Analysis of Structures, Functions, and Epitopes of Aminopeptidase from *Trichinella spiralis*

Zhang, Y.L.[1,2], Zhang, H.W.[2], Zhang, X.[1], Liu, L.N.[1], Liu, R.D.[1], Xu, B.L.[2], Wang, Z.Q.[1*] and Cui, J.[1*]
[1]Department of Parasitology, Medical College, Zhengzhou University, Zhengzhou 450052, China
[2]Institute of Parasitic Diseases, Center for Disease Control and Prevention of Henan Province, Zhengzhou 450016, China
*Correspondence should be addressed to Zhong Quan Wang; wangzq@zzu.edu.cn or Jing Cui; cuij@zzu.edu.cn

Abstract. We have previously reported that the recombinant *T. spiralis* aminopeptidase (rTsAP) could induce a partial protective immunity against *T. spiralis* infection in mice. The aim of this study was to predict the structures and functions of TsAP protein by using the full length cDNA sequence of TsAP gene. TsAP sequence was 1515 bp length with a 1515 bp biggest ORF encoding 504-amino acid protein. The molecular weight and isoelectric point of TsAP were 54.7 kDa and 6.69, respectively. TsAP structure domains contained a Peptidase_M17_N and a Peptidase_M17 domain, which has the function of catalysis of the hydrolysis of N-terminal amino acid residues. TsAP had no signal peptide site and transmembrane domain, and located in cytoplasm. The secondary structure of TsAP contained 16 α-helix, 14 β-strand and 29 coils. The TsAP had 11 and 21 potential antigenic epitopes of T cell and B cell, respectively. Based on the phylogenetic analyses of TsAP, *T. spiralis* have the closest relationship with *Plasmodium falciparum*. TsAP was a kind of proteolytic enzyme with a variety of biological functions and its antigenic epitopes could provide important insights on the diagnostic antigens and target molecular of anti-*Trichinella* drugs.

INTRODUCTION

Trichinella spiralis is an intracellular parasitic nematode of mammalian skeletal muscles. The invasion of host intestinal epithelial cells (IECs) by the infective larvae is the first step during *T. spiralis* infection. The infective larvae invade the intestinal epithelium, where they mature to the adult stage, mate, and produce newborn larvae, which enter the blood and migrate to skeletal muscles where they grow and mature to the infective stage, thus completing the life cycle (Despommier, 1998, Wang *et al.*, 2013b). However, the mechanisms by which *T. spiralis* infective larvae recognize, invade, and migrate within the intestinal epithelia are unknown.

Our previous studies showed that when the *T. spiralis* infective larvae were inoculated onto the monolayer of IECs, they invade the IECs and produced several proteins, and some of these proteins entered the IECs (Wang *et al.*, 2011, Wang *et al.*, 2012). Out of the proteins produced by the infective larvae after co-culture with IECs, *T. spiralis* aminopeptidase (TsAP, GenBank accession No. EFV57052) was identified by shotgun LC-MS/MS (Wang *et al.*, 2013a). The TsAP gene was transcribed and expressed during all the different developmental stages of *T. spiralis*, suggesting that the TsAP is an indispensable protein and plays an important role in the life cycle of *T. spiralis*. TsAP appears to be a cytoplasm protein located primarily at the cuticle and internal organs of this parasite (Zhang *et al.*, 2013).

TsAP is a kind of proteolytic enzymes, which can catalyze the amino acid released from the N-terminal of polypeptide chain and plays an important role in the degradation of some bioactive peptides (Taylor, 1993),

suggesting that the TsAP might be related with the larval invasion of IECs and that these proteins might mediate or facilitate the entry into cells. In the present study, the full-length cDNA sequence of TsAP (GenBank accession No. EFV57052) was analyzed; its structure and function were predicted by using bioinformatics techniques.

MATERIALS AND METHODS

The full-length cDNA sequence of TsAP was used in this study. The analysis methods of structures, functions, and epitopes of TsAP applied here were performed as described previously (Liu *et al.*, 2015). The structure domain and function domain were predicted by online analysis http://smart.embl-heidelberg.de/. The secondary structures were predicted by one program named Protean in DNAstar (Garnier *et al.*, 1978). The antigenic index was predicted by Jameson-Wolf method (Jameson & Wolf, 1988), the flexible regions were predicted by Karplus-Schulz method (Plasterer, 1997). The surface probability was predicted by the method of Plot-Emini (Emini *et al.*, 1985), and used to value the immunogenicity. The 3D models of proteins were constructed by SWISS-MODEL, a protein structure server on the website http://swissmodel.expasy.org/, which is considered to predict protein 3D structures that have more than 100 amino acids (Arnold *et al.*, 2006, Guex & Peitsch, 1997, Schwede *et al.*, 2003). Other aminopeptidase amino sequences of model organisms of other parasites used in this study were obtained from GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html) and listed as follows: *Wuchereria bancrofti* (EJW88036.1), *Caenorhabditis remanei* (EFP02419.1), *Caenorhabditis elegans* (CCD73205.1), *Brugia malayi* (EDP39509.1), *Ascaris suum* (ERG84966.1), *Necator americanus* (ETN85978.1), Loa loa (EFO24712.1), *Haemonchus contortus* (CDJ86298.1), *Plasmodium falciparum* (AAN37052.1), *Fasciola hepatica* (AAV59016.1). The multiple sequence alignment of TsAP and the above-mentioned sequences were carried out by Clustal X v2.0 (Larkin *et al.*,

2007); Then, the phylogenetic relationship of *T. spiralis* with other related species was estimated through four phylogenetic inference methods: neighbor-joining (NJ), maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI), respectively. NJ, MP and ML analyses were performed in MEGA v6.0 (Tamura *et al.*, 2007). BI analysis was performed in MrBayes v3.1 (Ronquist and Huelsenbeck, 2003) with 5,000,000 generations, sampling trees every 100 generations.

RESULTS

The Basic Properties of TsAP Sequence
The TsAP sequence was of 1 515 bp length with a 1515 bp biggest OFR from 1 bp (ATG) to 1515 bp (TGA), which encoded 504-amino acid protein. Nucleotide sequence and deduced amino acid sequence were shown in Figure 1.

Physical and Chemical Properties of TsAP
The TsAP had the molecular weight of 54.7 kDa and theoretical isoelectric point (pI) of 6.69. Extinction coefficients are 50600 M^{-1} cm^{-1}, at 280 nm measured in water, assuming all pairs of Cys residues form cysteines. The half-life was 30h, >20h, and >10h in mammalian reticulocytes (*in vitro*), yeast (*in vivo*), and *Escherichia coli* (*in vivo*), respectively. The instability index (II) was computed to be 31.37. This classifies the protein as stable. Aliphatic index is 86.37. Grand average of hydropathicity (GRAVY) is –0.125.

The basic characteristics of TsAP and subcellular localization
The confidently predicted TsAP structure domains contained a Peptidase_M17_N located at 7aa–144aa and a Peptidase_M17 domain located at 180aa–495aa, which has the function of catalysis of the hydrolysis of N-terminal amino acid residues from a polypeptide chain.

The results predicted by DNAStar showed that α-helix was the main secondary structure. Using the scale Hphob./Kyte & Doolittle, the result of hydrophilicity was

Figure 1. Sequences and amino acid residues of TsAP. The TsAP sequence was of 1515 bp length with a 1015 bp biggest OFR from 1 bp (ATG) to 1515 bp (TGA), which encoded 504-amino acid protein.

showed in Figure 2. Above the abscissa axis were hydrophilic regions; under the axis were hydrophobic regions. The results of all antigenic index, flexible regions and surface probability were shown in Figure 2. The flexible region is easy to bind to antibodies, so it might have high potential to be the epitope. PSIPRED v. 3.3 was also used to predict the secondary structures of TsAP which had 16 α-helix, 14 β-strand and 29 coils (Figure 3).

The prediction results of TsAP signal peptide by Signal SignalP 4.0 Server showed that there was no peak fraction and the TsAP protein had no signal peptide. The results also indicate that TsAP did not belong to the secretory protein. Prediction of transmembrane domain of TsAP with TMHMM Server v2.0 suggested that the TsAP had no transmembrane domain. Results of the k-NN prediction of TsAP suggested that the peptide chain was located in the cytoplasmic, nuclear, vacuolar, extracellular (including cell wall), cytoskeletal, endoplasmic reticulum, peroxisomal, and mitochondrial with the possibility of 52.2%, 17.4%, 8.7%, 4.3%, 4.3%, 4.3% and 4.3%, respectively. The maximum possible location was in the cytoplasmic (k = 23).

Antigenic Epitopes of TsAP

The sequence of TsAP was submitted to the SYFPEITHI online soft to predict the T cell epitopes which could be recognized by the MHC molecules. The TsAP had 11 potential antigen epitopes (aa 155-169, 183-197, 204-213, 225-239, 298-312, 316-330, 331-345, 370-384, 396-410, 438-452, and 490-504). The B cell epitopes were analyzed by BepiPred online. There were 21 potential antigen epitopes (aa 16-19, 31-36, 47-51, 63-63, 75-93, 118-133, 146-157, 164-171, 175-188, 229-235, 245-253, 264-269, 285-293, 336-346, 359-365, 381-384, 405, 407-411, 422-429, 457-465, 494-512) on the sequence of TsAP.

Construction of 3D Model

The SWISS-MODEL a predicted server for protein tertiary structure was used to build the tertiary structure of TsAP. SWISS-MODEL is a fully automated protein structure homology-modeling server. The SWISS-
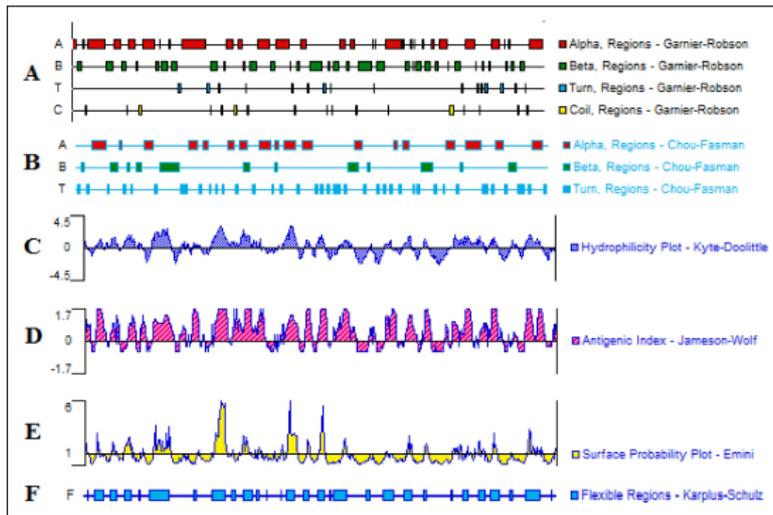
Figure 2. The basic characteristics of TsAP predicted by the software DNAstar. **A-B**: Secondary structure prediction respectively by Garnier-Robson and Chou-Fasman method. **C:** Hydrophilicity prediction by Kyte-Doolittle method. **D**: Antigenic index prediction by Jameson-Wolf method. **E**: Surfsace probability prediction by Plot-Emini method. **F**: Flexible regions prediction by Karplus-Schulz method.
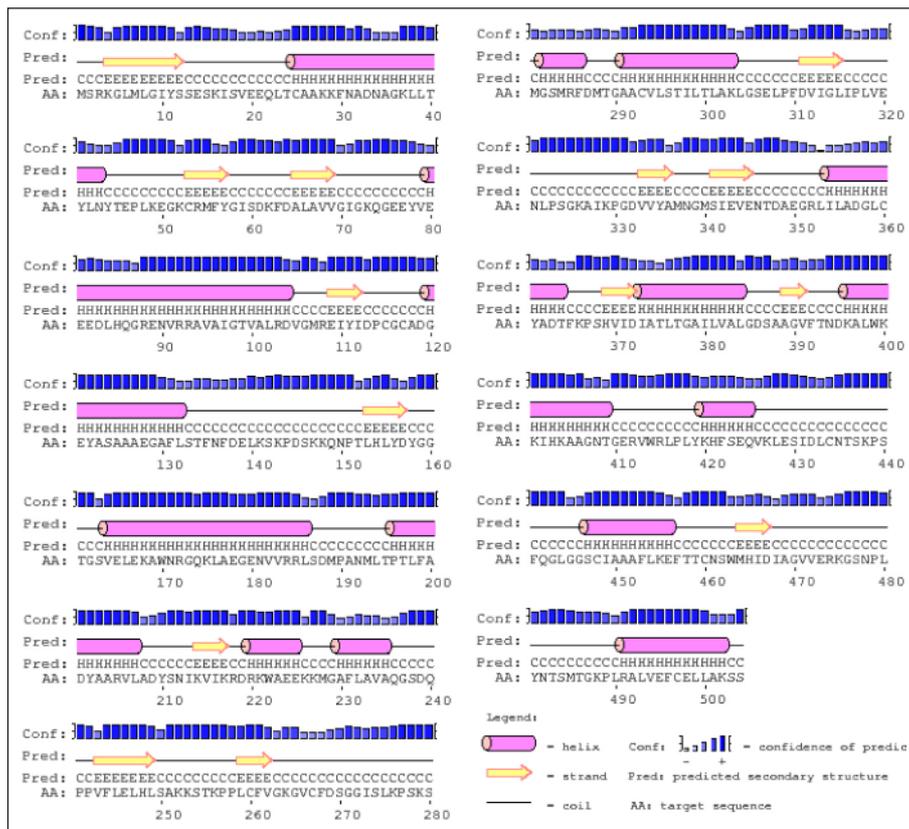
Figure 3. The predicted secondary structure of TsAP by using PSIPRED. There were 16 α-helix, 14 β-strand and 29 coils of the predicted secondary structure of TsAP.

MODEL repository is a database of annotated three-dimensional comparative protein structure models generated by the fully automated homology-modelling pipeline SWISS-MODEL. The modeling of TsAP was built on the basis of the crystal structure of bovine lens leucine aminopeptidase in complex with zofenoprilat (Figure 4).

Molecular Evolution of TsAP

As shown in Figure 5, phylogenetic trees based on four methods (NJ, MP, ML and BI) using TsAP sequences all supported the sibling relationship between *T. spiralis* and *P. falciparum* with high support values (bootstrap values 75, 94, 87 and Bayesian posterior probability 1.0).
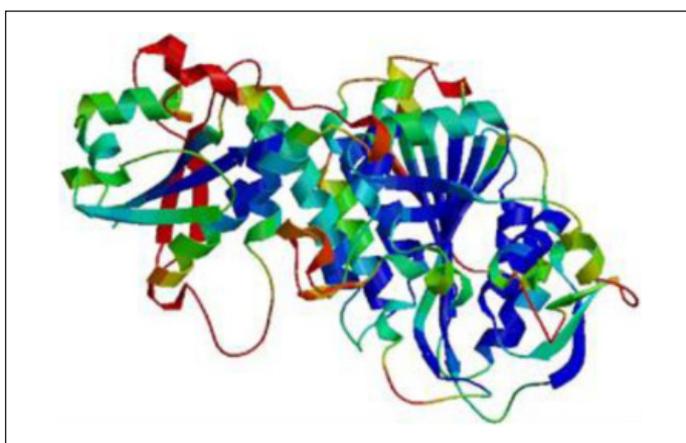
Figure 4. The 3D structure of TsAP predicted by SWISS-MODEL.
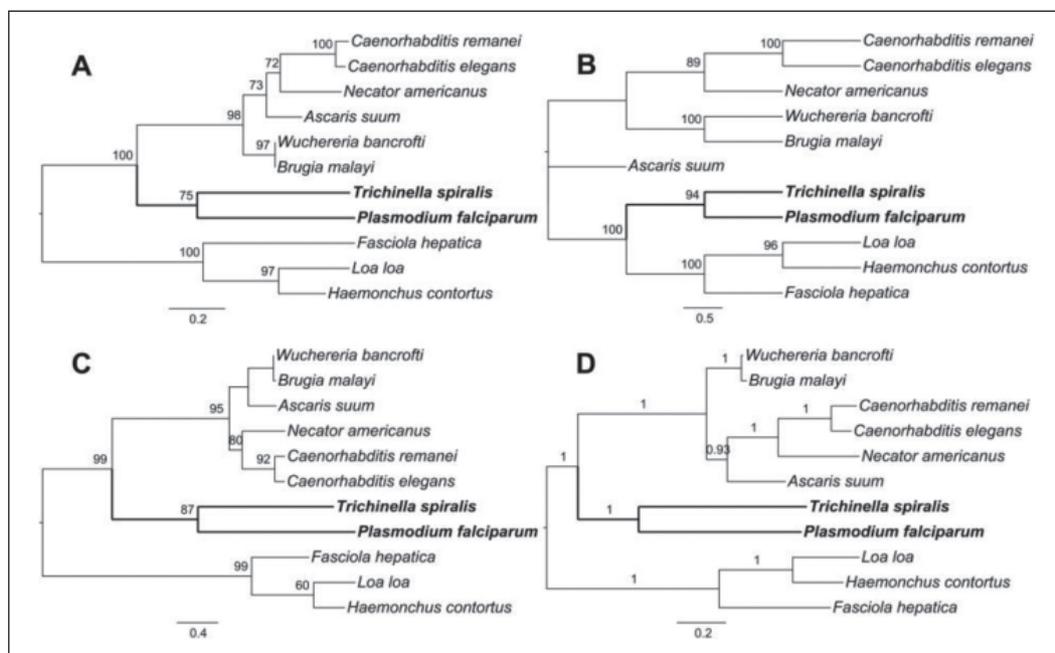
Figure 5. Phylogenetic analyses referred from TsAP based on neighbor-joining (Fig. 5A), maximum parsimony (Fig. 5B), maximum likelihood (Fig. 5C) and Bayesian inference (Fig. 5D), respectively. The numbers along branches indicate bootstrap values and posterior probabilities resulting from different analyses. Only bootstrap values above 60 and posterior probabilities above 0.6 are shown.

DISCUSSION

Based on the full-length cDNA sequence of TsAP, the sequence of TsAP gene was 1515 bp length with a 1515 bp biggest ORF encoding 504-amino acid protein. The predicted molecular weight and isoelectric point of TsAP were 54.7 kDa and 6.69, respectively. The predicted TsAP structure domains contained a Peptidase_M17_N and a Peptidase_M17 domain, which has the function of catalysis of the hydrolysis of N-terminal amino acid residues from a polypeptide chain. Aminopeptidases are exopeptidases that catalyze the sequential removal of amino acids from the N termini of peptides and play a major role in regulating the balance between catabolism and anabolism in all living cells.

Aminopeptidase belongs to the peptidase M17 family, which can be also classified into different aminopeptidases according to its major substrate. Leucine aminopeptidases (LAP) are the representative group of amino-peptidases (Taylor, 1993). They have been identified, puried and characterized in many helminth and protozoal parasites (Acosta *et al.*, 1998, Rhoads & Fetterer, 1998) and shown to play important roles such as moulting, surface membrane remodeling, egg hatching and digestion for the survival of parasites within host (McCarthy *et al.*, 2004, Rogers, 1982, Xu & Dresden, 1986). The tissue localization and functional analysis of LAP in *S. cervi* and several other nematodes also suggested that this enzyme is involved in parasite feeding, nal digestion of the partially hydrolysed peptide fragments within gastrodermal cells, cuticle remodeling, egg hatching and embryogenesis (Pokharel *et al.*, 2006, Richer *et al.*, 1992). Based on the phylogenetic analysis of TsAP, *T. spiralis* has the closest evolutionary status with *P. falciparum.*

The previous studied showed that TsAP gene was transcribed and expressed during the *T. spiralis* different developmental stages (adult worms, newborn larvae, pre-encapsulated larvae and muscle larvae), suggesting that the TsAP is an indispensable protein and plays an important role in the life cycle of *T. spiralis* (Zhang *et al.*, 2013).

The predicted results of TsAP in this study indicated that the TsAP had a good immunogenicity and might be valuable to further develop the diagnostic antigen of trichinellosis and target molecular of anti-*Trichinella* drugs (Wang *et al.*, 2013b).

REFERENCES

Acosta, D., Goni, F. & Carmona, C. (1998). Characterization and partial purification of a leucine aminopeptidase from *Fasciola hepatica. Journal of Parasitology* **84**: 1-7.

Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**: 195-201.

Despommier, D.D. (1998). How does *Trichinella spiralis* make itself at home? *Parasitology Today* **14**: 318-323.

Emini, E.A., Hughes, J.V., Perlow, D.S. & Boger, J. (1985). Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *Journal of Virology* **55**: 836-839.

Garnier, J., Osguthorpe, D.J. & Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *Journal of Molecular Biology* **120**: 97-120.

Guex, N. & Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**: 2714-2723.

Zhang, Y.L., Wang, Z.Q., Li, L.G. & Cui, J. (2013). Molecular characterization of *Trichinella spiralis* aminopeptidase and its potential as a novel vaccine candidate antigen against trichinellosis in BALB/c mice. *Parasites & Vectors* **6**: 246.

Jameson, B.A. & Wolf, H. (1988). The antigenic index: a novel algorithm for predicting antigenic determinants. *Computer Applications in the Biosciences* **4**: 181-186.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. & Higgins, D.G. (2007). Clustal W and clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

Liu, L.N., Wang, Z.Q., Zhang, X., Jiang, P., Zhang, Z.F., Zhang, G.Y. & Cui, J. (2015). Analysis of structures, functions and epitopes of *Spirometra erinaceieuropaei* casein kinase I. *Tropical Biomedicine* **32**: 167-175.

McCarthy, E., Stack, C., Donnelly, S.M., Doyle, S., Mann, V.H., Brindley, P.J., Stewart, M., Day, T.A., Maule, A.G. & Dalton, J.P. (2004). Leucine aminopeptidase of the human blood flukes, *Schistosoma mansoni* and *Schistosoma japonicum*. *International Journal of Parasitology* **34**: 703-714.

Plasterer, T.N. (1997). PROTEAN. Protein sequence analysis and prediction. *Methods in Molecular Biology* **70**: 227-239.

Pokharel, D.R., Rai, R., Kumar, P., Chaturvedi, C.M. & Rathaur, S. (2006). Tissue localization of collagenase and leucine aminopeptidase in the bovine filarial parasite *Setaria cervi*. *Filaria Journal* **5**: 7.

Rhoads, M.L. & Fetterer, R.H. (1998). Purification and characterisation of a secreted aminopeptidase from adult *Ascaris suum*. *International Journal of Parasitology* **28**: 1681-1690.

Richer, J.K., Sakanari, J.A., Frank, G.R. & Grieve, R.B. (1992). *Dirofilaria immitis*: proteases produced by third- and fourth-stage larvae. *Experimental Parasitology* **75**: 213-222.

Rogers, W.P. (1982). Enzymes in the exsheathing fluid of nematodes and their biological significance. *International Journal of Parasitology* **12**: 495-502.

Schwede, T., Kopp, J., Guex, N. & Peitsch, M.C. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Research* **31**: 3381-3385.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**: 1596-1599.

Taylor, A. (1993). Aminopeptidases: structure and function. *Federation of American Societies for Experimental Biology Journal* **7**: 290-298.

Wang, L., Wang, Z.Q. & Cui, J. (2013a). Proteomic analysis of the changed proteins of *Trichinella spiralis* infective larvae after co-culture in vitro with intestinal epithelial cells. *Veterinary Parasitology* **194**: 160-163.

Wang, L., Wang, Z.Q., Hu, D.D. & Cui, J. (2013b). Proteomic analysis of *Trichinella spiralis* muscle larval excretory-secretory proteins recognized by early infection sera. *Biomed Research International* **2013**: 139745.

Wang, S.W., Wang, Z.Q. & Cui, J. (2011). Protein change of intestinal epithelial cells induced in vitro by *Trichinella spiralis* infective larvae. *Parasitology Research* **108**: 593-599.

Wang, Z.Q., Wang, L. & Cui, J. (2012). Proteomic analysis of *Trichinella spiralis* proteins in intestinal epithelial cells after culture with their larvae by shotgun LC-MS/MS approach. *Jouranl of Proteomics* **75**: 2375-2383.

Xu, Y.Z. & Dresden, M.H. (1986). Leucine aminopeptidase and hatching of *Schistosoma mansoni* eggs. *Journal of Parasitology* **72**: 507-511.