
50,48–54,20–26,49–55,19–25,21–27,24–30,23–29,22–2  
8,18–24,67–73,68–74 |
| Antigenicity | 17–35,37–41,4  
5–63,65–79 | 9–15,7–13,14–20,13–19, 3–9,4–10,8–14,59–65, 15–21,60–66,5–11,61–67, 62–68,2–8,63–69,49–55 |
| Surface   | 19–31,49–50,6  
7–73,77–79 | 21–26,28–33,19–24,69–74,20–25,27–32,68–73,26–31,24  
29–22,27,66–71,18–23,25–30,23–28,29–34,48–53,70–7  
5,31–36,67–72,47–52 |
| Accessibility | 20–24,29–33,4  
7–51 | 19–25,20–26,47–53,46–52,21–27,45–51,18–24,22–28,48  
54,23–29,43–49,44–50,17–23,24–30,42–48,27–33,28–3  
4,49–55,59–65,29–35 |

4. Discussion

China is one of the countries with high incidence of CE. The disease seriously affects the health of farmers and herdsmen in Western China, and brings heavy burden to the economic development in the west of China. CE is usually insidious and can be asymptomatic within a few years. Most patients have been found to be late stages of chronic infection, so that patients miss the best treatment period. Human infection with CE is often through exposure to canine faeces, while canines as the final host of Eg have not yet been effectively diagnosed. Therefore, it is necessary to early diagnosis and control of CE through immunization. EgAgB is recognized as a highly immunogenic and highly specific cystic antigen in Eg. A study found that EgAgB1 was abundantly expressed in the germinal layer of echinococcus[9]. The main stage of human infection of Eg for larva, this period contains the germinal layer and protoscolex, Therefore, the detection of specific antibodies in serum by EgAgB1 recombinant protein can achieve the purpose of early diagnosis, so that early treatment can be achieved. By real-time quantitative PCR, it was found that EgAgB3 was abundantly expressed in the adult stage of Eg[10]. Moreover, EgAgB is a secreted antigen of worms, which can be detected easily with the feces of infected dogs. Therefore, it can be used as the best target protein for detection of Eg infection in dog feces antigen.

Using bioinformatics techniques to analyze and compare protein-related information and predict their epitopes has become one of the important means of immunological research. In this study, DNAstar software and IEDB online software were used to analyze and predict the secondary structure of EgAgB1 and EgAgB3 proteins and the characteristics of B-cell epitopes. The higher the hydrophilicity, the greater the probability of exposure to the surface, and the greater the possibility of becoming an epitope. The greater the surface accessibility, the greater the probability of amino acid residues in the protein antigen being contacted by solvent molecules. The higher the flexibility of the protein skeleton, the greater the chances of distorted folds. Peptides with good antigenicity may serve as antigenic epitopes. Combined with the above factors, predict the B-cell epitopes of EgAgB1 are located at positions: 8–15 and 31–37. Which Sequences are DDGLTSTS and RDPLGQK. Predict the B-cell epitopes of EgAgB3 are located at positions: 20–27 and 47–53. Which Sequences are ADDDDDEV and SDPLGKK. In combination with previous studies on the expression levels of EgAgB1 and EgAgB3 in human and animal serum, and RNA data, we further determined the important role of EgAgB1 and EgAgB3 in the early diagnosis of echinococcosis.
5. Acknowledgments
This study was supported by the National Nature Science Foundation of China (81660343, 81460307, and 31560262)

6. References
[1] Pang N, Zhang F, Ma X, Zhu Y, Zhao H, Xin Y, et al. TGF-β/Smad signaling pathway regulates Th17/Treg balance during Echinococcus multilocularis infection. Int Immunopharmacol. 2014; 20:248–57.
[2] Pang N, Zhang F, Ma X, Zhang Z, Zhao H, Xin Y, et al. Th9/IL-9 profile in human echinococcosis: their involvement in immune response during infection by Echinococcus granulosus. Mediators Inflamm. 2014; 2014: 781649.
[3] Alvarez Rojas CA, Romig T, Lightowlers MW. Echinococcus granulosus sensu lato genotypes infecting humans--review of current knowledge [J]. Int J Parasitol, 2014, 44(1): 9-18
[4] Nakao M, Lavikainen A, Yanagida T, et al. Phylogenetic systematics of the genus Echinococcus (Cestoda: Taeniidae)[J]. Int J Parasitol, 2013, 43(12-13):1017-1029.
[5] Wang H, Zhang F, Ma X, et al. Prokaryotic expression and identification of B- and T-cell combined epitopes of Em95 antigen of Echinococcus multilocularis[J]. Int J Clin Exp Pathol, 2014, 7(8):5117-5122.
[6] Koz S, Sahin I, Kayabas U and Kuzucu C. Brucella and peritoneal dialysis related peritonitis:case report and review of literature. Clin Nephrol 2014; 82: 283-286.
[7] Zheng XL. B-cell epitope KT-12 of enterohemorrhagic Escherichia coli O157:H7: a novel peptidovaccine candidate. Microbiol Immunol2011; 55: 247-253.
[8] Tae H, Shallom S, Settlage R, Hawkins GN, Adams LG and Garner HR. Complete genome sequence of Brucella suis VBI22, isolated from bovine milk. J Bacteriol 2012; 194: 910.
[9] Zhang W, Li J, Jones MK, er al. The Echinococcus granulosus antigen B gene family comprises at least 10 unique genes in five subclasses which are differentially expressed [J]. PLoS Negl Tro Dis, 2010, 4(8):e784.
[10] Silva-Álvarez V, Folle AM, Ramos AL, et al. Echinococcus granulosus Antigen B binds to monocytes andmacrophages modulating cell response to inflammation[J]. Parasit Vectors, 2016 Feb 4; 9:69.