ORIGINAL PAPER

The adaptive evolution divergence of triosephosphate isomerases between parasitic and free-living flatworms and the discovery of a potential universal target against flatworm parasites

Bing Chen · Jian-Fan Wen

Received: 3 April 2010 / Accepted: 28 December 2010 / Published online: 19 January 2011 © Springer-Verlag 2011

Abstract Triosephosphate isomerase (TIM) is an important drug target or vaccine candidate for pathogenetic organisms such as schistosomes. Parasitic and free-living flatworms shared their last common ancestor but diverged from each other for adapting to parasitic and free-living lives afterwards, respectively. Therefore, adaptive evolution divergence must have occurred between them. Here, for the first time, TIMs were identified from three free-living planarian flatworms, namely Dugesia japonica, Dugesia ryukyuensis, and Schmidtea mediterranea. When these were compared with parasitic flatworms and other organisms, the following results were obtained: (1) planarian TIM genes each contain only one intron, while parasitic flatworm genes each contain other four introns, which are usually present in common metazoans, suggesting planarian-specific intron loss must have occurred; (2) planarian TIM protein sequences are more similar to those of vertebrates rather than to their parasitic relatives or other invertebrates. This implies that relatively rapid evolution

Nucleotide sequence data reported in this paper are available in the GenBank[™] database under the accession numbers EU288186 and EU288187.

Electronic supplementary material The online version of this article (doi:10.1007/s00436-010-2249-4) contains supplementary material, which is available to authorized users.

B. Chen \cdot J.-F. Wen (\boxtimes)

State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China e-mail: wenjf@mail.kiz.ac.cn

B. Chen

Graduate School of the Chinese Academy of Sciences, Beijing 100039, China

occurred in parasitic flatworm TIMs; (3) All the investigated parasitic flatworm TIMs contain a unique tripeptide insert (SXD/E), which may imply its insertion importance to the adaptation of parasitic life. Moreover, our homology modeling results showed the insert region was largely surface-exposed and predicted to be of a B cell epitope location. Finally, the insert is located within one of the three regions previously suggested to be promising immunogenic epitopes in *Schistosoma mansoni* TIM. Therefore, this unique insert might be significant to developing new effective vaccines or specific drugs against all parasitic flatworm diseases such as schistosomiasis and taeniosis/ cysticercosis.

Abbreviations

TIM Triosephosphate isomerase EST Expressed sequence tag

Introduction

Flatworms (platyhelminthes) comprise planarians, trematodes, and cestodes. Planarians are mostly free-living, while the latter two are exclusively parasitic. Parasitic flatworms can cause serious diseases such as schistosomiasis and taeniosis/cysticercosis. Although great progresses have been made to prevent and treat these diseases, people still face various challenges. Schistosomiasis, for example, is still epidemic all over the world associated with high disability and mortality (Chitsulo et al. 2000; Gryseels et al. 2006). Praziquantel (PZQ), as the most effective chemotherapy drug currently used, has limitations such as not being effective enough against juvenile schistosomes to prevent reinfection (Ribeiro-dos-Santos et al. 2006). Moreover, schistosomes subjected to the drug pressure of PZQ can develop resistance (Fallon and Doenhoff 1994). Therefore, it is urgent to find new targets so as to develop specific drugs, effective vaccines, and new strategies to control these parasites.

Triosephosphate isomerase (TIM) is a vital glycolytic enzyme, which is responsible for the reversible isomerization between glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP). Due to its key role in energy metabolism, as well as its well-understood enzymatic and structural features, TIM of a pathogenetic organism is often considered as drug target or vaccine candidate (Gomez-Puyou et al. 1995; Velanker et al. 1997; Wu et al. 2005). TIMs of several parasitic flatworms have been identified (dos Reis et al. 1993; Hooker and Brindley 1996; Jimenez et al. 2000; Wei et al. 2006), and schistosomal TIM been identified as a protective antigen (Wu et al. 2005). Furthermore, Shoemaker et al. (1992) indicated three fragments within Schistosoma mansoni TIM, rather than the whole molecule, might be used as vaccine candidates so as to avoid human autoimmune responses. However, until now, neither drug nor vaccine aimed at parasitic flatworm TIMs has been successfully applied. To compare the TIMs from the two lifestyle flatworms might be helpful to identify new efficient target vaccines or drugs.

Dugesia japonica, Dugesia ryukyuensis, and Schmidtea mediterranea are well-studied triclad planarians. To date, expressed sequence tag (EST) databases of *D. ryukyuen*sis (Ishizuka et al. 2007), *D. japonica* (Mineta et al. 2003), and *S. mediterranea* (Alvarado et al. 2002; Zayas et al. 2005) have already been constructed, and the genome sequencing of *S. mediterranea* is also ongoing. Therefore, in the present work, first the full TIM gene and protein sequences of the three free-living flatworms were identified. Then, they were compared with those of parasitic flatworms as well as other metazoans. Some distinct features between TIMs of the two lifestyle flatworms were found, in particular, a novel parasitic flatworm TIMspecific insert. Its potential as a control target was investigated.

Materials and methods

D. japonica

The triclad planarian *D. japonica* was collected from a spring pool in a park, and cultured at 20–25°C. The identity was established by its shape, size, and its distribution in China (Kawakatsu et al. 1995). Adult worms used in our experiments had undergone 1 week of starvation.

Acquisition of the cDNA and DNA sequences of D. japonica TIM by experiments

Total RNA was isolated by using a TaKaRa Trizol Kit (TaKaRa Biotechnology (Dalian) Co., Ltd., Dalian, China) according to the manufacturer's instructions. Genomic DNA was extracted by a modified method described by Liu and Jiang (2005).

According to the reported EST sequence of the *D. japonica* TIM gene (GenBank accession number BP189215), a primer (5'-CTAATATAAAAgATTggACTCg-3') was designed to performed 3'-RACE experiment. The amplification parameters were as follows: 94°C for 2 min, followed by 32 cycles at 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min, followed by 72°C for 10 min. The amplification product of the expected size was gel isolated and sequenced after cloning into the pMD19-T cloning vector (TaKaRa).

According to the sequences obtained by 3'-RACE and of the known 5'-end, another pair of primers was designed as follows: sense primer, 5'-TCCAATATTTACTCAAAAA TATgTC-3' and antisense primer, 5'-ATgATAATAATATT CACTAAACC-3', in order that the product contains the entire coding region of the TIM gene. Then, PCR reactions were performed by using cDNA reverse transcribed from total RNA and genomic DNA as templates, respectively. PCR products were cloned and sequenced as described above.

Search and identification of two other planarian TIM sequences

EST database of D. ryukyuensis was downloaded from the Planarian EST Database (http://planaria.bio.keio.ac. ip/) and genomic sequences of S. mediterranea from the ftp server site at Washington University in St. Louis (http://genome.wustl.edu/). Two other EST databases of S. mediterranea (http://planaria.neuro.utah.edu/ and http://www.life.uiuc.edu/planaria) were searched online. Human TIM (GenBank accession number P60174) was used as a query to carry out local or online BLAST searches against each of the four databases at default settings, respectively (Altschul et al. 1997). Then, the best hits were used as queries to reciprocally BLASTP search against the GenBank non-redundant (nr) protein database. If the best hit of the reciprocal BLASTP was annotated as TIM, the obtained sequence was considered as a candidate sequence.

To further confirm the candidate planarian TIM sequences, Pfam program (http://pfam.sanger.ac.uk/) and ClustalW2 at EBI were used to analyze whether they contain the characteristic TIM domain and active sites conserved in TIMs.

Sequence alignment of metazoan TIMs

Metazoan proteins with unambiguous annotations of TIM in Swiss-Prot/TrEMBL (http://www.expasy.org/sprot/) were collected. TIMs of two trematodes, *Schistosoma haematobium* and *Orientobilharzia turkestanicum*, as well as TIM fragment of polyclad planarian, *Stylochus* sp., were collected from GenBank. In addition, TIM fragments of two parasitic nematodes, *Wuchereria bancrofti* and *Necator americanus*, were identified from Parasitic Helminth Genomes (the Sanger Center). For the sequences with high similarities from a common lineage (except flatworms and the mammalian hosts of parasitic flatworms), only one sequence was selected as a representative. (For the GenBank accession numbers of these selected sequences, please see Electronic Supplementary Material (ESM 1.pdf).)

Homology modeling and B cell epitope prediction

To generate a protein structure model of parasitic flatworm TIM, the sequence was submitted to CPHmodels-2.0 Server (Lund et al. 2002) and the modeling template was chosen automatically. Then the model was superposed with human TIM (PDB: 1HTI) by using the SSAP Server (http://www.cathdb.info/cgi-bin/cath/GetSsapRasmol.pl).

Based on the parasitic flatworm TIM models, linear, and discontinuous B cell epitopes were predicted by using the DiscoTope 1.2 server and BepiPred 1.0b server, respectively, with default settings (Haste-Andersen et al. 2006; Larsen et al. 2006), and solvent accessible surface areas were also calculated (Fraczkiewicz and Braun 1998).

Results

Identification of TIMs from three planarians, *D. japonica*, *S. mediterranea*, and *D. ryukyuensis*

D. japonica TIM (DjaTIM)

After a ~500-bp DNA sequence of *D. japonica* TIM, cDNA was obtained through the 3'-RACE experiment, the unknown part of the coding region, and the complete 3'untranslated region (UTR) sequence of this gene were obtained. Then a ~850-bp cDNA amplicon, which contains the full-length coding region, was obtained and sequenced. The electrophoresis patterns of related PCR products are shown in Fig. 1a. The open reading frame (ORF) encodes a 248-aa protein, which is most similar to chicken TIM when BLASTP searching against non-redundant protein database (*e* value: $6e^{-103}$, and ESM 2.pdf). Pfam program predicted that the protein sequence contains TIM domain (from Phe7 to Ile245). Moreover, it contains all the conserved amino acid



Fig. 1 a Electrophoresis patterns of PCR products of *D. japonica* TIM. *M*, marker; *Lane 1*, genomic DNA; *Lane 2*, PCR products of our 3'-RACE experiment of TIM; *Lanes 3* and 4, DjaTIM cDNA and DNA amplified from total RNA and genomic DNA, respectively. **b** A diagram illustrating *D. japonica* TIM gene structure. The coding sequence is showed by *black box*, which is disrupted by only one intron (from 116 to 167 nt) showed by *open box*

residues of known TIMs (e.g., Asn12, Lys14, His96, Glu166) and conserved motifs (e.g., 169-WAIGTG-174, which is critical for the substrate binding to the active center). Thus, *D. japonica* TIM cDNA and protein sequences were identified, and we named the protein DjaTIM.

The DjaTIM DNA sequence was obtained by PCR against genomic DNA, and then compared with the cDNA sequence obtained above. It showed that this gene contains a single 52-bp intron (from 116 nt to 167 nt), at the ends of which the typical GT/AG intron splicing signal can be found (Fig. 1b).

Both DjaTIM cDNA sequence and DNA sequence identified experimentally in this work have been deposited in GenBank (EU288186 and EU288187).

S. mediterranea TIM (SmeTIM)

High scoring homologs were obtained through database BLAST against both EST and genome databases of *S. mediterranea*. All displayed EST sequences were found to be almost identical in sequence, thus the longest one, PL06021B1D09 (GenBank accession number DN315650), was chosen for further analyses. It contains an ORF encoding a 248-aa protein, which has the highest scoring homology with chicken TIM (e value $5e^{-105}$) and has a positional identity of 87% with DjaTIM mentioned above. TIM-conserved residues and motifs were also found in the encoded peptide. Thus, *S. mediterranea* TIM cDNA and protein sequences were identified, and we named the protein SmeTIM.

The SmeTIM DNA sequence was also obtained by blasting the genome database, and then compared with the cDNA sequence. The result indicated that the SmeTIM gene also includes only one 47-bp intron, which shares a similar size and identical location with the DjaTIM gene. In addition, the genome database searching result indicated there is only one copy of the TIM gene in the *S. mediterranea* genome.

D. ryukyuensis TIM (DryTIM)

As described for SmeTIM, the *D. ryukyuensis* TIM cDNA sequence was identified (EST number Dr_sW_016_F10 and GenBank accession number BW639877), which encodes a

248-aa protein (named DryTIM). DryTIM has the highest scoring homolog with other two planarian TIMs (97% and 88% with DjaTIM and SmeTIM, respectively) and the second highest scoring homolog with chicken TIM (e value $1e^{-101}$), rather than with those of parasitic flatworms.

A unique tripeptide insert in TIMs of parasitic flatworms

The alignment of 35 metazoan TIM sequences revealed that the three newly identified planarian TIMs are quite conserved with all known metazoan TIMs. However, interestingly, within the region between the two conserved residues, Q147 and W161 (numbered according to *S. japonicum* TIM), all the collected TIMs of parasitic flatworms (trematode and cestode) exclusively possess a unique tripeptide insert—"SAD" in three trematodes (*S. mansoni*, *S. haematobium*, and *O. turkestanicum*), "SID" in another trematode *S. japonicum*, and "SKE" in cestode *Taenia solium*, and neither free-living flatworm TIM nor other known metazoan TIM shares such an insert (Fig. 2).

	: : *	* **	:*:	* :	*		*	:*:*	******	*** *:
S.purpuratus	KLKVVA	CIGEKL	DEREKGQ	TNEVVY	RQMR	ALADVIT	' <mark>DW</mark> (NVVIA	YEPVWAI	JTGKTATP
D.melanogaste	GLKVIA	CIGETL	EE <mark>REA</mark> GK	TNEVVA	RQMC	AYAQKI <mark>K</mark>	<mark>DW</mark> F	(NVVVA	YEPVWAI	JTGQTATP
M.prolifera	GLKVIA	CIGEKI	NE <mark>REA</mark> GQ	TEEVVK	TQL K	AISDRIQ	<mark>DW</mark>	NVVVA	YEPVWAI	JTGLSATP
X.maculatus	GLGVIA	CIGEKL	DEREGGI	TEKVVF.	AQTK	VIADNV <mark>K</mark>	DWS	KVVLA	YEPVWAI	JTGKTASP
D.rerio	GLGVIA	CIGEKL	DE <mark>REA</mark> GI	TEKVVF.	AQTKI	FIADNVK	DWS	KVVLA	YEPVWAI	JTGKTASP
Human	GLGVIA	CIGEKL	DE <mark>REA</mark> GI	TEKVVF	EQTK	VIADNV <mark>K</mark>	<mark>DWS</mark>	KVVLA	YEPVWAI	JTGKTATP
P.pygmaeus	GLGVIA	CIGEKL	DE <mark>REA</mark> GI	TEKVVF	EQTK	VIADNVK	<mark>DWS</mark>	KVVLA	YEPVWAI	JTG<mark>K</mark>TATP
M.fasciculari	GLGVIA	CIGEKL	DE <mark>REA</mark> GI	TEKVVF	EQTK	VIADNV <mark>K</mark>	<mark>DWS</mark>	KVVLA	YEPVWAI	FIGKTATP
Dog	GLGVIA	CIGEKL	DE <mark>REA</mark> GI	TEKVVF	EQTK	VIADNVK	DWS	KVVLA	YEPVWAI	JTGKTATP
Bovine	GLGVIA	CIGEKL	DE <mark>REA</mark> GI	TEKVVF	EQTK	VIADNV <mark>K</mark>	DWS	KVVLA	YEPVWAI	JTGKTATP
Mouse	GLGVIA	CIGEKL	DE <mark>REA</mark> GI	TEKVVF	EQTK	VIADNVK	<mark>DWS</mark>	KVVLA	YEPVWAI	JTGKTATP
P.colchicus	GLGVIA	CIGEKL	DE <mark>REA</mark> GI	TEKVVF	EQTK	AIADNVK	DWS	KVVLA	YEPVWAI	JTGKTATP
Chicken	GLGVIA	CIGEKL	DEREAGI	TEKVVF	EQTK	AIADNVK	DWS	KVVLA	YEPVWAI	JTGKTATP
A.brevirostru	GLGVIA	CIGEKL	DEREAGI	TEKVVF	E <mark>Q</mark> T <mark>K</mark> A	AIADNVK	DWS	KVVLA	YEPVWAI	JTG<mark>K</mark>TASP
Rat	GLGVIA	CIGEKL	DE <mark>REA</mark> GI	TEKVVF	EQTK	AIADNVK	DWS	KVVLA	YEPVWAI	JTGKTATP
P.aethiopicus	NLGVIA	CIGEKL	DEREAGI	TEQVVF	QQTK	AIAARV <mark>K</mark>	<mark>DW</mark> F	(<mark>P</mark> VVLA	YEPVWAI	GTGKTATP
B.germanica	GLNVVA	CIGEKL	EEREGGK	TEEVVF	QQTK	AIADKIK	DWS	KVVIA	YEPVWAI	GTGKTATP
C.tarsalis	GLKVIA	GETL	QEREAGQ	TEAVCE	RQTK	AIADKVK	DWS	NVVIA	YEPVWAI	GTGKTASP
A.merus	GLKVIA	CIGETL	QEREAGQ	TEAVCE	RQTK	AIAAQVK	DWS	NVVIA	YEPVWAI	GTGKTATP
T.solium	GLNVIP	CIGELL	SEREAGK	TNDVCF	AQMD.	AIAKNVP	SKEAWI	KVVIA	YEPVWAI	GTGKTATP
S.haematobium	GLSVIA	CIGETL	SERESNK	TEEVCV	RQLK	AIANKIK	SADEW	RVVVA	YEPVWAI	GTGKVATP
S.mansoni	GLSVIA	CIGETL	SERESNK	TEEVCV	RQLK	AIANKIK	SADEW	RVVVA	YEPVWAI	GTGKVATP
S.japonicum	GLSVVA	CIGETL	SERESGK	TEEVCV	RQLN	AIANKIK	SIDEW	RVVVA	YEPVWAI	GTGKVASP
0.turkestanic	GLNVIA	CIGETL	SERESGK	TEEVCV	RQLK	AIANKIK	SADQWH	RIVIA	YEPVWAI	GTGKVATP
D.ryukyuensis	GLSIIPO	GIGEKL	DERESNK	TNEVCF	RQLK	EIAANIK	DW	RVVIA	YEPVWAI	GTGKTASP
D.japonica	GLSIIP	CIGEKL	DERESNK	TNEVCF	R <mark>QL</mark> K	EIAANIK	DWJ	RVVVA	YEPVWAI	GTG <mark>K</mark> TASP
S.mediterrane	GLSVIPO	CIGEKL	EERENNK	TNEVCF	R <mark>QL</mark> K	EIAANIK	<mark>EWS</mark>	HVVIA	YEPVWAI	GTGRTATP
Stylochus sp.	GLSIIPO	CVGEKL	EERQAIK	TEEVVF	HQMK	AIADNVS	DWS	RVVIA	YEPVWAI	JTGQTATP
C.fornicata	GLSLIPO	CIGEKL	DERDAGK	TEEVVF	K <mark>Q</mark> M <mark>K</mark> I	FIA <mark>DN</mark> VS	<mark>DW</mark> F	RVVIA	YEPVWAI	GTGKTATP
A.miniata	GLKVIA	GETL	EQRQQGQ	TQEIVF	HQTK	AIA <mark>DN</mark> VT	' <mark>DW</mark> I	QVVIA	YEPVWAI	JTGVTATP
M.senile	GVKVIA	CVGELL	AE <mark>rea</mark> gk	TTEVVF	RQIA	AIAEHVS	<mark>DWS</mark>	KVVIA	YEPVWAI	JTGRTATP
C.lacteus	GLKIIPO	CIGEKR	EEREAGK	TEEVCF	RQL	AIVNNVS	<mark>DWS</mark>	NVVLA	YEPVWAI	JTGLTASP
C.elegans	GIKVVF	GIGEKL	EEREAGH	TKDVNF	RQLQ	AIVDKGV	'S <mark>W</mark> E	IVIA	YEPVWAI	JTGKTAS G
N.americanus	KINVIF	GIGEKL	EEREAGK	TKEVNF	R <mark>Q</mark> MQ/	ALVDQKV	<mark>DW</mark>	NIVIA	YEPVWAI	JTGKTATP
W.bancrofti	GLQTIF	CCGEKL	DEREAGK	TKAVNF	ROLO	AVIDKKV	NWN	KIVIA	YEPVWAI	GTGKTASP

Fig. 2 Partial results of multi-sequence alignment of metazoan TIMs. It is showed that parasitic flatworm TIMs possess a unique tripeptide insert. Species in the *two boxes* belong to the phylum Platyhelminthes.

The *upper box* contains five parasitic flatworms, and the *nether box* contains four free-living planarians

Furthermore, the insert sequence is quite conserved (the first position is a serine and the third one is an acidic amino acid, aspartic acid, or glutamic acid) and can be defined as SXD/E ("X" here stands for alamine, isoleucine, or lysine).

Comparative protein structure modeling and prediction of *B* cell epitopes

Our structure models of the five parasitic flatworm TIMs all display the canonical conserved TIM β -barrel structure, and the unique tripeptide insert is located between the secondary structural elements α -helix 5 and β -strand 6 (ESM 3.pdf).

In terms of *S. japonicum* TIM, the automated homology modeling server chose the chicken TIM as its modeling template (PDB: 8TIM) (its positional identity with *S. japonicum* TIM is 66.4%). When superposed with the crystal structure of human TIM (PDB: 1HTI), the model shows exactly like structure except the insert region (Fig. 3). The insert region was far from the active site and the dimer interface; furthermore, this region is nearly completely surface-exposed (ESM 4.pdf). Similar results were observed in other four parasitic flatworm TIMs in spite of choosing different templates by the server (data not shown).

Our prediction indicated that the tripeptide insert of parasitic flatworm TIMs is the main part of either a discontinuous B cell epitope (For *S. japonicum* TIM see ESM 5.pdf) or linear B cell epitopes, or both. All the three cases are suitable for vaccine development.



Fig. 3 Results of superposition of the modeled *S. japonicum* TIM (through automated comparative protein modeling and shown in *solid ribbon* presentation) with the TIM A-chain structure of *Homo sapiens* (1HTI, shown in *line ribbon* presentation). The tripeptide insert "SID" of *S. japonicum* TIM and the catalytic residue Glu 169 were marked and shown in *CPK* and *ball* and *stic view*, respectively

Discussion

In the present work, for the first time, full TIM gene and deduced protein sequences from three free-living flatworms were identified, which made it possible to compare this important glycolytic enzyme between parasitic and freeliving flatworms, as well as with those of other organisms. Based on these data, we investigated the divergence of TIMs between the two different lifestyle flatworms, and then attempted to identify some distinct features as potential targets against parasitic flatworms.

The adaptive evolution divergence of TIMs between the different lifestyle flatworms

The two identified planarian TIM genes each contain only one intron of a similar size and identical location among species. However, TIM genes of S. mansoni (dos Reis et al. 1993) and S. japonicum (of which the intron information was identified from its genome database by us) both contain five introns, of which the first one has similar size and identical location to the intron of planarian TIM genes. All five introns located in the schistosomal TIM gene, also arise in that of mammals and birds (chicken), as well as placozoa (Trichoplax adhaerens), Cnidaria (Nematostella vectensis), and even a choanoflagellate (Monosiga brevicollis), which was considered to be the closest relative of metazoans in protozoa (King et al. 2008). In addition, our phylogenetic analysis excluded the possibility of lateral gene transfer (LGT) between schistosomes and chickens/ mammals (data not shown). Therefore, the TIM gene in the last common ancestor of extant flatworms most likely contained at least the five conserved introns, and the absence of four of them in planarian TIM is mostly due to intron loss in the planarian lineage rather than numerous independent intron acquisitions in schistosomes, and various other animals. It is generally considered that parasitic organisms streamline their structures (including gene structure) under high selective pressure, but here we find an inverse process in intron numbers.

Like other TIMs, planarian TIMs are very conserved, sharing high positional identity with other metazoan TIMs. But unexpectedly, their highest scoring homolog is the chicken TIM followed by many other vertebrate TIMs, instead of those of their parasitic relatives (trematode and cestode), or other invertebrates. To test whether this is totally due to the unique tripeptide insert, another alignment was performed after deleting the insert, but the result was almost the same. This might imply a relatively rapid evolution occurred in parasitic flatworms. Such a rapid evolution might be an adaptive evolution for parasitic flatworms to fit the parasitic life.

Morphological and molecular evidence supports that Rhabditophora, which contains all the parasitic flatworms and most planarians (including all the four species studied here), is a monophyly (Carranza et al. 1997; Ruiz-Trillo et al. 1999), namely, they share a last common ancestor. However, our work showed that compared to their freeliving relatives, parasitic flatworms contain an additional tripeptide insert in the conserved TIMs. Our sequence analyses indicated that such an insert does not exist in all other known TIMs yet, including those from invertebrates (including Porifera, Cnidaria, Nematoda, Nemertea, Mollusca, Arthropoda, and Echinodermata in our analysis) and vertebrates-almost all the representative lineages of the Metazoa (Fig. 2 and ESM 1.pdf). Therefore, this tripeptide insert is specific to parasitic flatworm TIMs, suggesting that the insertion event must have occurred after the divergence of parasitic flatworms from free-living ones in evolutionary history. The insert sequences themselves are quite conserved among trematodes, and even between trematodes and cestodes, which suggests that the acquisition of the insert was a single evolutionary event and there might be some functional constraints against losing the insert or changing its sequence remarkably. Therefore, this unique insert of TIM might result from adaptive evolution and be of importance to enzymatic activity and even to the parasitic life of these parasitic flatworms.

A universal target against parasitic flatworms

Understanding the modifications that occurred during coevolution with hosts is not only fundamental to the understanding of parasitism, but also highly relevant for the design of antiparasitic drugs and vaccines (Brehm et al. 2006). In term of schistosomiasis, TIM is one of the most important vaccine candidates (Harn et al. 1992; Wu et al. 2005). However, TIM sequences are so conserved between schistosomes, and their mammalian hosts that autoimmunity of the host body would occur if using the whole schistosomal TIM as an antigen, and deficiency of TIM in the host (e.g., human) will result in hemolytic anemia and neuromuscular dysfunction (Daar et al. 1986). To avoid this, Shoemaker et al. (Shoemaker et al. 1992) recommended three regions within S. mansoni TIM to develop vaccines, which are rich in charges, hydrophilic, and share low identity with corresponding regions of the human TIM. Later, the fusion peptide of the first two regions from the S. japonicum TIM has been demonstrated to be an effective antigen and schistosomal specificity by other authors (Lou et al. 2001). Our identified tripeptide insert is within the second region, and the residues of the insert were predicted to be linear and/or discontinuous B cell epitopes. Therefore, we propose the region that contains the unique insert can be the first choice to develop more specific and effective vaccines, especially multiple antigenic peptide (MAP) vaccines, not just against *S. mansoni* but against all parasitic flatworms. Besides, this parasitic flatwormspecific insert might also be considered to be an ideal target candidate for developing specific chemotherapy drugs against the parasitic flatworm diseases such as schistosomiasis and taeniosis/cysticercosis.

Acknowledgments We wish to thank Yuan Lu for his help in culturing animals and Guang-wen Chen (Henan Normal University, China) for helping us to identify *Dugesia japonica*. We also want to thank Guo-li Zhang, Hai-feng Tian, and Gui-ling Sun for their very helpful discussions. This work was supported by grants (2006C0014Z; 2007C098M; 2005C0055M) from Yunnan province, 973 program (2007CB815705), and grants (30021004; 30623007) from the NSFC.

References

- Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402
- Alvarado AS, Newmark PA, Robb SMC, Juste R (2002) The Schmidtea mediterranea database as a molecular resource for studying platyhelminthes, stem cells and regeneration. Development 129:5659–5665
- Brehm K, Spillotis M, Zavala-Gongora R, Konrad C, Frosch M (2006) The molecular mechanisms of larval cestode development: First steps into an unknown world. Parasitol Int 55:S15– S21
- Carranza S, Baguna J, Riutort M (1997) Are the Platyhelminthes a monophyletic primitive group? An assessment using 18S rDNA sequences. Mol Biol Evol 14:485–497
- Chitsulo L, Engels D, Montresor A, Savioli L (2000) The global status of schistosomiasis and its control. Acta Trop 77:41–51
- Daar IO, Artymiuk PJ, Phillips DC, Maquat LE (1986) Human triosephosphate isomerase deficiency: a single amino-acid substitution results in a thermolabile enzyme. Proc Natl Acad Sci U S A 83:7903–7907
- dos Reis MG, Davis RE, Singh H, Skelly PJ, Shoemaker CB (1993) Characterization of the *Schistosoma mansoni* gene encoding the glycolytic enzyme, triosephosphate isomerase. Mol Biochem Parasitol 59:235–242
- Fallon PG, Doenhoff MJ (1994) Drug-resistant schistosomiasis: resistance to praziquantel and oxamniquine induced in *Schistosoma mansoni* in mice is drug specific. Am J Trop Med Hyg 51:83–88
- Fraczkiewicz R, Braun W (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. J Comput Chem 19:319–333
- Gomez-Puyou A, Saavedra-Lira E, Becker I, Zubillaga RA, Rojo-Dominguez A, Perez-Montfort R (1995) Using evolutionary changes to achieve species-specific inhibition of enzyme action - Studies with triosephosphate isomerase. Chem Biol 2:847–855
- Gryseels B, Polman K, Clerinx J, Kestens L (2006) Human schistosomiasis. Lancet 368:1106–1118
- Harn DA, Gu W, Oligino LD, Mitsuyama M, Gebremichael A, Richter D (1992) A protective monoclonal antibody specifically recog-

nizes and alters the catalytic activity of schistosome triosephosphate isomerase. J Immunol 148:562-567

- Haste-Andersen P, Nielsen M, Lund O (2006) Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Sci 15:2558–67
- Hooker CW, Brindley PJ (1996) Cloning and characterisation of strain-specific transcripts encoding triosephosphate isomerase, a candidate vaccine antigen from *Schistosoma japonicum*. Mol Biochem Parasitol 82:265–269
- Ishizuka H, Maezawa T, Kawauchi J, Nodono H, Hirao Y, Nishimura O, Nakagawa H, Sekii K, Tasaka K, Tarui H, Agata K, Hoshil M, Kobayashi K, Sakakibara Y, Matsumoto M (2007) The *Dugesia ryukyuensis* database as a molecular resource for studying switching of the reproductive system. Zoolog Sci 24:31–37
- Jimenez L, Vibanco-Perez N, Navarro L, Landa A (2000) Cloning, expression and characterisation of a recombinant triosephosphate isomerase from *Taenia solium*. Int J Parasitol 30:1007–1012
- Kawakatsu M, Oki I, Tamura S (1995) Taxonomy and geographical distribution of *Dugesia japonica* and *D. ryukyuensis* in the Far East. Hydrobiologia 305:55–61
- King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I, Marr M, Pincus D, Putnam N, Rokas A, Wright KJ, Zuzow R, Dirks W, Good M, Goodstein D, Lemons D, Li WQ, Lyons JB, Morris A, Nichols S, Richter DJ, Salamov A, Bork P, Lim WA, Manning G, Miller WT, McGinnis W, Shapiro H, Tjian R, Grigoriev IV, Rokhsar D, Sequencing JGI (2008) The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. Nature 451:783–788
- Larsen JEP, Lund O, Nielsen M (2006) Improved method for predicting linear B-cell epitopes. Immunome Res 2:2
- Liu CY, Jiang H (2005) Improved CTAB Method Suitable for Extracting Genomic DNA from Planaria [J]. Amino Acids & Biotic Resources 27:32–35
- Lou PA, Zhu YC, Yu CX, Yin XR, Hua WQ, He W, Liu YJ (2001) Cloning, expression and characterization of the specific recom-

binant peptide of triose-phosphate isomerase of *Schistosoma japonicum* Chinese strain [J]. Chinese Journal of Schistosomiasis Control 13:9–13

- Lund O, Nielsen M, Lundegaard C, Worning P (2002) CPHmodels 2.0: X3M a Computer Program to Extract 3D Models. Abstract at the CASP5 conference A102
- Mineta K, Nakazawa M, Cebria F, Ikeo K, Agata K, Gojobori T (2003) Origin and evolutionary process of the CNS elucidated by comparative genomics analysis of planarian ESTs. Proc Natl Acad Sci U S A 100:7666–7671
- Ribeiro-dos-Santos G, Verjovski-Almeida S, Leite LCC (2006) Schistosomiasis–a century searching for chemotherapeutic drugs. Parasitol Res 99:505–521
- Ruiz-Trillo I, Riutort M, Littlewood DTJ, Herniou EA, Baguna J (1999) Acoel flatworms: Earliest extant bilaterian metazoans, not members of Platyhelminthes. Science 283:1919–1923
- Shoemaker C, Gross A, Gebremichael A, Harn D (1992) cDNA cloning and functional expression of the *Schistosoma mansoni* protective antigen triose-phosphate isomerase. Proc Natl Acad Sci U S A 89:1842–1846
- Velanker SS, Ray SS, Gokhale RS, Suma S, Balaram H, Balaram P, Murthy MRN (1997) Triosephosphate isomerase from *Plasmodium falciparum*: The crystal structure provides insights into antimalarial drug design. Structure 5:751–761
- Wei J, Xu MQ, He GS, Yao BA (2006) Cloning and sequence analysis of triosephosphate isomerase gene of Orientobilharzia turkestanicum. Journal of Pathogen Biology 1:27–31
- Wu ZD, Lu ZY, Yu XB (2005) Development of a vaccine against Schistosoma japonicum in China: a review. Acta Trop 96:106– 116
- Zayas RM, Hernandez A, Habermann B, Wang YY, Stary JM, Newmark PA (2005) The planarian *Schmidtea mediterranea* as a model for epigenetic germ cell specification: Analysis of ESTs from the hermaphroditic strain. Proc Natl Acad Sci U S A 102:18491–18496