Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Medicine

journal homepage:www.elsevier.com/locate/apjtm

Document heading doi:

Identification, structure and function of a novel tetraspaninhomologue from *Spirometra erinaceieuropaei*

Fu Qiongyao[1], Lu Yajun[2], Rao Langyu[2], Chen Jinglong[2], Wu Qiang[3], Cai Qunfang[3], Li Lihua[2], Wu Lixian[2], Lu Gang[2*]

[1] *Department of Clinical Laboratoy, Hainan Medical College, Haikou 571101, P.R.China*
[2] *Department of Pathogen Biology, Hainan Medical College, Haikou 571101, P.R.China*
[3] *Department of Clinical Biochemistry, Hainan Medical College, Haikou 571101, P.R.China*

ARTICLE INFO

Article history:
Received 11 February 2011
Received in revised form 11 April 2011
Accepted 15 June 2011
Available online 20 September 2011

Keywords:
Spirometra erinaceieuropaei
Tetraspanin
Prediction
Structure
Function

ABSTRACT

Objective: To identify a full length cDNA sequence of a novel tetraspanin (TSP) homologue from *Spirometra erinaceieuropaei* and to predict the structure and function of its encoding protein using bioinformatics methods. **Methods:** Using the NCBI, EMBI, Expasy and other online sites, the open reading frame (ORF), conserved domain, physical and chemical parameters, signal peptide, transmembrane domain, epitope, topological structures of the protein sequences were predicted. And Vector NTI software was used for multiple sequence alignment and phylogenetic tree construction. **Results:** The target sequence was 1 132 bp length with a 681 bp biggest ORF encoding 226 amino acids protein with typical TSP conserved domain. It was confirmed as full length cDNA of TSP16 from *Spirometra erinaceieuropaei* and named as *Se*TSP16 (GenBank accession number: JF728872). The predicted molecular weight and isoelectric point of the deduced protein were 24 750.5 Da and 7.88 Da, respectively. Compared with TSP16s from *Schistosoma japonicum* and *Schistosoma mansoni*, it showed similarity of 59% and 59%, respectively. *Se*TSP16 contained four transmembrane domains (TM1-4), intracellular N and C-termini, one short small extracellular loop and one large extracellular loop. Four major epitopes that were significant different from the corresponding epitope regions of TSP16 from *Schistosoma mansoni* and *Schistosoma japonicum* were predicted. **Conclusions:** The full length cDNA sequences of *Se*TSP16 are identified. It encodes a transmembrane protein which might be an ideal diagnosis antigen and target molecule for antiparasitic drugs.

1. Introduction

Plerocercoidosis is a parasitic infection caused by the plerocercoid larvae, the 3rd stage larva of diphyllobothroid tapeworms belonging to the genus Spirometra, most commonly *Spirometra erinaceieuropaei* (*S. erinaceieuropaei*), also called *Spirometra mansoni* (*S. mansoni*) in Asia and *Spirometra mansonoides* in North America. Humans are the accidental hosts in the life cycle, while dogs, cats, and other mammals are definitive hosts. Copepods (freshwater crustaceans) are the first intermediate hosts, and various amphibians and reptiles are second intermediate hosts[1]. The first human plerocercoidosis case was reported by Patrick Manson in Xiamen, China in 1882[2]. The parasite is transmitted to humans in three different ways. First, humans may acquire the infection by drinking water that is contaminated with copepods housing *Spirometra* larvae. Second, humans may acquire the infection by consuming the raw flesh of one of the second intermediate hosts, such as frogs or snakes. Third, humans may acquire the infection by placing raw poultices of the second intermediate hosts on open wounds, lesions, and/or the eyes for medicinal or ritualistic reasons[3]. China is the main epidemic country of plerocercoidosis in the world. Most cases occur in Southern China and Eastern China. Main risk factors include applying meat of infected frog as poultice, eating raw meat of frog/snake and drinking contaminated water. Risk factors are differenct in different areas in China for different tradition and eating habits. Plerocercoidosis caused by applying frog meat as poultice is deduced greatly and that caused by eating raw meat of frog/snake is increased

*Corresponding author: Gang Lu, Prof., PhD, Department of Pathogen Biology, Hainan Medical College, Haikou, Xueyuan Road, 571101, Hainan, China.
Tel: +86-0898-66893745
Fax: +86-0898-66984309
E-mail: luganghn@yahoo.com.cn

greatly (especially for CNS plerocercodosis), indicating plerocerocoidosis has been one of main food-born parasitic diseases in China[4].

Based on construction of full-length cDNA library from adult *S. erinaceieuropaei*[5], large-scale sequencing of 5′ end expressed sequence tags (EST) and Blastn/Blastx searching for the ESTs, an EST from clone GDTC006_D09 having high homologus ratio (58%) with tetraspanin (TSP) 16 from *S. mansoni* (*Sm*TSP16) is identified, which has a 5′ end but without 3’ end or polyA. It is presumed as a cDNA fraction coding TSP 16 of *S. erinaceieuropaei*. Then the full-length cDNA sequence with polyA is obtained by the method of walking sequencing and gene splicing. This study is to identify the obtained cDNA sequence and to predict the structure and functions of its encoding protein using bioinformatics methods in the hope of getting valuable information for the further study.

2. Materials and methods

2.1. Materials

Construction of full-length cDNA plasmid library for adult *S. erinaceieuropaei*[2], large-scale sequencing of 5′ end EST, Blastn/Blastx searching and analysis, walking sequencing and gene splicing had been done by the cooperation between our team and United Gene Group Limited, Shanghai.

The full-length cDNA sequence of clone GDTC006_D09 was used in this study. Other TSP16 or TSPs amino sequences of model organisms and other parasites used in this study were obtained from GenBank (http: //www.ncbi.nlm.nih.gov/Genbank/index. html) and listed as follows: *S. mansoni* (*Sm*TSP16, XP_002573925.1), *Schistosoma japonicum* (*S. japonicum*) (*Sj*TSP16, AAW24863.1), *Danio rerio* (*Dr*CD9, NP_997784.1), *Xenopus laevis* (*Xl*CD9, NP_001085461.1), *Gallus gallus* (*Gg*CD9, NP_990093.1), *Mus musculus* (*Mm*CD9, NP_031683.1), and *Homo sapiens* (*Hs*CD9, NP_001760.1).

2.2. Methods

Methods applied here were the same as reported in our previous work[6]. Briefly, with the help of NCBI, EMBI, Expasy and other online sites, the protein information was obtained including the open reading frame (ORF), conserved domain, physical and chemical parameters, signal peptide, transmembrane structure, epitope, topological structures of the protein sequences. Vector NTI software was used for sequence alignment and phylogenetic tree construction.

3. Results

3.1. Gene identification

Target sequence was 1 132 bp length with a 681 bp biggest OFR started from 76 bp (ATG) and ended at 756 bp (TGA), which encoded 226 amino acid protein with polyA locating at the positions 1 115-1 132 bp. Nucleotide sequence and deduced amino acid sequence were shown in Figure 1. The deduced protein sequence contained a complete TSP conserved domain and a TSP-LEL domain(Figure 2). Blast results showed it had highest similarity (59%) with TSP family protein 16 of *S. mansoni* (*Sm*TSP16) and *S. japonicum* (*Sj*TSP16), suggested that the insert sequence of clone GDTC006_D09 was full-length cDNA of a novel TSP family protein 16 homologue of *S. erinaceieuropaei*, named as *Se*TSP16. Its accession number of GenBank was JF728872.

```
1    CCGCTGAGGTCGCAGCTAAGGTCTAGGTGACAATTTAGGAATTTGTAATAGAAGGTGGAT
61   TCTAGTCCTTAAACG ATG ACCAAAATATCACTTACCTGCTGTTTTAAATGTGTCAAGTAC
                     M  T  K  I  S  L  T  C  C  F  K  C  V  K  Y
121  GCGCTCTTTGTCTTCTGTCTGGTCGCTTGGCTTCTTGGTCTCATCGCGTTCGTCTGGGGT
       A  L  F  V  F  C  L  V  A  W  L  L  G  L  I  A  F  V  W  G
181  ATTGTCGCGAGGACGACAGGCAGTTTTCGACCACTGGATCAGCATTTTCCCGCAGTAAAC
       I  V  A  R  T  T  G  S  F  G  P  L  D  Q  H  F  P  A  V  N
241  TCTGGCGCGAATCTCCTGATTACAGTTGGCTTCTTCATTATGCTGGTGGCGACCCTGGGC
       S  G  A  N  L  L  I  T  V  G  F  F  I  M  L  V  A  T  L  G
301  TGCTGTGGAGCTATCCGCGAAAGCCAATGCATGCTTATGTCGTTCTTCTTCTCCATATTC
       C  C  G  A  I  R  E  S  Q  C  M  L  M  S  F  F  F  S  I  F
361  ATGTGCTTTACGCTTCTCATGGCGGCTGGGCTTTGGGCAGTTGCTTGGCAACCTAGGCTG
       M  C  F  T  L  L  M  A  A  G  L  W  A  V  A  W  Q  P  R  L
421  GGCGTCTACCTCTTCAACAGTCTCGAAAAGCTTGTAAATAACTACAAACTGAATGAAGAC
       G  V  Y  L  F  N  S  L  E  K  L  V  N  N  Y  K  L  N  E  D
481  CCCGATGATATCAAGCTACTCGACTTCATTCAAACCAAGTTCGTCTGCTGCGGAGTGACG
       P  D  D  I  K  L  L  D  F  I  Q  T  K  F  V  C  C  G  V  T
541  GGAGCCGTCGACTTCGTGCCTAAGAAGGCGCCAGAAAGTTGTGGAGGAACAACGAAGGAA
       G  A  V  D  F  V  P  K  K  A  P  E  S  C  G  G  T  T  K  E
601  ATTTCTCAGAGGCCGGGATGCCACTCGATGATACAACAGGCTGGACAGGACAACCTTTCA
       I  S  Q  R  P  G  C  H  S  M  I  Q  Q  A  G  Q  D  N  L  S
661  ACCATTACGGGCATTGGCATCGGATTTGCCATGATAATGATTTGCGGAATGCTATTCTCG
       T  I  T  G  I  G  I  G  F  A  M  I  M  I  C  G  M  V  F  S
721  TTGATGCTCTGCTGTTCGTTACGAGAAATAAAT TGA GATATTCGCGACCAACCTATCGGA
       L  M  L  C  C  S  L  R  E  I  N
781  TCTTTTCTTCTCTACTTCTCACATGACATTAAACCAATACGTATTGTTTTAAAGGAAAGC
841  ACGCACTGCGCCTCCTGCTGTTAGGCAAATGGTCTTCACTAACACTTGTAGAGAAGTCTC
901  TGGAGCCCTTCCCCGAGCTGTTTTGCATTTGACAAATACGAGTACACTTACTATCTCCCG
961  CTTCTTTTTTCCGCATCACCTGTGCTTTGCAATTTTTGGCCTTCTGCATGTCTTCTACACA
1021 AAAAAGATCCTGACAGTGCAAACTGCAAAACATCGGAAACACGAGTGCTTCTTGTCGATG
1081 TTTATGCCTTTTAATAAGATTATCAAAGACTTGTGAAAAAAAAAAAAAAAA
```

Figure 1. Full length cDNA sequence and deduced amino acid sequence of *Se*TSP16.

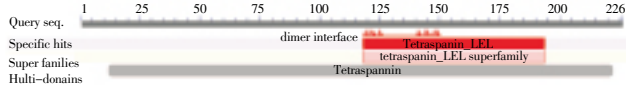

Figure 2. Putative conserved domains of *Se*TSP16.

3.2. Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignment and similarities of amino acid sequences of *Se*TSP16 with other TSPs were showed as Figure 3 and Table 1, respectively. The highest homology was with *Sm*TSP16 (59%) and *Sj*TSP16 (59%), while the lowest was in *Hs*CD9 (28%). Phylogenetic analysis showed that *Se*TSP16, *Sm*TSP16 and *Sj*TSP16 originated from the same ancestor. The evolutionary distance between *Se*TSP16 and *Hs*CD9 was far compared to that of between *Sm*TSP16/*Sj*TSP16 and *Hs*CD9 (Figure 4).

Table 1
Similarity of *Se*TSP16 and TSPs from other species.

	*Se*TSP16	*Sm*TSP16	*Sj*TSP16	*Dr*CD9	*Xl*CD9	*Gg*CD9	*Mm*CD9	*Hs*CD9
*Se*TSP16	100	59	59	30	32	30	28	28
*Sm*TSP16		100	98	32	31	30	29	27
*Sj*TSP16			100	33	31	31	29	27
*Dr*CD9				100	63	61	59	60
*Xl*CD9					100	69	71	71
*Gg*CD9						100	73	72
*Mm*CD9					100	89		
*Hs*CD9						100		

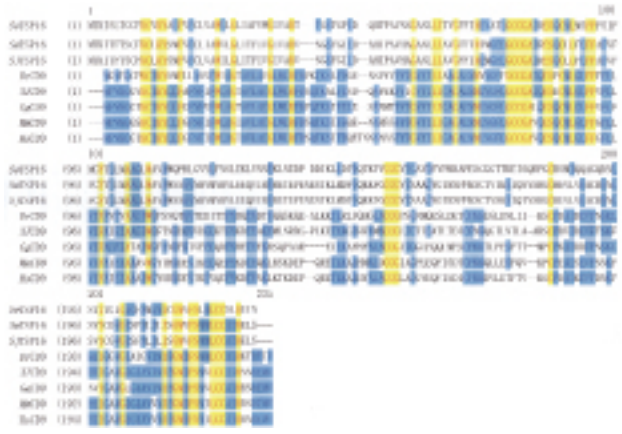

Figure 3. Multiple sequence alignment of *Se*TSP16 and TSPs from other species.

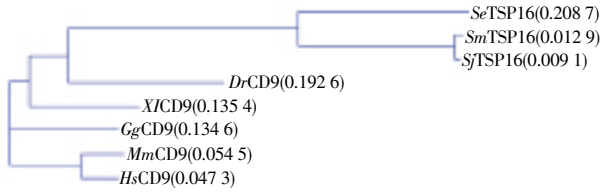

Figure 4. Phylogenetic analysis of *Se*TSP16 and TSPs from other species.

3.3. Physical and chemical parameters, signal peptide and subcellular localization prediction

The putative gene product of *Se*TSP16 had a predicted isoelectric point of 7.88 and a predicted molecular weight of 24 750 Da, its extinction coefficient value for solution at 280 nm was 27 345 units/M/cm. The estimated half-life of protein predicted in mammals, yeast and *Escherichia coli* were 30 h, >20 h and >10 h respectively. Its instability coefficient was 22.95, thus *Se*TSP16 was predicted as a stable protein. A signal peptide cleavage site was located between 35aa-36aa. No target signal of mitochondria, nucleus or peroxisome was observed. Subcellular localization prediction showed *Se*TSP16 to be a cytoplasmic protein (94.1% reliability) which located in the endoplasmic reticulum, vacuolar, extracellular, Golgi, mitochondria with possibility of 55.6%, 11.1%, 11.1%, 11.1%, 11.1%, respectively.

3.4. Analysis of secondary structure and topology of protein, post-translation modification site

*Se*TSP16 had four transmembrane domains (TM1 16aa-38aa, TM2 58aa-80aa, TM3 87aa-109aa and TM4 200aa-222aa), intracellular N and C-termini and two extracellular domains, one short small extracellular loop (SEL) and one large extracellular loop (LEL) (Figure 5). Post-prediction modification site prediction showed that here were different types of phosphorylation sites on *Se*TSP16 including 1 *N*-glycosylation site (193-aa-196aa), 6 protein kinase C phosphorylation sites (172aa-174aa, 177aa-179aa, 221aa-223aa), 2 casein kinase Ⅱ phosphorylation sites (172aa-175aa, 221aa-224aa) and 4 *N*-myristoylation sites (75aa-80aa, 153aa-158aa, 199aa-204aa, 211aa-216aa). There were four cysteine (C) residues and a highly conserved 'CCG' motif in the *LEL* domain. These results also strongly implied that *Se*TSP16 was a transmembrane protein, belonging to tetraspans superfamily.

3.5. Linear-B cell epitopes prediction

Linear B-cell epitopes prediction showed that *Se*TSP16 had four epitopes (41aa-55aa, 131aa-138aa, 160aa-181aa, 188aa-195aa), 1 located in SEL and 3 located in LEL. But, only two epitopes were existed on *Sm*TSP16/*Sj*TSP16 (131aa-140aa and 156aa-172aa) and *Hs*CD9 (42aa-56aa and 130aa-141aa), respectively(Figure 5). Compared these epitopes of *Se*TSP16 with the correspond epitopes of *Sm*TSP16, *Sj*TSP16 and *Hs*CD9, low similarity were observed among them, which implied that these epitopes might be ideal diagnostic antigen epitopes.

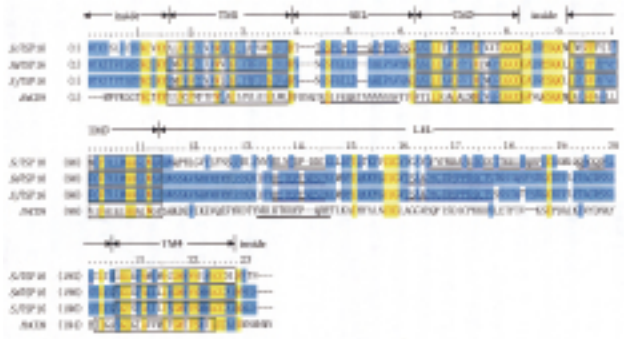

Figure 5. Transmembrane domains and epitope prediction of *Se*TSP16 and *Sm*TSP16, *Sj*TSP16 and *Hs*CD9.
inside: intracellular terminal; epitope: = ; TM:□ .

4. Discussion

Since Manson found the first case of plerocerocoidosis in Xiamen, China in 1882, there are several reports of plerocerocoidosis all over the world. Currently there are more than 1 400 cases of global record[7], and it is more common in East and Southeast Asian countries. According to incomplete statistics of plerocerocoidosis reported in official journals prior to April 2010 in mainland China by Lu *et al*[4], there were 963 cases of plerocerocoidosis, which accounting for approximately 1/2-2/3 of the global record, distributing at 24 provinces and cities. Nearly 60% of the cases distribute in the southeast coastal provinces. Provinces with high incidence are Guangdong, Hainan, Hunan, Fujian, Sichuan and Jilin. While Hainan has the highest incidence (1/100,000) (1.30). Plerocercoid can parasitize in all kinds of tissues and organs of the human body and it mainly parasitize in the subcutaneous muscles, eyes and central nervous system, accounting for 45.07%, 33.75% and 14.47%, respectively. Immunological diagnostic method for the auxiliary diagnosis of plerocerocoidosis is not widely used due to the sporadic condition and few cases[8,9]. Clinical misdiagnosis is common, and most are confirmed by biopsy or postoperative analysis.

The tegument is a dynamic host-interactive layer of schistosoma that is involved in nutrition, excretion, osmoregulation, immune evasion and modulation, sensory reception, and signal transduction[6]. Therefore, proteins exposed on the tegument surface are potential antigen targets for diagnosis, vaccines and drugs[10]. TSPs are a family of membrane proteins that are prominent in the outer tegument of schistosome[6]. They are characterized by 4 transmembrane domains (TM1-4) and 2 extra cytoplasmic regions, including a SEL, flanked by TM1 and TM2 domains, and a LEL, flanked by TM3 and TM4 domains. The first TSP gene identified in *S. japonicum* is *Sj*23, which becomes the traditional vaccine candidate for protection against the parasite[11]. Another member of the TSP family, TSP-2, expressed on the tegument of *S. mansoni* has received more attention as a protective antigen[12-17]. TSPs of other species have been suggested to play important roles in a series of biological processes[18,19]. But no study on TSPs of cestode was reported.

Based on construction of full-length cDNA library from adult *S. erinaceieuropaei*[7], large-scale 5′-EST sequencing, Blastn/Blastx, walking sequencing and gene splicing, a full-length cDNA (1 132 bp length with a 681 bp biggest OFR) having high homologus ratio (59%) with *Sm*TSP16 was obtained. It encoded a 226aa protein contained a complete TSP conserved domain and a TSP-LEL domain. The protein predicted has 4 transmembrane domains (TM1-4), intracellular N and C-terminal, 1 epitopes located in SEL and 3 located in LEL. Epitopes of *Se*TSP16 and that of *Sm*TSP16, *Sj*TSP16 and *Hs*CD9 have low similarity. These predictions implies that *Se*TSP16 is a typical transmembrane protein located on the tegument of *S. erinaceieuropaei* like TSPs of schistosoma. It might be an ideal diagnostic antigen and target molecule for antiparasitic drugs.

Conflict of interest statement

We declare that we have no conflict of interest.

References

[1] John DT, Petri WA. *Markell and Voge's medical parasitology*. 9th ed. St. Louis: Saunders Elsevier; 2006.

[2] Manson P, Manson-Bahr P, Wilcocks C. *Manson's tropical diseases: A manual of the diseases*. New York: William Wood and Company; 1921.

[3] Wiwanitkit V. A review of human plerocerocoidosis in Thailand. *Int J Infect Dis* 2005; **9**(6): 312-316.

[4] Lu G, Shi DZ, Lu YJ, Wu LX. Retrospective analysis for epidemiological factors of plerocercoidosis in mainland China. *J Hain Med Col* 2011.

[5] Lu G, Lu YJ, Fan AG. Construction and identification of a full-Length cDNA library from *Spirometra erinaceieuropaei*. *Chin J Parasitol Parasit Dis* 2010; **28**(5): 393-394.

[6] Van Hellemond JJ, Retra K, Brouwers JF, van Balkom BW, Yazdanbakhsh M, Shoemaker CB, et al. Functions of the tegument of schistosomes: clues from the proteome and lipidome. *Int J Parasitol* 2006; **36**(6): 691-699.

[7] Qiu MH, Qiu MD. Plerocercoidosis and plerocercoidosis Ⅰ: A historical review on aetiology. *Chin J Parasitol Parasit Dis* 2009; **27**(1): 54-60.

[8] Ding YX, Guo LL, Liu DW. Detection of anti *Spirometra erinaceieuropaei* antibody using ELISA. *Chin J Parasitol Parasit Dis* 2001; **19**(5): 303-304.

[9] Yeo IS, Yong TS, Im K. Serodiagnosis of human plerocerocoidosis by a monoclonal antibody-based competition ELISA. *Yonsei Med J* 1994; **35**(1): 43-48.

[10] Da′dara AA, Li YS, Xiong T, Zhou J, Williams GM, McManus DP, et al. DNA-based vaccines protect against zoonotic schistosomiasis in water buffalo. *Vaccine* 2008; **26**(29-30): 3617-3625.

[11] Tran MH, Pearson MS, Bethony JM, Smyth DJ, Jones MK, Duke M, et al. Tetraspanins on the surface of *Schistosoma mansoni* are protective antigens against schistosomiasis. *Nat Med* 2006; **12**(7): 835-840.

[12] Teixeira de Melo T, Michel de Araujo J, Do Valle Duraes F, Caliari MV, Oliveira SC, Coelho PM, et al. Immunization with newly transformed *Schistosoma mansoni* schistosomula tegument elicits tegument damage,reduction in egg and parasite burden. *Parasite Immunol* 2010; **32**(11-12): 749-759.

[13] Okoror LE, Eniolorunda TA, Okoror OI. Molecular evolutionary studies of Lassa virus nucleoprotein 2. *Asian Pac J Trop Dis* 2011; **1**(1): 28-34.

[14] Ramachandran R, Lakshmi R, Kumar RD, Devika K, Rahman F, Wares DF. Fast track method for the identification of multi-drug resistant tuberculosis on direct clinical specimen using combined drug media. *Asian Pac J Trop Dis* 2011; **1**(1): 47-49.

[15] Magana-Arachchi DN, Medagedara D, Thevanesam V. Molecular characterization of *Mycobacterium tuberculosis* isolates from Kandy, Sri Lanka. *Asian Pac J Trop Dis* 2011; **1**(3): 181-186.

[16] Alaya-Bouafif NB, Chahed MK, Bez HE, Bellali H, Ayari L, Achour N. Completeness of malaria notification in Tunisia assessed by capture recapture method. *Asian Pac J Trop Dis* 2011; **1**(3): 187-191.

[17] Rahimi MT, Sharifdini M, Ahmadi A, Laktarashi B, Mahdavi SA, Kia EB. Hydatidosis in human and slaughtered herbivores in Mazandaran province, northern Iran. *Asian Pac J Trop Dis* 2011; **1**(3): 212-215.

[18] Hemler, ME. Tetraspanin proteins mediate cellular penetration, invasion, and fusion events and define a novel type of membrane microdomain. *Annu Rev Cell Dev Biol* 2003; **19**: 397-422.

[19] Levy S, Shoham T. The tetraspan in web modulates immune-signalling complexes. *Nat Rev Immunol* 2005; **5**(2): 136-148.