

RESEARCH ARTICLE

Open Access



Fast evolutionary rates associated with functional loss in class I glucose transporters of *Schistosoma mansoni*

Alejandro Cabezas-Cruz¹, James J. Valdés², Julien Lancelot¹ and Raymond J. Pierce^{1*}

Abstract

Background: The trematode parasite, *Schistosoma mansoni*, has evolved to switch from oxidative phosphorylation to glycolysis in the presence of glucose immediately after invading the human host. This metabolic switch is dependent on extracellular glucose concentration. Four glucose transporters are encoded in the genome of *S. mansoni*, however, only two were shown to facilitate glucose diffusion.

Results: By modeling the phase of human host infection, we showed that transporter transcript expression profiles of recently transformed schistosomula have two opposing responses to increased glucose concentrations. Concurring with the transcription profiles, our phylogenetic analyses revealed that *S. mansoni* glucose transporters belong to two separate clusters, one associated with class I glucose transporters from vertebrates and insects, and the other specific to parasitic Platyhelminthes. To study the evolutionary paths of both groups and their functional implications, we determined evolutionary rates, relative divergence times, genomic organization and performed structural analyses with the protein sequences. We finally used the modelled structures of the *S. mansoni* glucose transporters to biophysically (i) analyze the dynamics of key residues during glucose binding, (ii) test glucose stability within the active site, and (iii) demonstrate glucose diffusion. The two *S. mansoni* Platyhelminthes-specific glucose transporters, which seem to be younger than the other two, exhibit slower rates of molecular evolution, are encoded by intron-poor genes, and transport glucose. Interestingly, our molecular dynamic analyses suggest that *S. mansoni* class I glucose transporters are not able to transport glucose.

Conclusions: The glucose transporter family in *S. mansoni* exhibit different evolutionary histories. Our results suggested that *S. mansoni* class I glucose transporters lost their capacity to transport glucose and that this function evolved independently in the Platyhelminthes-specific glucose transporters. Finally, taking into account the differences in the dynamics of glucose transport of the Platyhelminthes-specific transporters of *S. mansoni* compared to that of humans, we conclude that *S. mansoni* glucose transporters may be targets for rationally designed drugs against schistosomiasis.

Keywords: *Schistosoma mansoni*, Glucose transporters, Transcriptional regulation, Phylogeny, Biophysics

Background

Upon contact with host mammalian skin, the free-living cercaria of *Schistosoma mansoni* undergoes dramatic biological [1] and metabolic transformations [2]. The metabolic transformation is a switch from an oxidative metabolism to a glycolytic metabolism that is reversibly

dependent on external glucose concentrations [2]. In the mammalian host, glucose transport is facilitated by diffusion through the *S. mansoni* tegument [3]. Two closely related (61 % identity) schistosome glucose transporter proteins (SGTPs) have been functionally characterized in *S. mansoni*, named SGTP1 and SGTP4 [4, 5]. Both show stereospecificity for glucose, relaxed specificity for different hexoses and sodium-independent activity [4]. Differential protein expression profiles show that SGTP4 is expressed in the host-interactive outer tegument [6], while SGTP1 is mainly located in the basal membranes

* Correspondence: raymond.pierce@pasteur-lille.fr

Alejandro Cabezas-Cruz and James J. Valdés joint first authorship.

¹Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 8204 - CIL - Centre d'Infection et d'Immunité de Lille, F-59000 Lille, France
Full list of author information is available at the end of the article

of the tegument [7]. Furthermore, Western blot analysis showed that SGTP1 is expressed in the egg, sporocyst, cercaria, schistosomula, and in adult male and female worms, but SGTP4 is only detected in mammalian-stage parasites (schistosomula and adults) [6]. Studies using RNAi showed that both SGTP1 and SGTP4 were critical for exogenous glucose uptake in schistosomula *in vitro* and for parasite development *in vivo* [5].

A third glucose transporter gene, *sgtp2*, has also been identified. The *sgtp2* open reading frame was interrupted by a missing base ~260 bp downstream the first initiator ATG codon. The hypothetical addition of a base at this downstream position resulted in a translation product homologous to glucose transporters [4]. Functional assays using *Xenopus* oocytes, however, showed that this hypothetical SGTP2 does not transport glucose (at 1 mM and 10 mM of substrate). In addition, transcripts of the *sgtp2* gene were only found in adult females [4], possibly due to a specific biological function at this stage. Although expressed sequence tags exist for all *S. mansoni* *sgtp* genes [5, 8], a fourth glucose transporter, *sgtp3*, is rarely mentioned in the literature.

The human glucose transporter protein family is divided into three classes: class I comprises GLUT1 to GLUT4 and GLUT14; class II comprises GLUT5, GLUT7, GLUT9, and GLUT11; and, class III comprises GLUT6, GLUT8, GLUT10 and GLUT12 (Augustin [9]). These three human glucose transporter classes are phylogenetically differentiated and show distinct molecular properties [9]. Among the human glucose transporters, GLUT1 and GLUT4 have been well studied and are transcriptionally upregulated under glucose deprivation conditions [10]. In terms of primary amino acid sequence, SGTP1 and SGTP4 were reported to be homologous to human GLUT1 [11], while SGTP2 is more similar to GLUT4 [4].

Glucose transporters undergo conformational shifts during glucose diffusion [12]. The resolved crystal structure of the proton symporter, Xyle, of *Escherichia coli* has a glucose-bound, occluded structural conformation towards the intracellular compartment [13]. This occluded conformation disallows glucose diffusion into the cell. The crystal structure of human GLUT1, possesses a glucose-bound, open-inward structural conformation permitting glucose diffusion within the cell [12]. Comparing these two homologous transporter structures, Deng et al. [12] were able to hypothesize four intermediate conformational shifts during glucose binding and diffusion. Xyle and GLUT1 represent the two respective, intermediate and sequential, glucose-bound conformations. The other two are glucose-free that represent conformations before glucose binding and after glucose diffusion [12].

In this study, the evolution of *S. mansoni* glucose transporters is explored and the effects of glucose-induced

transcriptional regulation of the encoded genes were determined. Our findings led us to investigate glucose stability and migration with *S. mansoni* glucose transporters using a publicly available, state-of-the-art algorithm. These results provide insights into the molecular properties, evolution and biophysics of glucose transport in trematodes. Unveiling the structural and dynamic differences in glucose transport by parasitic worms can establish the basis for the rational design of schistosome-specific glucose transporter inhibitors.

Results and discussion

Glucose induces transcriptional changes in *S. mansoni* glucose transporter genes in schistosomula larvae

Regulation of nutrient transporters by nutrient availability is a well-known phenomenon in microorganisms, such as yeast [14] and bacteria [15]. The regulation of mammalian transporters by their substrates, however, is less understood [10]. Nevertheless, the effect of glucose on the transcriptional regulation of glucose transporters has been studied both *in vitro* and *in vivo* [10]. From these studies it is clear that mRNA levels of class I glucose transporters *glut1* and *glut4*, the most extensively studied in this regard, change in response to the glucose concentration in the medium. As a consensus, *glut1* and *glut4* mRNA levels are higher under glucose deprivation conditions than in the presence of glucose [10]. Despite extensive characterization of SGTP1 and SGTP4 from *S. mansoni* [5], no study on transcriptional regulation in response to glucose has been conducted. In addition, knowledge of the properties of SGTP2 and SGTP3 is limited.

Upon infection, changes in glucose concentration have been shown to be essential for *S. mansoni* [2]. Here we used an infection relevant model to evaluate the transcriptional regulation of the four glucose transporter genes from *S. mansoni* in presence of glucose. All glucose transporters were expressed in schistosomula 3 h after mechanical transformation in the presence of 0.05 mM glucose. A significant reduction in expression was observed after 8 h incubation in glucose deprivation conditions (p -value <0.01). However, when the medium was supplemented with 10 mM glucose we observed a significant change in expression pattern. While the transcriptional levels of *sgtp1* and *sgtp4* were downregulated (p -value <0.025), the transcriptional levels of *sgtp2* and *sgtp3* were upregulated (p -value <0.025) in the presence of 10 mM glucose (Fig. 1). The downregulation of *sgtp1* and *sgtp4* transcripts was specific to glucose, as the addition of galactose or maltose to the medium resulted in a significant increase (p -value <0.025) in the transcriptional levels of both genes (Fig. 1).

Our results complement those obtained by Krautz-Peterson et al. [5] since they showed that RNA

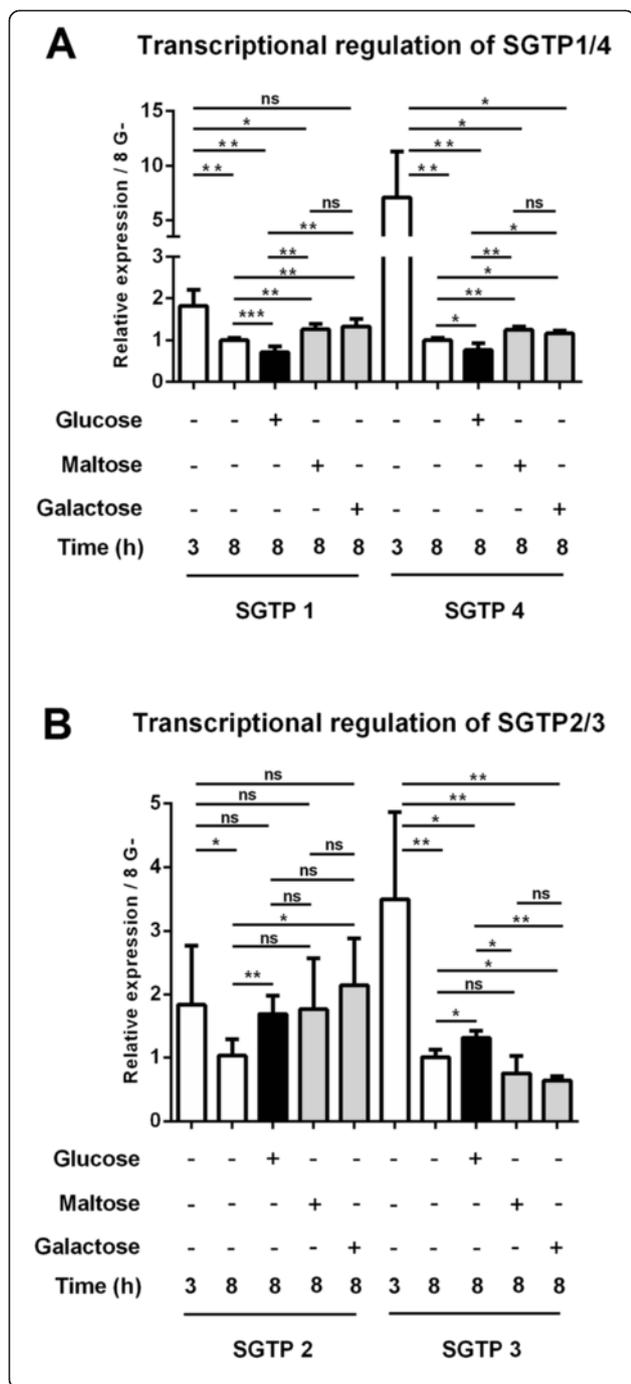


Fig. 1 Relative expression of *sgtp1*, *sgtp2*, *sgtp3* and *sgtp4* after transformation and adding glucose. The relative expression of *sgtp1*, *sgtp4* (Panel a), *sgtp2* and *sgtp3* (Panel b) in schistosomula was determined using quantitative RT-PCR. Cercariae were mechanically transformed into schistosomula. A total of 10 000 schistosomula were incubated for 3 h with traces of glucose (0.05 mM). Subsequently, glucose, maltose and galactose were added to a final concentration of 10 mM. After 8 h of incubation, total RNA was extracted, cDNA was synthesized and quantitative RT-PCR was performed. Transcripts of *sgtp1* (*p*-value = 0.0006) and *sgtp4* (*p*-value = 0.02) were significantly downregulated after adding 10 mM glucose, while *sgtp2* (*p*-value = 0.002) and *sgtp3* (*p*-value = 0.02) were significantly upregulated after adding 10 mM glucose. Results shown are the means of three independent experiments. Gene expression data is presented relative to 8 h with 0.05 mM of glucose which was used as "calibrator sample"

interference (RNAi) of both SGTP1 and SGTP4 more strongly affected schistosomula when they were maintained in glucose-poor medium than in glucose-rich medium. In glucose deprivation conditions the parasite may increase the expression of these two glucose transporters to maximize the uptake of glucose. Thus, inhibiting the glucose transporters in glucose deprivation conditions will be more detrimental to the parasite than in the presence of glucose. The transcriptional expression levels of *sgtp2* significantly increased in the presence of galactose (*p*-value = 0.010), but not maltose (*p*-value = 0.163). In contrast, the transcriptional expression levels of *sgtp3* significantly decreased in the presence of galactose (*p*-value = 0.004), but not maltose (*p*-value = 0.125). These observations suggest that the transcriptional regulation of *sgtp2* and *sgtp3* may not be glucose-specific.

S. mansoni glucose transporters cluster separately and evolve at different rates

Maximum likelihood (ML), maximum parsimony (MP) and neighbor joining (NJ) analyses were performed to phylogenetically characterize the *S. mansoni* glucose transporters SGTP1, SGTP2, SGTP3 and SGTP4. Amino acid sequences from glucose and trehalose transporters from the taxonomic classes Insecta, Mammalia, Teleostei, Cestoda and Trematoda were included in the phylogenetic analyses. All phylogenetic trees had the same topology with each of the above methods. Therefore, only one topology (ML) is displayed (Fig. 2).

Three classes of glucose transporters have been described in humans that form different phylogenetic clusters [9]. We found that *S. mansoni* SGTP2 and SGTP3 clustered together with vertebrate and *D. melanogaster* class I glucose transporters (Fig. 2, roman numeral I), while SGTP1 and SGTP4 formed a Platyhelminthes-specific glucose transporter cluster. The phylogeny presented in Fig. 2 is consistent with the monophyly of both Cestoda and Digenea [16] since the two clusters

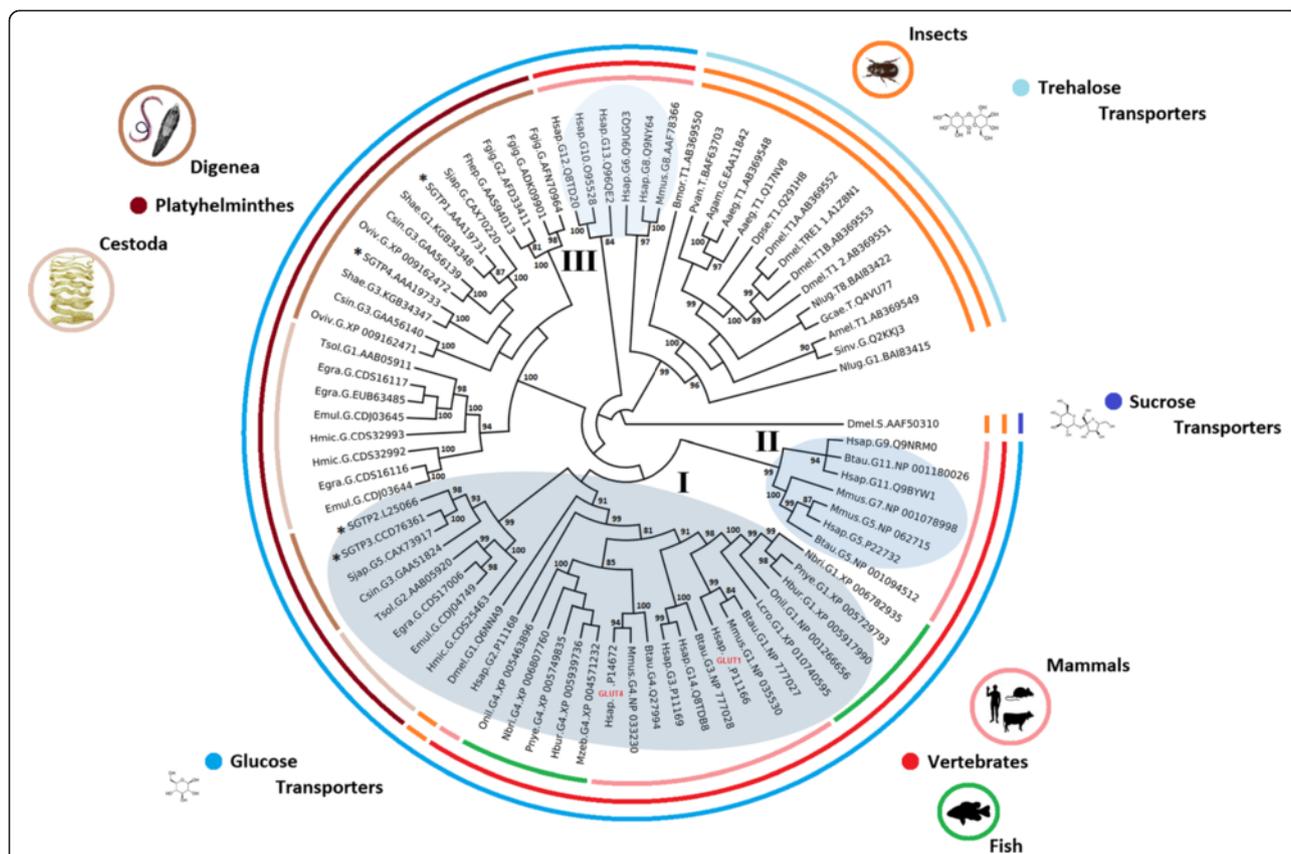
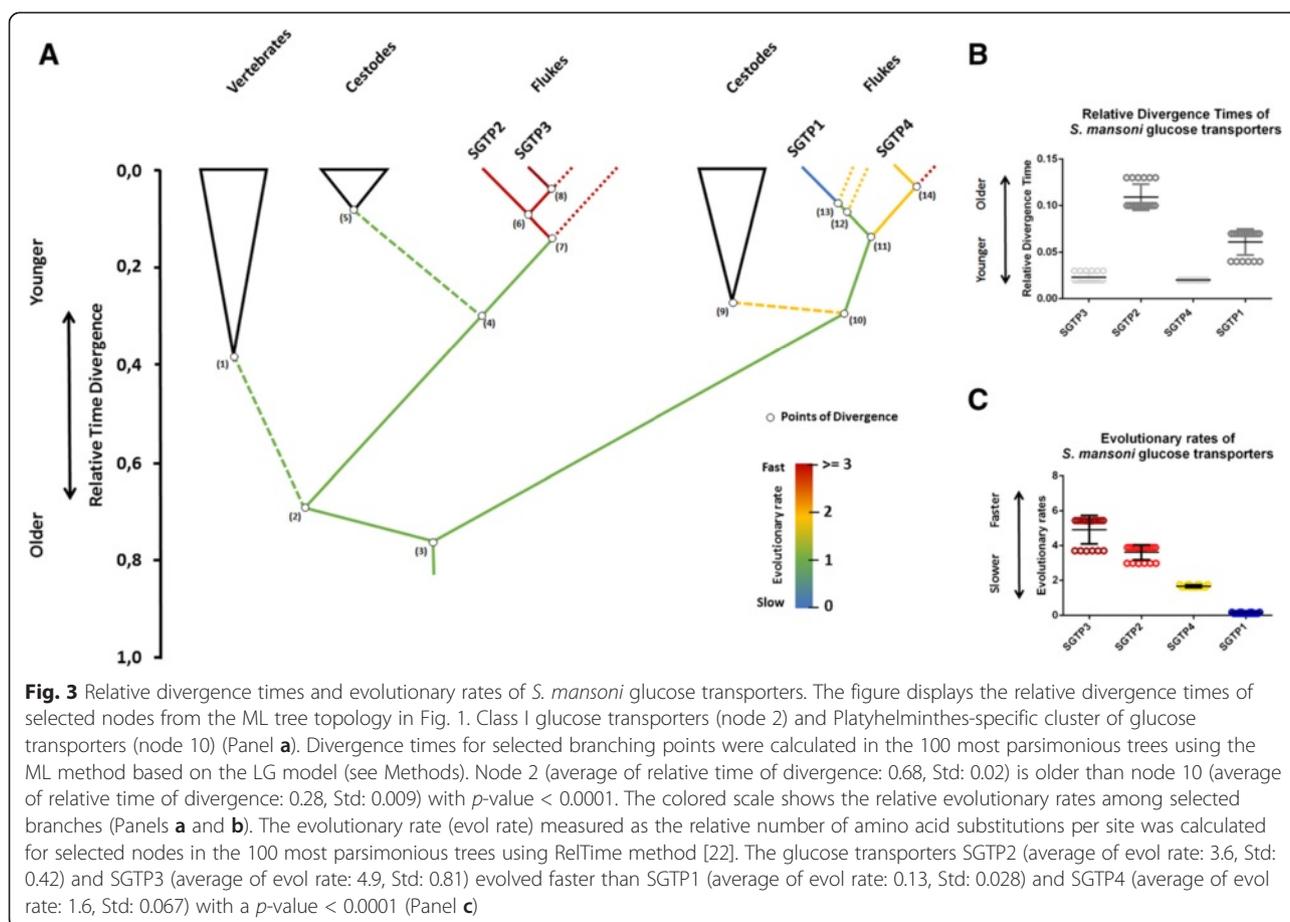


Fig. 2 Phylogeny of the trehalose and glucose transporter family. The figure shows a phylogenetic tree of trehalose and glucose transporters from vertebrates, trematodes and insects. The topologies obtained with ML, MP and NJ methods were similar, and thus only ML is displayed. Numbers on internal branches are the bootstrap values (only >80 % are shown); the statistical support obtained with approximate likelihood ratio test (aLRT) was similar. Sequences are labeled with data collected from GenBank as name of species (four first letters), type of transporters (T: trehalose, G: glucose), number (if reported) of the sugar transporter and accession numbers. *S. mansoni* glucose transporters are indicated by asterisks. *Homo sapiens* GLUT1 and GLUT4 are highlighted in red. The classes of glucose transporters are numbered with roman numerals. *D. melanogaster* sucrose transporter was used as the outgroup. The labels for the species used are as follow: Mammals: *H. sapiens* (Hsap), *Mus musculus* (Mmus), *Bos Taurus* (Btau); Fish: *Oreochromis niloticus* (Onil), *Neolamprologus brichardi* (Nbr), *Haplochromis burtoni* (Hbur), *Pundamilia nyererei* (Pnye), *Larimichthys crocea* (Lcro), *Dicentrarchus labrax* (Dlab), *Maylandia zebra* (Mzeb); Insects: *Anopheles gambiae* (Agam), *Aedes aegypti* (Aeag), *Apis mellifera* (Amel), *Bombyx mori* (Bmor), *Drosophila melanogaster* (Dmel), *Drosophila pseudoobscura pseudoobscura* (Dpse), *Gastric caeca* (Gcae), *Solenopsis invicta* (Sinv); Cestodes: *Taenia solium* (Tsol), *Echinococcus granulosus* (Egra), *Echinococcus multilocularis* (Emul), *Hymenolepis microstoma* (Hmic), *Fasciola gigantica* (Fgig), *Fasciola hepatica* (Fhep); Flukes: *Clonorchis sinensis* (Csin), *S. haematobium* (Shae), *Opisthorchis viverrini* (Oviv), *S. japonicum* (Sjap)

containing glucose transporters from Platyhelminthes, cestodes and trematodes clustered separately (Fig. 2). The phylogenetic analyses allowed us to classify SGTP2 and SGTP3 together with some glucose transporters from the flukes *S. japonicum*, *Clonorchis sinensis* and the cestodes *Taenia solium*, *Echinococcus granulosus*, *E. multilocularis* and *Hymenolepis microstoma* as class I glucose transporters. In particular, we found homologs of SGTP1, SGTP2, SGTP3 and SGTP4 in both *S. japonicum* and *S. haematobium* (Additional file 1).

Mammalia, Platyhelminthes and Insecta shared a common ancestor circa 695 MYA [17]. The fact that class I glucose transporters included sequences from Platyhelminthes, insects and mammals suggested that this

cluster is ancestral to the Platyhelminthes-specific cluster. To test this hypothesis we determined the relative divergence times for selected nodes in the 100 most parsimonious phylogenetic trees (see Methods). Indeed, the results from this analysis showed that the cluster containing class I glucose transporters is older than the Platyhelminthes-specific cluster (Fig. 3a, node 2 vs. node 10) with *p-value* <0.0001. In agreement with this, SGTP2 seems to be the oldest glucose transporter in *S. mansoni* (Fig. 3b, *p-value* <0.0001). The relative divergence time of SGTP3 and SGTP4 suggested that these phylogenetic events coincide in time, and they are recently evolved glucose transporters compared to SGTP2 and SGTP1. The phylogenetic position of SGTP3 and SGTP4 (Fig. 2),



and the relative divergence times (Fig. 3b) of these genes support the idea that they originated from SGTP2 and SGTP1 respectively, possibly due to gene duplication.

Gene duplications were reported in glucose transporters from *Homo sapiens* and *Saccharomyces cerevisiae* [18, 19]. After gene duplication, relaxation of selective constraints on duplicated genes has been observed, resulting in copies that evolve faster than their original orthologs [20, 21]. To test variations in evolutionary rates, we estimated these along branches of the glucose transporters phylogeny using the RelTime method [22]. Figure 3c shows that SGTP2 and SGTP3 evolved faster than SGTP1 and SGTP4 (p -value = 0.0001). In addition, SGTP3 and SGTP4 were found to evolve faster than SGTP2 and SGTP1, respectively with p -value < 0.0001.

We further compared the evolutionary rates of SGTP3/SGTP2 and SGTP4/SGTP1 using Tajima's relative rate test [23]. The χ^2 test statistic was 6.50 (p -value = 0.01079) for SGTP4/SGTP1 and 0 (p -value = 1.0000) for SGTP3/SGTP2. Thus, the null hypothesis of equal rates was rejected for SGTP4/SGTP1, but not for SGTP3/SGTP2. This suggests that Tajima's relative rate test may not be suited for distinguishing different rates of evolution in

fast-evolving sequences like SGTP3/SGTP2 or simply that after divergence SGTP3 and SGTP2 evolved at the same rates. In agreement with the latter, other sequences from *S. japonicum* and *C. sinensis*, in the cluster where SGTP2 and SGTP3 are located, are also fast-evolving sequences (Fig. 3). Interestingly, the transcriptional regulation of SGTP1 and SGTP4 compared with SGTP2 and SGTP3 in the presence of glucose (Fig. 1) is in agreement with these phylogenetic and evolutionary observations.

The rapidly-evolving glucose transporters SGTP2 and SGTP3 are encoded by intron-rich genes

The cDNA sequence encoding SGTP2 originally described by Skelly et al. [4] was constructed from a cloned sequence by the insertion of a single nucleotide in order to restore the open reading frame. The authors speculated that *sgtp2* might therefore be an expressed pseudogene and not encoded as a functional protein. Comparison of the *sgtp2* (L25066) cDNA sequence to the genome sequence allowed the reconstruction of a "genomic" predicted cDNA sequence. Three base changes were found within the encoding region, one leading to an amino acid change (Asn88Lys). Since the cDNA sequence encodes a full-length protein reconstructed by assembling the exons

of the genomic sequence, we speculated that the missing nucleotide in the cDNA clone obtained by Skelly et al. [4] might represent a cloning artifact. We therefore amplified the full-length coding sequence of SGTP2 from cDNA derived from adult worms and miracidia and sequenced the clones obtained. However, the examination of twenty such sequences failed to identify a single clone corresponding to the predicted genomic sequence (not shown). A number of these contain unspliced intronic sequences, whilst others contained short indels. None of the clones encoded a full-length SGTP2 protein sequence. We are therefore unable to verify our speculation concerning the status of the *sgtp2* gene as a pseudogene or a viable coding gene. One possibility is that the functional transcript might only be produced at specific stages during the life cycle. To compare the *sgtp2* gene organization and the potential structure and function of the SGTP2 protein, we used the transcript and protein sequence predicted by the genome. In contrast we confirmed that the sequence of SGTP3 predicted from the genome corresponds to viable transcripts present in adult worms. The transcript sequences were found to be identical to the genome prediction, but have a small deletion at the N-terminal end and to encode

a further 86 amino acids at the C-terminal end. However, these sequence modifications have no effect on the core glucose transporter structure. The transcript sequence was submitted to GenBank [GenBank: KT895372].

The schistosome genome sequence and the derived exon/intron structure shown in Fig. 4 and Table 1 are aligned with published transcript sequences of SGTP1 and SGTP4, the transcript sequence of SGTP2 predicted from the genome, and the predicted transcript sequence of SGTP3. We analyzed the gene structure of human *glut1*, and collected that of *glut4* from the literature [24], for comparison, since our phylogenetic analyses placed the human GLUT1 and GLUT4 within the same clade as SGTP2 and SGTP3. The analysis shows that the gene structures of *sgtp1* and *sgtp4* are similar to each other, but different from those of *sgtp2* and *sgtp3*, which are also similar to each other. Both *sgtp1* and *sgtp4* have 5 introns, four of which are large (>1 000 bp), the size of the fifth was not evaluated due to missing sequence data. In contrast, *sgtp2* has 10 introns and *sgtp3* has 7 introns. Apart from the first intron of *sgtp2*, which is large and not conserved in *sgtp3*, the first introns of these two genes are small, between 31 bp and 41 bp, whereas the

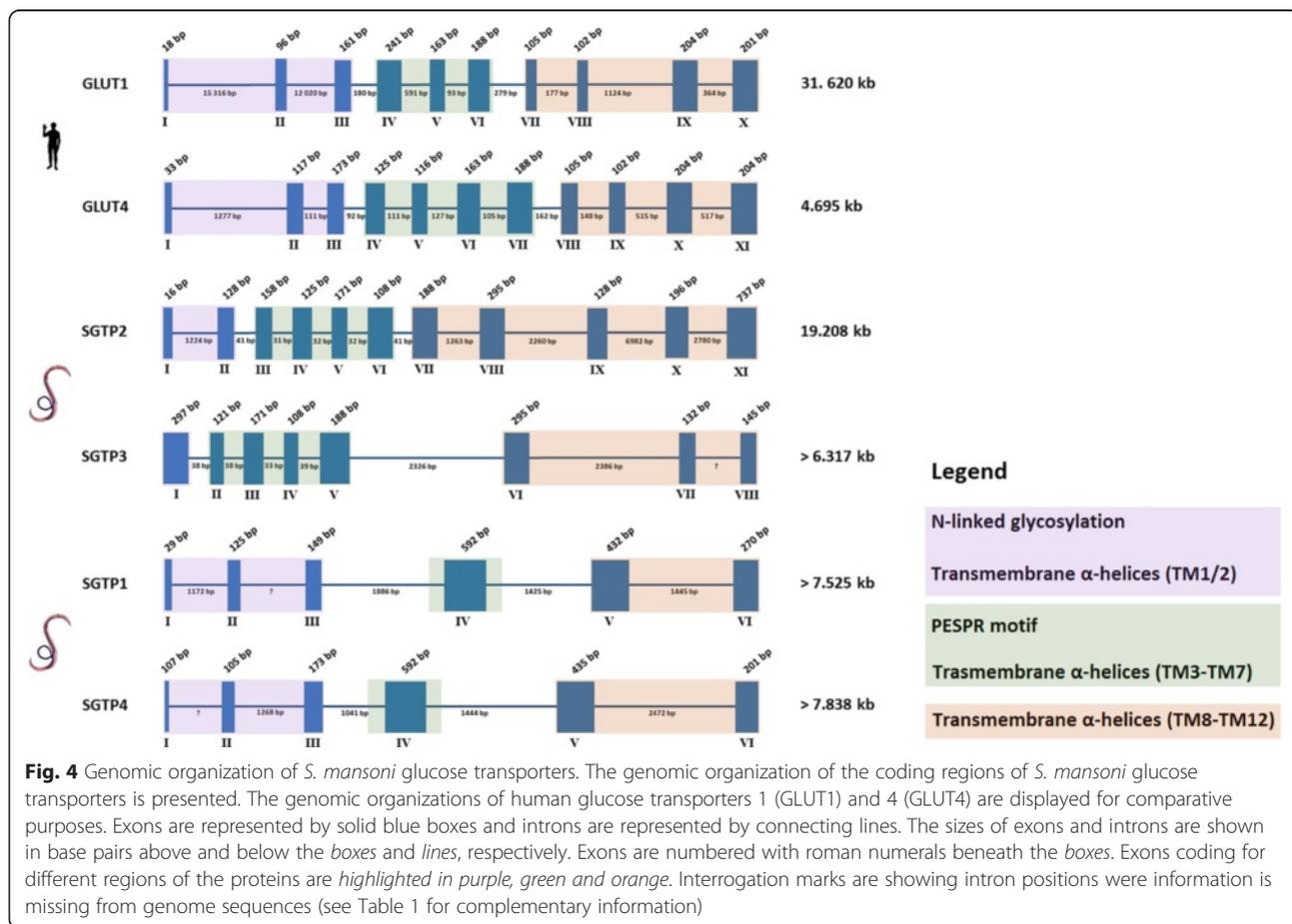


Table 1 Genomic structure and organization of *S. mansoni* glucose transporters

| <i>glut1</i> | | | | <i>sgtp1</i> | | | | <i>sgtp4</i> | | | | <i>sgtp2</i> | | | | <i>sgtp3</i> | | | |
|--------------|------------------------|----------------|-------|--------------|------------|---------|------|--------------|------------|---------|------|----------------|------------|---------|------|--------------|------------|---------|------|
| Exons | Size | Introns | Size | Exons | Size | Introns | Size | Exons | Size | Introns | Size | Exons | Size | Introns | Size | Exons | Size | Introns | Size |
| 1 | 18¹ | 1 | 15316 | 1 | 29 | 1 | 1172 | 1 | 107 | 1 | n.e. | 1 | 16 | 1 | 1224 | - | - | | |
| 2 | 96 | 2 | 12020 | 2 | 125 | 2 | n.e. | 2 | 105 | 2 | 1268 | 2 | 128 | 2 | 41 | - | - | | |
| 3 | 161² | 3 | 180 | 3 | 149 | | | 3 | 173 | | | 3 | 158 | 3 | 31 | 1 | 297 | 1 | 38 |
| 4 | 241 | | | | | | | | | | | 4 ⁴ | 125 | | | 2 | 121 | | |
| 5 | | 4 ³ | 591 | | | | | | | | | 5 | 171 | 4 | 32 | 3 | 171 | 2 | 38 |
| 6 | 188 | 5 | 93 | | | 3 | 1886 | | | 3 | 1041 | 6 | 108 | 5 | 32 | 4 | 108 | 3 | 33 |
| 7 | 105 | 6 | 279 | 4 | 592 | | | 4 | 592 | | | 7 | 188 | 6 | 41 | 5 | 188 | 4 | 39 |
| 8 | 102 | 7 | 177 | | | | | | | | | | | | | | | | |
| | | 8 | 1124 | | | 4 | 1425 | | | 4 | 1444 | 8 | 295 | 7 | 1263 | | | 5 | 2326 |
| 9 | 204 | | | 5 | 432 | | | 5 | 435 | | | 9 | 128 | 8 | 2260 | 6 | 295 | 6 | 2386 |
| | | 9 | 364 | | | 5 | 1445 | | | 5 | 2472 | 10 | 196 | 9 | 6982 | 7 | 132 | | |
| 10 | 201 | | | 6 | 1599 | | | 6 | 820 | | | 10 | 196 | 10 | 2780 | | | 7 | n.e. |
| | | | | | | | | | | | | 11 | 737 | | | 8 | 145 | | |

Exon and intron sizes are in base pairs

In bold: Exons for which the exon-intron junction at the 3' end is conserved between different glucose transporter genes

In blue: Exons for which the exon-intron junction at the 3' end is conserved among all selected glucose transporter genes

In red: Exons for which the exon-intron junction at the 3' end is conserved only between human *glut1*, *sgtp2* and *sgtp3*

In green: Exons for which the exon-intron junction at the 3' end is conserved only between *sgtp2* and *sgtp3*

n.e. not evaluated (information missing from genome sequences)

introns at the 3' end of both genes are large. Moreover, the sizes of exons 4 to 9 of *sgtp2* and 2 to 7 of *sgtp3* are identical and the positions of these introns with respect to the encoded amino acid sequence of the corresponding proteins are conserved. Comparison of *sgtp1*, *sgtp4*, *sgtp2*, *sgtp3* and *glut1* genes (Table 1) shows that 3 intron positions are perfectly conserved between all, and that the position of introns 1 and 2 is conserved between all but *sgtp3*. In addition, *glut1* shares an intron position (intron 5) with *sgtp2* (intron 6) and *sgtp3* (intron 4) that is not present in *sgtp1* or *sgtp4*. Overall, this analysis underlines the close relationship between *sgtp2* and *sgtp3*, and *sgtp1* and *sgtp4* that correlates well with our phylogenetic analysis.

S. mansoni glucose transporters show typical molecular signatures

Despite SGTP1 and SGTP4 being clustered separately from class I and II glucose transporters, they share several molecular properties with these classes (Fig. 5). Human class I glucose transporters comprise GLUT1, GLUT2, GLUT3, GLUT4 and GLUT14 with GLUT1 being the first isoform cloned described by Mueckler et al. [28]. All members of GLUT family isolated from humans possess 12 transmembrane (TM) α -helices [9], structural patterns that are also present in some glucose transporters from Insecta [29]. Using the TMHMM server we predicted the presence of the 12 TM α -helices in SGTP1-4 and compared their positions with human class I glucose transporters (Fig. 5). The positions of these TMs were conserved in the four glucose transporters from *S. mansoni*, except for the lack of TM7 in SGTP3. The only remarkable difference between the amino acid sequence of SGTP3 and the other glucose

transporters at TM7 is the presence of two phenylalanine residues (Fig. 5). However, the modelling analysis (see below) showed that SGTP3 has the typical 12 TM α -helices. The discrepancy between TMHMM and modelling predictions regarding TMs in SGTP3 can only be solved through the experimental determination of the structure of this transporter. Probably due to differences in the algorithm used, the positions of our predicted TM differ slightly from those reported by Skelly et al. [4] for SGTP1 and SGTP4. We then compared the sugar transporter signatures [9, 30] of human and *S. mansoni* glucose transporters. Some glycine (Gly) residues critical for structure stabilization in the GLUT family [28] were also conserved in the TM α -helices 1, 4, 5, 7 and 10 of *S. mansoni* glucose transporters (Fig. 5). The glutamic acid (Glu) and arginine (Arg) residues implicated in conformational alteration [31] and membrane topology [32] in human GLUTs were also present in cytoplasmic loops 2, 4, 8, and 10 of *S. mansoni* glucose transporters (Fig. 5).

Some exclusive characteristics of class I/II glucose transporters in humans are conserved in *S. mansoni*. For example, the N-linked glycosylation site (Fig. 5) in the first extracellular loop between TM1 and TM2 [33] and the motif PESPR in the sixth intracellular loop between TM6 and TM7 (Fig. 5). SGTP2 and SGTP3 showed an amino acid change in the last Arg of the motif PESPR to threonine (Thr) and serine (Ser), respectively. Despite the improbability of deducing substrate specificity or transport kinetics from the primary amino acid structure of class I glucose transporters, GLUT1, 3 and 4 are known to transport glucose, and not fructose, via a QLS sequence present in TM 7 [34]. This triad sequence has been implicated in substrate binding and is conserved in *S. mansoni* SGTP1, SGTP4 and SGTP3 (Fig. 5), but not

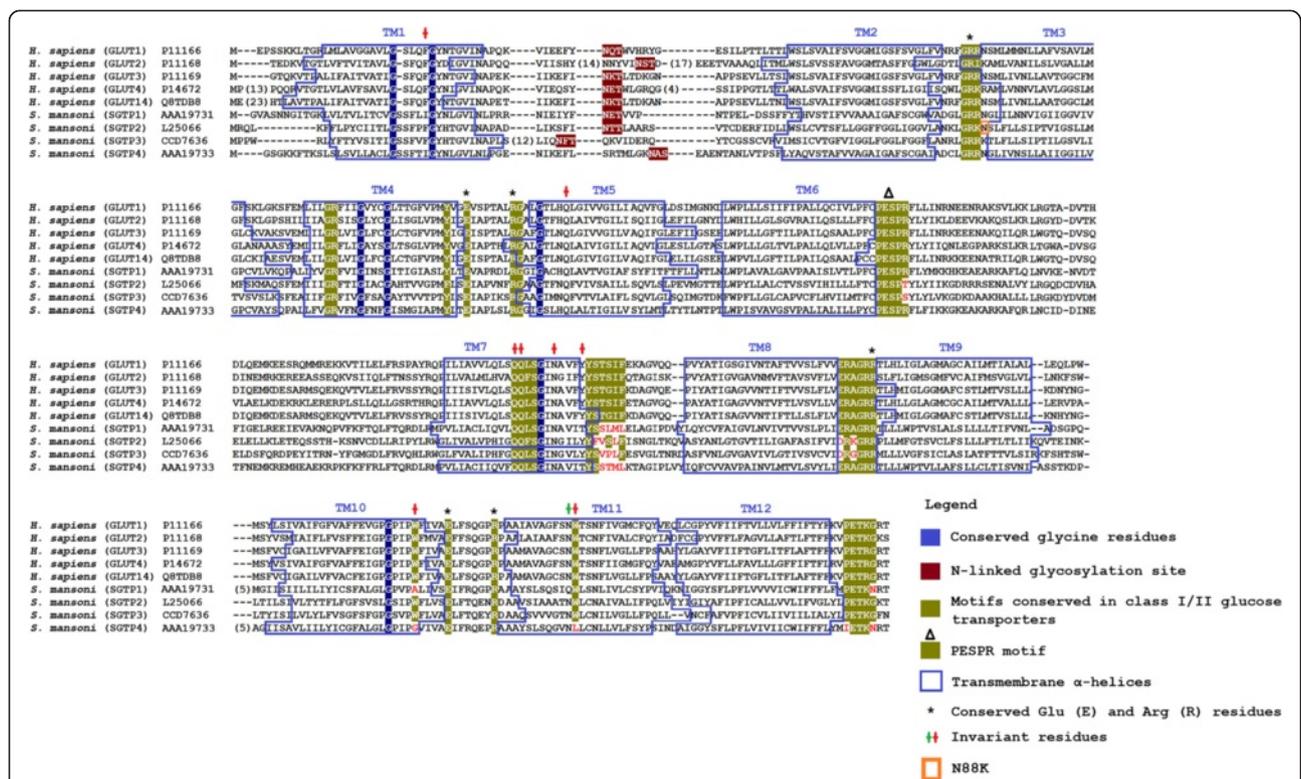


Fig. 5 Alignment of members of glucose transporters class I. Class I glucose transporters amino acid sequences from human (GLUT1-4 and GLUT-14) and *S. mansoni* (SGTP1, SGTP2, SGTP3 and SGTP4) were aligned using MAFFT. Relevant motifs among human class I/II glucose transporters and *S. mansoni* are highlighted. Annotation of relevant motifs was done using previous reports [21, 25, 26]. The 12 transmembrane (TM) topologies, predicted using TMHMM Server v. 2.0 [27], are indicated. The positions of the invariant residues of Xyle (green) and GLUT1 (red) are marked. The substitution of Asn(N)88 by Lys(K)88 found in the genome for SGTP2 is boxed (orange). The number of amino acid insertions that resulted from the alignment are shown in brackets

in SGTP2. In SGTP2 there is a substitution of leucine (Leu) by phenylalanine (Phe) and the same substitution was found in the human GLUT2. Human GLUT2 has low affinity for glucose ($K_m = 17$ mM), which is uncommon among the known members of the GLUT family [35], but has a high affinity for glucosamine ($K_m \approx 0.8$ mM) [36]. *S. mansoni* SGTP2 does not seem to transport glucose in *Xenopus* oocytes at 10 mM glucose [4], however, further studies should clarify whether this molecule shows a higher K_m for glucose or whether it has an affinity for glucosamine as does GLUT2. The presence of glucose transporters with different affinities for glucose is extremely important for the regulation of glucose uptake. For instance, under high glucose concentration conditions an increase in intracellular glucose reduces the glucose influx 50 % in yeast cells expressing high-affinity glucose transporters [37]. Thus, by regulating the expression profile of the different glucose transporters with different kinetic properties, yeasts avoid affecting their glucose uptake, which is directly dependent on extracellular glucose concentration [38]. Despite the fact that the hypothesis of glucose transporters with

different affinities for glucose in *S. mansoni* is appealing and matches the expression profile in Fig. 1, our biophysical analysis (below) suggests that neither SGTP2 nor SGTP3 transport glucose.

The predicted tertiary structural conformations of *S. mansoni* glucose transporters are homologous to GLUT1 and Xyle

Upon substrate binding, the tertiary structure of most proteins undergoes conformational changes that coordinate their function. By comparing the structural differences of Xyle and GLUT1, Deng et al. [12] were able to hypothesize on the conformational changes of glucose transporters during glucose binding and diffusion. This working model of glucose diffusion predicts that transporters prefer an outward (towards the extracellular), open, glucose-free conformation (i.e., the initial conformation). Glucose association (i.e., conformation of Xyle) and dissociation (i.e., conformation of GLUT1) trigger the migration of glucose towards an environment with a lower concentration – i.e., the intracellular compartment [12]. The final conformation, glucose-free

occluded (by the intracellular α -helices), remains to be structurally resolved.

As proof of principle for our simulations, we submitted the modelled *S. mansoni* glucose transporters to the DALI server [39] to find homologous crystal structures with the best matching conformation. The crystal structures GLUT1 of *H. sapiens* [PDB: 4PYP] and Xyle of *E. coli* [PDB: 4GBZ] were among the top candidates. This provided two opportunities to control our simulations since, (i) glucose is not the substrate of Xyle, it is a competitive inhibitor and (ii) Xyle and GLUT1 are the two sequential intermediate structural conformations during glucose binding and diffusion, respectively [12]. The conformational changes during binding occur mainly in TM α -helices 1, 4, 7 and 10 that cause the occlusion by the intracellular α -helices to open, thereby permitting diffusion [12]. Root mean square deviation (RMSD) is frequently used in structural bioinformatics to measure the average distances between atoms from two or more protein structures. The superposition in Fig. 6a graphically shows these overall structural differences – the RMSD between Xyle [PDB: 4GBZ] and GLUT1 [PDB: 4PYP] is 3.2 Å.

Figure 6b shows the predicted tertiary structures of *S. mansoni* glucose transporters SGTP1-SGTP4 and their structural similarity to GLUT1 and Xyle. Both SGTP1 (RMSD 1.0 Å) and SGTP4 (RMSD 1.3 Å) are structurally homologous to Xyle. SGTP2 (RMSD 2.3 Å) and SGTP3 (RMSD 0.6 Å) are structurally homologous to GLUT1. The higher RMSD for SGTP2 is mainly caused by differences in TM α -helices 9 and 10, intracellular α -helices and the C-terminus. These opposing conformations allow us to examine the stability of glucose within the active site of SGTP1 and SGTP4 and the inability of glucose migration (or diffusion) by SGTP2 and SGTP3.

Residue dynamics reveal how glucose affects invariant residues of glucose transporters involved in binding

The dynamics of key residues during substrate binding vary depending on the substrate and the protein. Since 1958, the key-lock theory to explain protein-substrate complexes has been revised by Koshland [40] who proposed an induced-fit process. The induced-fit theory demonstrates that during protein-substrate interactions, conformational changes occur between both the protein and substrate to accommodate the complex. These molecular interactions subsequently cause the global conformational changes aforementioned in Fig. 6. To understand how glucose will interact with key residues of glucose transporters, we performed an induced-fit docking simulation using the PELE server (see Methods).

Several conserved, invariant residues among glucose transporters (namely, GLUT1- GLUT4) were reported

for Xyle that, when mutated to alanine (Ala), reduce the affinity of substrate binding (highlighted in Fig. 5). These Xyle invariant residues are responsible for the competitive inhibitory properties of glucose [13] and showed altered conformations between Xyle and GLUT1 crystal structures. The differences in conformation between Xyle (green) and GLUT1 (red) invariant residues are clearly depicted in the structural representations of Fig. 7. Large residue conformational shifts between Xyle and GLUT1 are seen in Gln161/168 (TM5), Tyr292/298 (TM7) and Trp388/392 (TM10). Both TM7 and TM10 are involved in the two sequential intermediate structural conformations during glucose binding [12]. These homologous residues for SGTP1 and SGTP4 are, however, mutated – in TM7 as Tyr292/298Thr (SGTP1 and SGTP4), and in TM10 as Trp388/392Ala (SGTP1) and Trp388/392Gly (SGTP4) (Fig. 5). To simplify matters we refer to the invariant residue positions in the order GLUT1/Xyle, hereafter.

The residue dynamics depicted in the line graphs of Fig. 7 show that Gln161/168, Gln282-283/288-289, Asn288/294, Tyr292/298 and Trp388/392 undergo large conformational changes during induced-fit docking of glucose (compared with the dynamics of the negative control, Asn317/325 – see Fig. 7 legend). Fig. 7, however, shows that these dynamics differ depending on the glucose transporter. Xyle and GLUT1 show high perturbations for Gln161/168 and Asn288/294, but have opposing dynamics for Tyr292/298 (Xyle) and Trp388/392 (GLUT1). The *S. mansoni* glucose transporter SGTP4 showed extremely low perturbations for all binding residues, perhaps due to its lower affinity for glucose compared with SGTP1 [4]. In contrast, SGTP2 possesses the most invariant residues with high perturbations, namely for Gln161/168, Gln282-283/288-289, Asn288/294 and Trp388/392. Both SGTP1 (Gln282-283/288-289 and Asn288/294) and SGTP3 (Gln161/168 and Trp388/392) only showed high perturbations for two invariant residues (Fig. 7). There is a kink in TM7 that contributes to the coordination of the substrate in the active site of glucose transporters in the proximity of polar residues Gln282-283/288-289 and Asn288/294 and the aromatic residue, Tyr292/298 [13]. The benzene ring of Tyr292/298 may act as “trapdoor” switch for glucose since it is thought to maintain the open-inward conformation of GLUT1 by interacting with TM4 [12].

Biophysical properties of *S. mansoni* glucose transporters provide insights into glucose affinity

The induced-fit simulations revealed how invariant residues behave upon binding (since the algorithm constrains glucose migration within the active site), but what would happen if glucose were unconstrained? To answer this question, we simulated an unconstrained

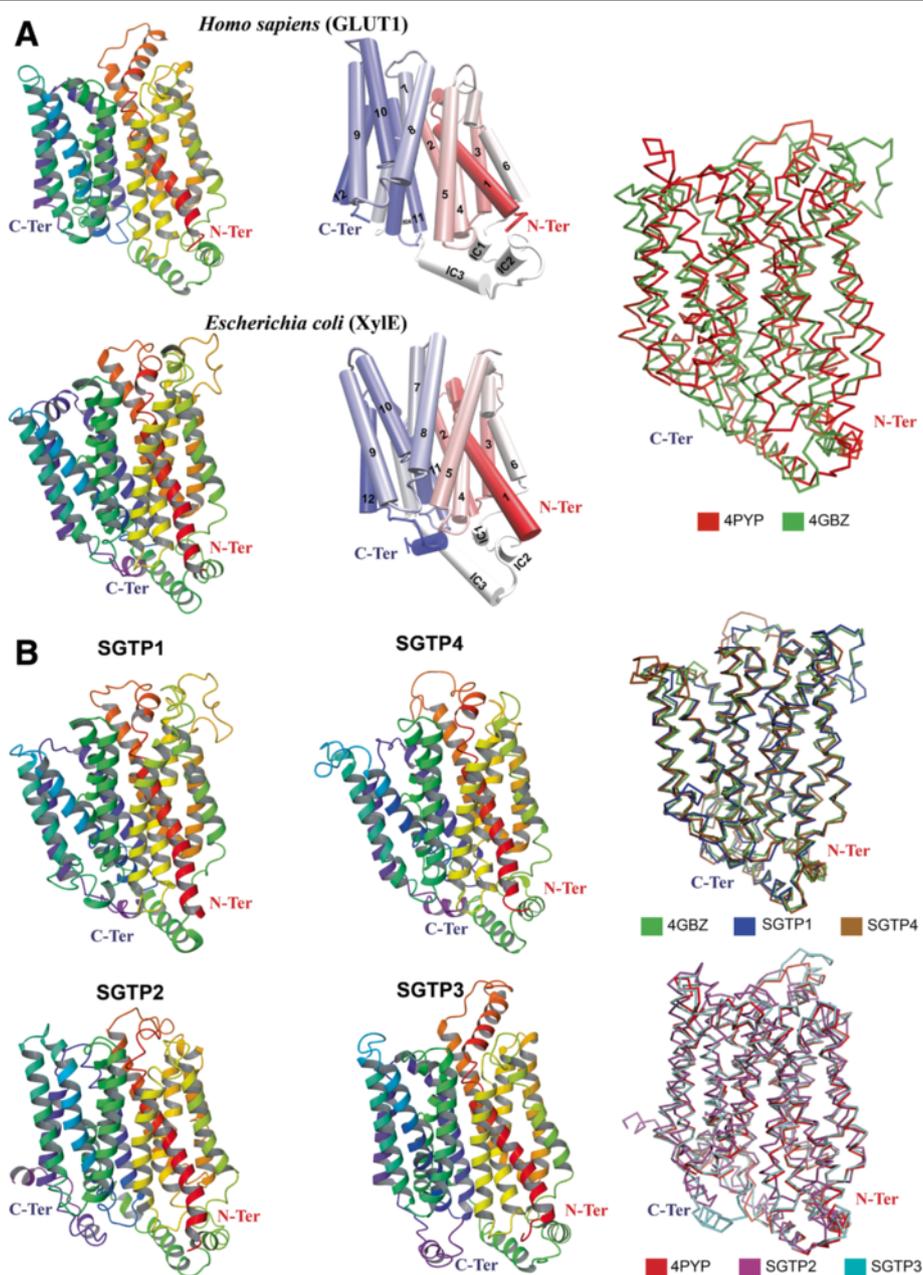
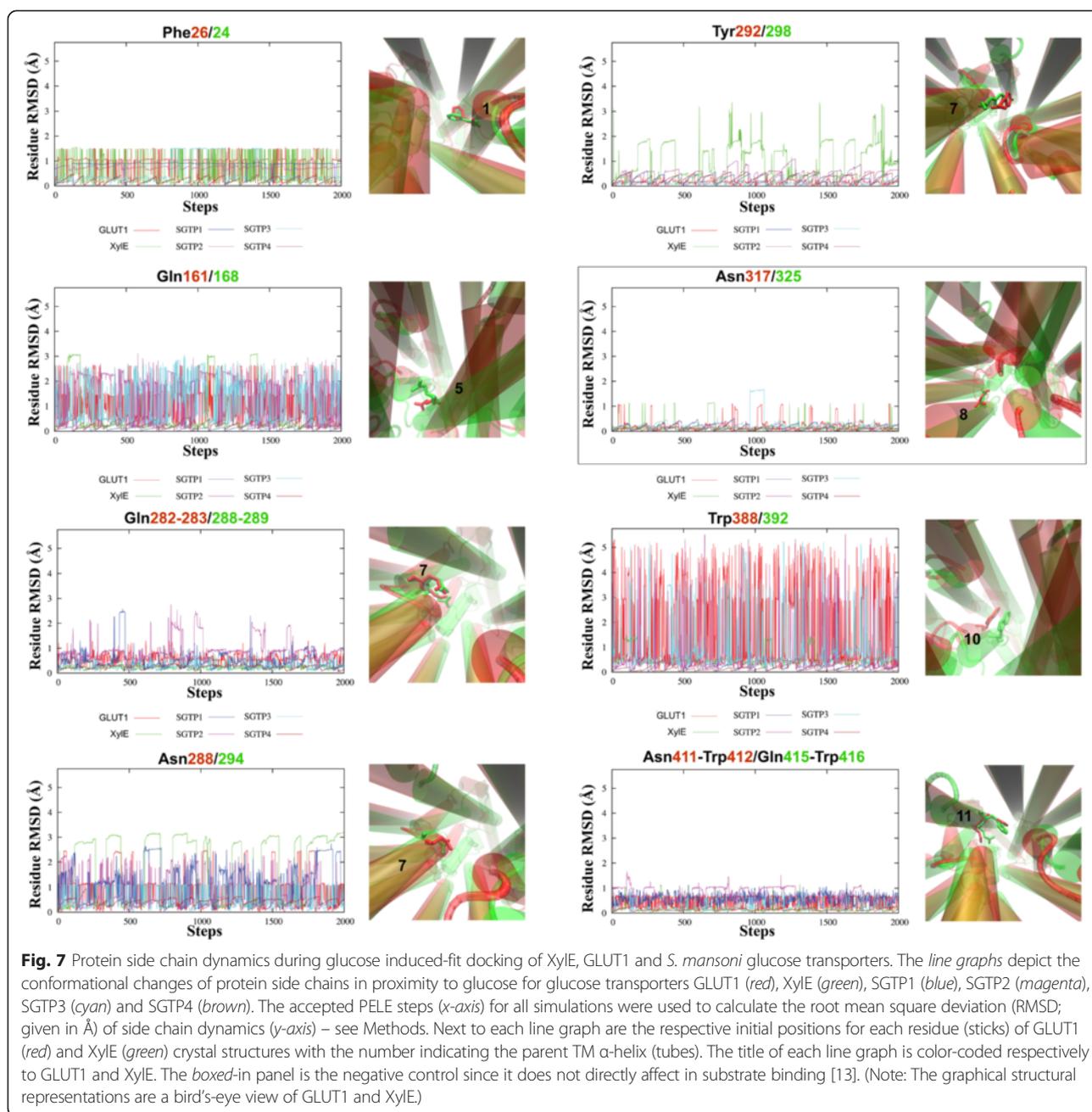


Fig. 6 Xyle and GLUT1 compared to the tertiary predicted structures of *S. mansoni* glucose transporters. Panel **a** (from left to right) shows the crystal structures of glucose transporters GLUT1 of *H. sapiens* [PDB: 4PYP] and Xyle of *E. coli* [PDB: 4GBZ]. The transmembrane (TM) α -helices (middle) are numbered (1–12) accordingly (IC = intracellular) with the α -carbon backbone structural alignment of both crystal structures (GLUT1 = red and Xyle = green; far right). Panel **b** shows the tertiary predicted structures of *S. mansoni* glucose transporters (left) with the structural alignment (right) of SGTP1 (blue) and SGTP4 (brown) with Xyle of *E. coli* (green), and SGTP2 (magenta) and SGTP3 (cyan) with GLUT1 of *H. sapiens* (red). Structures are color-coded from the amino-terminus (N-ter; red) to the carboxyl-terminus (C-ter; blue). The structural alignment of the α -carbon backbone was performed using the default tool of Schrodinger's Maestro program

glucose migration starting from the active site. Due the importance of Tyr292/298 in maintaining the openward conformation of glucose transporters [12], we used it to triangulate glucose migration from its initial position in the active site. Additional file 2 shows distinct glucose migration patterns for GLUT1 and Xyle. In

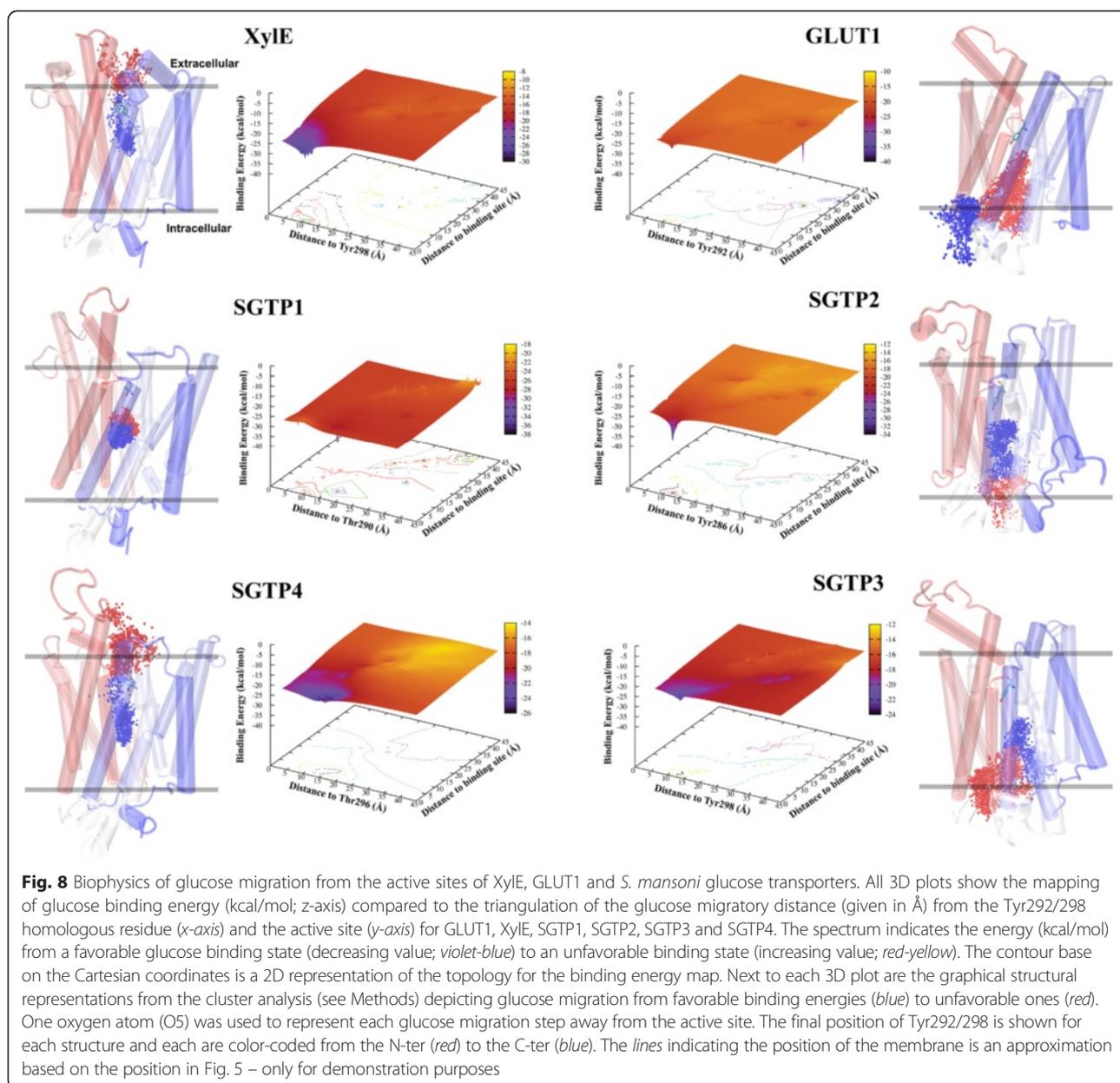
accordance with the structural similarities (Fig. 6), transcript expression (Fig. 1) and evolution (Figs. 2 and 3), *S. mansoni* glucose transporters group together in glucose migration patterns showing that in SGTP2 and SGTP3 glucose migrates similarly to GLUT1, while SGTP1 and SGTP4 are similar to Xyle. These migration patterns



seem to show glucose either slowly diffusing away from the active site (i.e., XylE and SGTP4) or being ejected from it (i.e., GLUT1, SGTP2 and SGTP3). In SGTP1, however, glucose seems to stay within 10 Å of the active site (Additional file 2). This concentrated positioning within the active site agrees with the higher affinity that SGTP1 has for glucose compared with other SGTPs [4].

In order to determine the biophysics of glucose affinity and the direction of migration patterns (towards the extracellular or intracellular compartments) we

performed two types of analyses: (i) a cluster analysis to visualize the direction, and, (ii) an incorporation of 3D energy mapping of glucose migration to localize energy favorable clusters. Figure 8 shows that XylE and GLUT1 possess conversely energy favorable glucose migration patterns. This is strictly caused by the structural conformation of these transporters since both are occluded by the extracellular TMs, but differ in the conformation of intracellular α -helices (closed for XylE and open for GLUT1). The open-inward conformation of GLUT1 is



the characteristic that permits the diffusion of glucose within the cell. The distinct migration patterns and energy signatures of *S. mansoni* glucose transporters, however, may explain their individual response to glucose (Fig. 8).

Although SGTP2 and SGTP3 are structurally similar to GLUT1, the energy signatures for both are not comparable to GLUT1. The energy signature difference between SGTP2 and SGTP3 may be due to their slight structural deviations (as mentioned above and shown in Fig. 6). Functional studies showed that SGTP2 does not transport glucose, even at 10 mM concentration [4]. Our transcriptional analysis shows that SGTP2 and SGTP3 have opposing responses to glucose compared with SGTP1 and SGTP4 (Fig. 1), and phylogenetically group

together (Fig. 2). These functional and evolutionary insights may explain glucose migration for SGTP2 and SGTP3 (Fig. 8), since it migrates from an energy favorable state (i.e., the active site) to an unfavorable energy state (i.e., towards the intracellular compartment). Our results showed that transcriptional regulation of SGTP2 and SGTP3 respond differently to maltose and galactose than do SGTP1 and SGTP4 (Fig. 1). All these results suggest that glucose is not the native substrate for SGTP2 and SGTP3, but may act as an inhibitor (as in XyleE) since both show energy favorable states within the active site (Fig. 8).

Glucose affinity is higher for SGTP1 ($K_m = 1.3$ mM) than for SGTP4 ($K_m = 2$ mM) [4] and this affinity is

clearly demonstrated in Fig. 8. For SGTP4 glucose migrates from an energy favorable state (i.e., the active site) to an unfavorable energy state (i.e., towards the extracellular compartment). In contrast, glucose spends the majority of the simulation within the active site of SGTP1 with unfavorable energy states. We do see, however, that a smaller glucose cluster migrates towards the intracellular compartment with energy favorable states (Fig. 8). This may be due to the similarity in structural conformation between SGTP1 and XyleE (with closed intracellular α -helices), thereby impeding its diffusion by trapping glucose within the active site. The fact that SGTP1 differs from XyleE and SGTP4 in energy and migration patterns may be due to the fact that SGTP1 possesses a higher affinity for glucose [4] and that xylose, not glucose, is the native substrate for XyleE [13].

Conclusions

SGTP2/3, not SGTP1/4, are relatively close to human GLUT1/4. This conclusion is based on our observations. The phylogenetic analysis shows that SGTP1/4 belong to a Platyhelminthes-specific glucose transporter class, while SGTP2/3 clearly belong to class I glucose transporters together with vertebrates and insect glucose transporters. Both SGTP2 and SGTP3 show a genome organization similar to that of GLUT1/4 and different to SGTP1/4. SGTP2 seems to be the ancestral glucose transporter in *S. mansoni* and GLUT1 is a structural homolog of SGTP2/3. Thus, the capacity of SGTP1/4 to transport glucose may have evolved independently, while SGTP2/3 apparently lost this capacity after the divergence from the last common ancestor with the other glucose transporters of class I. Overall, these observations may permit the development of specific inhibitors for *S. mansoni* glucose transporters.

A recent review [8] suggested that blocking glucose uptake, as the first step of glucose metabolism, represents an appealing strategy for developing drugs to produce starvation, energy supply deficiency, and mortality in parasitic worms. This is particularly necessary since praziquantel is currently the only treatment available for schistosomiasis and its massive use, notably in sub-Saharan Africa, has led to concerns about the potential for the development of resistance [41]. Glucosides, plant-derived compounds with glucose bound to a functional, non-sugar group (aglycone), have been shown to inhibit glucose diffusion in *S. mansoni* [4, 42] and could therefore represent lead compounds for developing drugs targeting glucose transport. With the advent of computational biology, the ability to rationally design drugs for specific protein targets has greatly advanced. For example, using the *S. mansoni* glucose transporter structures, the aglycone could be altered to computationally test for stability and affinity of the designed

glucoside(s). The top glucoside candidate(s) could then be experimentally analyzed for effectiveness. Since glucosides are non-transportable [43] they could “trap” themselves in the active site to maintain the structural conformation similar to SGTP1 in Fig. 8 – i.e., the glucose-bound, occluded conformation of XyleE. One of the caveats of the PELE algorithm is that it cannot perform large protein conformational changes, as those described by Deng et al. [12]. Such simulations to rationally design drugs will benefit from other molecular dynamics software such as Desmond [44].

Methods

Ethics statement

All animal experimentation was conducted in accordance with the European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (ETS No 123, revised Appendix A) and was approved by the committee for ethics in animal experimentation of the Nord-Pas de Calais region (Authorization No. AF/2009) and the Pasteur Institute of Lille (Agreement No. A59-35009).

Parasite material

A Puerto Rican strain of *S. mansoni* was maintained in the laboratory using the snail, *Biomphalaria glabrata* as its intermediate host and the golden hamster, *Mesocricetus auratus* as a definitive host. Schistosomula were obtained by mechanical transformation from cercariae released from infected snails as previously reported [45]. Schistosomula were washed three times in serum and glucose-free DMEM medium (Life sciences, Carlsbad, USA) supplemented with traces of glucose (0.05 mM) then incubated at 37 °C, in an atmosphere of 5 % CO₂. After three hours of incubation, glucose, maltose or galactose were added to the medium at a concentration of 10 mM each, and the schistosomula were incubated for another eight hours. A total of 10 000 parasites were used per condition and experiments were carried out in triplicate.

RNA isolation and quantitative RT-PCR

Total RNA was isolated from the different stages of *S. mansoni* with TRIzol reagent (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions. Complementary DNAs (cDNA) were obtained by reverse transcription of total RNA using the ThermoScript RT-PCR System (Invitrogen, Carlsbad, USA). The cDNAs were then used as templates in triplicate assays for RT-PCR amplification using the KAPA SYBR FAST ABI Prism kit (Kapa Biosystems, Boston, USA), and ABI PRISM 7000 sequence detection system. We used previously reported primers for *S. mansoni* *sgtp1* and *sgtp4* [5]. The primers for *sgtp2* and *sgtp3* were designed for

this study: *sgtp2F* 5' TTTACCTTCGAGGGCAAGAT 3' and *sgtp2R* 5' CACCGCAAGTATGGAATACG 3'; *sgtp3F* 5' GCAGCAACTCTCAGGAATCA 3' and *sgtp3R* 5'ACACAATAACCGCTCCAACC 3'. The ratios of relative expression were calculated using the $2^{-\Delta\Delta C_t}$ ratio [46] with *S. mansoni* α -tubulin as the endogenous control gene [47]. The statistical significance between groups was evaluated using the unpaired non-parametric Mann Whitney's test in the GraphPad 6 Prism program (GraphPad Software Inc.). Differences were considered significant when p -value <0.05.

Phylogenetic analyses

Sequences were collected from GenBank for phylogenetic analyses of sugar transporter proteins from mammals (3 species), fish (teleosts, 7 species), insects (8 species), cestodes (6 species) and flukes (4 species). A *Drosophila melanogaster* sucrose transporter was introduced as an outgroup. The species, sugar specificity and accession numbers of each sequence is provided in Fig. 2. Glucose transporter amino acid sequences were aligned using MUSCLE (v3.7) configured for highest accuracy [48] and non-aligned regions were removed with Gblocks (v 0.91b) [49]. Thus, the final alignment contained 333 gap-free amino acid positions. The best-fit model of the sequence evolution was selected based on Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (cAIC) and Bayesian Information Criterion (BIC) implemented in Datamonkey [50]. The LG [51] model, which had the lowest values of AIC, cAIC and BIC, was chosen for subsequent phylogenetic analyses. Maximum likelihood (ML), maximum parsimony (MP) and neighbor joining (NJ) methods, implemented in MEGA 6 [52], were used to obtain the best tree topologies for each method. A proportion of Gamma distributed and invariants sites (G + I) were estimated in MEGA for each phylogenetic method. Reliability of internal branches was assessed using the bootstrapping method (1 000 bootstrap replicates) and the approximate likelihood ratio test (aLRT – SH-Like) implemented in MEGA and PhyML, respectively [52, 53]. Graphical representation and editing of the phylogenetic tree was performed with EvolView [54].

Molecular clock and evolutionary rates

First, the likelihood ratio test (LRT) [55] was used to test the molecular clock hypothesis (i.e., all tips of the tree are equidistant from the root of the tree) on the aforementioned phylogenetic trees [52]. The LRT rejected the null hypothesis of equal evolutionary rate throughout the tree at a 5 % significance level (p -value = 0) for both topologies (ML and MP). Therefore, relative time divergence was determined [56]. Time trees were generated using the RelTime method [22], as implemented in

MEGA 6 [52]. RelTime is a useful method to estimate relative lineage-specific evolutionary rates and relative divergence times without requiring the pre-specification of statistical distribution of lineage rates and clock calibrations [22]. To calculate evolutionary rates, RelTime assumes that the elapsed time of two sister lineages from their most recent common ancestor is equal [22]. Divergence times for all branching points were calculated using the Maximum Likelihood method based on the LG model. The MP tree(s) were used to estimate the relative divergence times of glucose transporters in different taxa with special emphasis on *S. mansoni*. To account for statistical errors in our analysis, the 100 most parsimonious trees were searched using the Min-mini heuristic algorithm [57], and the relative divergence times of the selected nodes was determined in all the trees. The Shapiro-Wilk normality test [58] rejected normal distribution of the obtained relative divergence times (P <0.0001), therefore a paired Wilcoxon test was used to test whether the difference between divergence times of selected nodes were significant [56]. The GraphPad 6 Prism programme (GraphPad Software Inc.) was used to perform both Shapiro-Wilk and Wilcoxon tests. Evolutionary rates (measured as the relative number of amino acid substitutions per site) were calculated using the RelTime method [22], implemented in MEGA 6 [52]. To further compare the evolutionary rates of SGTP3/SGTP2 and SGTP4/SGTP1 we used Tajima's relative rate test [23]. This method tests whether two sequences have equal rates of evolution using a third sequence as outgroup. The equality is tested using the chi-square (χ^2) test. When the observed χ^2 is significantly higher (p -value <0.05) than expected, the null hypothesis of equal rates can be rejected. Thus, to perform Tajima's relative rate test, Fhep.-G.AAS94013 and Tsol.G2.AAB05920 sequences (see Fig. 2) were used as outgroups to compare the rates of SGTP3/SGTP2 and SGTP4/SGTP1, respectively.

Analysis of genome organization and transcript variants

The cDNA sequences of the four encoded *S. mansoni* glucose transporters were compared with genomic sequences available at the Wellcome Trust Sanger Institute Blast server (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_mansoni) and GeneDB (<http://www.genedb.org/Homepage/Smansoni>). Intron-exon junctions were manually detected (5'GT and 3'AG) using sequence alignments constructed with Megalign (DNASTar Inc.). For this analysis, the cDNA sequences for SGTP1 [GenBank: L25065], SGTP2 [GenBank: L25066] and SGTP4 [GenBank: L25067] published by Skelly et al. [4] were used. In order to verify and complete the predicted transcript sequences of SGTP2, we carried out 5' and 3' RACE (GeneRacer kit, Invitrogen, Carlsbad, USA) using

the oligonucleotides: *sgtp2F1* 5' GTAAACACAAT CAATGAGACAACACTG 3', *sgtp2R1* 5' GTAGAAAAT AACTGGATAGATGACGA 3' and *sgtp2R2* 5' ATGG GAAATAAAACAAAATAGAACAA 3', based on the predicted sequence. In the case of SGTP3, a similar strategy was adopted, based on the sequence (Smp_127200) predicted by genome annotation [59]. We carried out 5' and 3' RACE PCR using the oligonucleotides: *sgtp35.1* 5' CTGCCGCGCCACGTGACTTTATT 3', *sgtp35.2*, 5' TTGTTGGGATAGAAAAGAAGGAAT 3', *smgtp33.1* 5' ATCTTGGGTTGGAGCGTTATTGT 3', *smgtp33.2* 5' TCACTCAAGAATATAGGGATGC 3'. Subsequently, the full-length coding sequence was amplified using oligonucleotides: *smgtp3FL1* 5' CACTGACATGGACTGAAG GAGT 3', *smgtp3FL2* 5' TGCTACGAGTTTCTGCTTCT CATGC 3', *smgtp3FL3* 5' TTAATGATAGTACTGCACT GATTTA 3', *smgtp3FL4* 5' AGAATCGTTTTACCGG TATGATTGT 3'. Verification and the search for variants of the full-length coding sequences were carried out by PCR using the Advantage 2 Polymerase mix according to the manufacturer's instructions (Clontech, Mountain View, USA). PCR products were purified from agarose gels using the extraction kit Wizard SVGel and PCR clean-up system (Promega) and then inserted into pCR2.1-TOPO vector to chemically transform competent *E. coli* cells (One-Shot TOP10, Invitrogen). Eurofins Genomics GmbH sequenced the clones. Analysis and alignment of the sequences were performed using the LASERGENE package (DNASar).

Tertiary protein modelling

To approximate an accurate tertiary model of the *S. mansoni* glucose transporters SGTP1, SGTP2, SGTP3 and SGTP4 we used several protein structure prediction servers, namely FOLDpro [60], I-TASSER [61], 3D-Jigsaw [62], LOOPP [63], Phyre2 [64] and SwissModel [65]. To assess the quality of the output models and to choose the top candidates we used Resprox [66], Qmean [67], ModFOLD [68]. We then manually inspected the top three models of each SGTP to determine any unresolved secondary structures (i.e., α -helices).

The top candidate model structures were refined via minimization to remove steric clashes and optimization of the hydrogen-bond network by means of side chain sampling using the Schrödinger's Maestro Protein Preparation Wizard [69]. Briefly, the Protein Preparation Wizard analyzes the structure to build a cluster of hydrogen bonds and with the highest degree of sampling, the algorithm then performs 100 000 Monte Carlo orientations for each cluster. Based on the electrostatic and geometric scoring functions, the algorithm then determines an optimized structure. Glucose and the other glucose transporter crystal structures (see below) were also prepared and optimized in this manner.

Glucose induced-fit docking and migration

For the biophysical simulations, namely induced-fit docking and substrate migration, we used the state-of-the-art Protein Energy Landscape Exploration server (PELE) [70]. The PELE server provides ready-made scripts for substrate binding refinement (induced-fit docking) and unconstrained substrate binding site search (migration). The many uses of PELE can be accessed at <https://pele.bsc.es/pele.wt> and its algorithm is thoroughly explained in references [70, 71]. Briefly, the PELE algorithm performs three stages. First, localized substrate perturbations are performed. Protein perturbations of the α -carbon backbone are also performed using an anisotropic network model (ANM) [72]. Second, the amino acid side chains in proximity to the substrate are optimized using steric filters and a rotamer library [73]. Finally, a truncated Newton minimizer and a surface generalized Born implicit solvent for minimization [74] is used to achieve a local minimum after the initial perturbation. These three stages are performed for a number of steps and are analyzed in parallel with several computer-processing units. The results are a series of trajectories that represent protein side chain conformational changes and substrate migration. Based on the calculated energies of each step, a Monte Carlo Metropolis criterion implemented in PELE either accepts (if they are equal to and/or less than the initial energy) or rejects the steps (greater than the initial energy) [71]. The energy is calculated by using a standard force field to describe the potential energy of a particular molecular system, known as the optimized potentials for liquid simulations (OPLS-2005) [75].

Several parameters were optimized from the ready-made scripts provided by the PELE server to facilitate sampling. For the substrate binding refinement (induced-fit docking), (i) *steric_tr* ('100') and *tries* ('25') were reduced for the simulations to be less computationally expensive, (ii) translation (*tra_r* '2') and rotation (*rot_r* '0.3') of glucose were increased to permit more exploration of the active site, (iii) minimization radius (*mirad*) to allow minimization of the whole system, (iv) an ANM *type* to '4' was added and ANM *mode* changed to '5' for favorable protein perturbations; and, (v) an increased number of steps (1000) was used for sufficient overall sampling. For the unconstrained substrate binding site search (migration), (i) *waitfor* was increased to '4' to allow sufficient sampling of the active site, (ii) translation (*tra_r*) was reduced to '4' to avoid large translation of glucose, (iii) an ANM *type* to '4' was added since the default ANM *mode* produced favorable protein perturbations; (iv) an increased number of steps (1000) was used for sufficient overall sampling; and, (v) a focused ANM (*anmrad* 15 Å) and minimization (*mirad* 10 Å) in proximity of the substrate was used for localized interpretation of glucose transport.

As proof of principle, we used two experimentally and structurally characterized glucose transporters: GLUT1 of *H. sapiens* [PDB: 4PYP] and XyIE of *E. coli* [PDB: 4GBZ]. The *xyz* coordinates of glucose from the crystal structure of *E. coli* XyIE were used as the starting position for all simulations (i.e., the active site of glucose transporters). The trajectories of glucose induced-fit (>2 000) and migration (>10 000) produced by the PELE algorithm were viewed and analyzed using the Visual Molecular Dynamics program (VMD) [76]. Clustering analyses of glucose migration was performed using VMD. The clustering parameters were set for '10' clusters with a 10 Å cut-off distance for each cluster, since all simulations showed a concentrated glucose sampling of the active site (Additional file 2). SGTP1 was set with a 5 Å cut-off distance for each cluster since the majority of glucose exploration during the migration simulation was within the active site (Additional file 2). All plots were generated using GNU-PLOT (<http://gnuplot.sourceforge.net>).

Additional files

Additional file 1: Phylogeny of SGTP1, SGTP2, SGTP3 and SGTP4 homologs in *S. japonicum* and *S. haematobium*. The figure shows a phylogenetic tree of the SGTP1, SGTP2, SGTP3 and SGTP4 homologs in *S. japonicum* (Sjap) and *S. haematobium* (Shae). The topology was obtained using the ML method. Numbers on internal branches are the bootstrap values. GenBank accession numbers of each sequence are included. (TIFF 16 kb)

Additional file 2: Glucose migration from the active site of Xyle, GLUT1 and *S. mansoni* glucose transporters. The panels depict glucose migration Cartesian distances (given in Å) from the Tyr292/298 homologous residue (x-axis) and the active site (y-axis) for XyIE, GLUT1, SGTP1, SGTP2, SGTP3 and SGTP4. (TIFF 1023 kb)

Abbreviations

Å: angstrom; aLRT: approximate likelihood ratio test; bp: base pairs; cDNA: complementary deoxyribonucleic acid; LRT: likelihood ratio test; MEGA: molecular evolutionary genetics analysis; ML: maximum likelihood; mM: millimolar; MP: maximum parsimony; mRNA: messenger Ribonucleic acid; NJ: neighbor joining; PCR: polymerase chain reaction; PDB: protein data bank; PELE: protein energy landscape exploration; RACE PCR: rapid amplification of cDNA ends PCR; RMSD: root mean square deviation; RT-PCR: real-Time PCR; SH-like: Shimodaira-Hasegawa-like.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors developed the overall concept and drafted the manuscript. ACC performed and interpreted the phylogenetic, evolutionary analysis and transcription experiments. JJV performed and interpreted the bioinformatics experiments. JL performed and helped in the transcription experiments and sequence analyses. RP carried out the genomic analysis, made critical revisions of the manuscript and interpretation of the data. All authors have read and approved the manuscript in its present form.

Acknowledgments

This work was supported by funding from the European Union's Seventh Framework Programme for research, technological development and demonstration (grant numbers SETReND 241865, A-ParaDDiS 602080). This work was also supported by the Centre National de la Recherche Scientifique

(CNRS), the Institut National de la Santé et de la Recherche Médicale (INSERM) (to R.J.P and J.L), the Ministère de l'Éducation Supérieure et de la Recherche (to A.C.C) and European Social Fund and the state budget of the Czech Republic (CZ.1.07/2.3.00/30.0032 to JJV). We thank Jacques Trolet (Institut Pasteur de Lille) for technical assistance.

Author details

¹Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 8204 - CILIL - Centre d'Infection et d'Immunité de Lille, F-59000 Lille, France.

²Institute of Parasitology, Biology Centre of the Academy of Sciences of the Czech Republic, 37005 České Budějovice, Czech Republic.

Received: 13 May 2015 Accepted: 26 October 2015

Published online: 19 November 2015

References

- Stirewalt MA. *Schistosoma mansoni*: cercaria to schistosomule. Adv Parasitol. 1974;12:115–82.
- Horemans A, Tielens A, van den Bergh S. The reversible effect of glucose on the energy metabolism of *Schistosoma mansoni* cercariae and schistosomula. Mol Biochem Parasitol. 1992;51:73–80.
- Camacho M, Agnew A. Glucose uptake rates by *Schistosoma mansoni*, *S. haematobium*, and *S. bovis* adults using a flow in vitro culture system. J Parasitol. 1995;81:637–40.
- Skelly P, Kim J, Cunningham J, Shoemaker CB. Cloning, characterization, and functional expression of cDNAs encoding glucose transporter proteins from the human parasite *Schistosoma mansoni*. J Biol Chem. 1994;269:4247–53.
- Krautz-Peterson G, Simoes M, Faghiri Z, Ndegwa D, Oliveira G, Shoemaker CB, et al. Suppressing glucose transporter gene expression in schistosomes impairs parasite feeding and decreases survival in the mammalian host. PLoS Pathog. 2010;6, e1000932.
- Skelly P, Shoemaker C. Rapid appearance and asymmetric distribution of glucose transporter SGTP4 at the apical surface of intramammalian-stage *Schistosoma mansoni*. Proc Natl Acad Sci U S A. 1996;93:3642–6.
- Zhong C, Skelly P, Leaffer D, Cohn R, Caulfield J, Shoemaker C. Immunolocalization of a *Schistosoma mansoni* facilitated diffusion glucose transporter to the basal, but not the apical, membranes of the surface syncytium. Parasitology. 1995;110:383–94.
- You H, Stephenson R, Gobert G, McManus D. Revisiting glucose uptake and metabolism in schistosomes: new molecular insights for improved schistosomiasis therapies. Front Genet. 2014;5:176.
- Augustin R. The protein family of glucose transport facilitators: It's not only about glucose after all. IUBMB Life. 2010;62:315–33.
- Klip A, Tsakiridis T, Marette A, Ortiz P. Regulation of expression of glucose transporters by glucose: a review of studies in vivo and in cell cultures. FASEB J. 1994;8:43–53.
- Jiang J, Skelly P, Shoemaker C, Caulfield J. *Schistosoma mansoni*: the glucose transport protein SGTP4 is present in tegumental multilamellar bodies, discoid bodies, and the surface lipid bilayers. Exp Parasitol. 1996;82:201–10.
- Deng D, Xu C, Sun P, Wu J, Yan C, Hu M, et al. Crystal structure of the human glucose transporter GLUT1. Nature. 2014;510:121–5.
- Sun L, Zeng X, Yan C, Sun X, Gong X, Rao Y, et al. Crystal structure of a bacterial homologue of glucose transporters GLUT1–4. Nature. 2012;490:361–6.
- Forsberg H, Ljungdahl P. Sensors of extracellular nutrients in *Saccharomyces cerevisiae*. Curr Genet. 2001;40:91–109.
- Di Palo B, Ripa V, Santi I, Brettoni C, Muzzi A, Metruccio MM, et al. Adaptive response of Group B streptococcus to high glucose conditions: new insights on the CovRS regulation network. PLoS One. 2013;8, e61294.
- Perkins E, Donnellan S, Bertozzi T, Whittington I. Closing the mitochondrial circle on paralogy of the Monogenea (Platyhelminthes) infers evolution in the diet of parasitic flatworms. Int J Parasitol. 2010;40:1237–45.
- Douzery E, Snell E, Baptiste E, Delsuc F, Philippe H. The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? Proc Natl Acad Sci U S A. 2004;101:15386–91.
- Brown C, Todd K, Rosenzweig R. Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. Mol Biol Evol. 1998;15:931–42.
- Wu X, Freeze H. GLUT14, a duplcon of GLUT3, is specifically expressed in testis as alternative splice forms. Genomics. 2002;80:553–7.

20. Robinson-Rechavi M, Laudet V. Evolutionary rates of duplicate genes in fish and mammals. *Mol Biol Evol.* 2001;18:681–3.
21. Nguyen Ba A, Strome B, Hua J, Desmond J, Gagnon-Arsenault I, Weiss E, et al. Detecting functional divergence after gene duplication through evolutionary changes in posttranslational regulatory sequences. *PLoS Comput Biol.* 2014;10, e1003977.
22. Tamura K, Battistuzzi F, Billings-Ross P, Murillo O, Filipinski A, Kumar S. Estimating divergence times in large molecular phylogenies. *Proc Natl Acad Sci U S A.* 2012;109:19333–8.
23. Tajima F. Simple methods for testing molecular clock hypothesis. *Genetics.* 1993;135:599–607.
24. Zhao F, Keating A. Functional properties and genomics of glucose transporters. *Curr Genomics.* 2007;8:113–28.
25. Uldry M, Thorens B. The SLC2 family of facilitated hexose and polyol transporters. *Pflugers Arch.* 2004;447(5):480–9.
26. Kanamori Y, Saito A, Hagiwara-Komoda Y, Tanaka D, Mitsumasa K, Kikuta S, et al. The trehalose transporter 1 gene sequence is conserved in insects and encodes proteins with different kinetic properties involved in trehalose import into peripheral tissues. *Insect Biochem Mol Biol.* 2010;40(1):30–7.
27. Moller S, Croning M, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics.* 2001;17:646–53.
28. Mueckler M, Caruso C, Baldwin S, Panico M, Blench I, Morris H, et al. Sequence and structure of a human glucose transporter. *Science.* 1985;229:941–45.
29. Kikuta S, Kikawada T, Hagiwara-Komoda Y, Nakashima N, Noda H. Sugar transporter genes of the brown planthopper, *Nilaparvata lugens*: a facilitated glucose/fructose transporter. *Insect Biochem Mol Biol.* 2010;40:805–13.
30. Joost H, Thorens B. The extended GLUT-family of sugar/polyol transport facilitators: nomenclature, sequence characteristics, and potential function of its novel members (review). *Mol Membr Biol.* 2001;18:247e256.
31. Schurmann A, Doege H, Ohnimus H, Monser V, Buchs A, Joost H. Role of conserved arginine and glutamate residues on the cytosolic surface of glucose transporters for transporter function. *Biochemistry.* 1997;36:12897–902.
32. Sato M, Mueckler M. A conserved amino acid motif (R-X-G-R-R) in the Glut1 glucose transporter is an important determinant of membrane topology. *J Biol Chem.* 1999;274:24721–5.
33. Mueckler M, Thorens B. The SLC2 (GLUT) family of membrane transporters. *Mol Aspects Med.* 2013;34:121–38.
34. Seatter M, De la Rue S, Porter L, Gould G. QLS motif in transmembrane helix VII of the glucose transporter family interacts with the C-1 position of D-glucose and is involved in substrate selection at the exofacial binding site. *Biochemistry.* 1998;37:1322–6.
35. Johnson J, Newgard C, Milburn J, Lodish H, Thorens B. The high Km glucose transporter of islets of Langerhans is functionally similar to the low affinity transporter of liver and has an identical primary sequence. *J Biol Chem.* 1990;265:6548–51.
36. Uldry M, Ibberson M, Hosokawa M, Thorens B. GLUT2 is a high affinity glucosamine transporter. *FEBS Lett.* 2002;524:199–203.
37. Teusink B, Diderich J, Westerhoff H, van Dam K, Walsh M. Intracellular glucose concentration in derepressed yeast cells consuming glucose is high enough to reduce the glucose transport rate by 50 %. *J Bacteriol.* 1998;180:556–62.
38. Boles E. Yeast as a model system for studying glucose transport. In: Quick MW, editor. *Transmembrane transporters*. Hoboken: Wiley; 2002. p. 19–35.
39. Holm L, Rosenström P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 2010;38:545–9.
40. Koshland DE. Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci U S A.* 1958;44:98–104.
41. Doenhoff M, Cioli D, Utzinger J. Praziquantel: mechanisms of action, resistance and new derivatives for schistosomiasis. *Curr Opin Infect Dis.* 2008;21:659–67.
42. Uglem GL, Read C. Sugar transport and metabolism in *Schistosoma mansoni*. *J Parasitol.* 1975;61:390–7.
43. Toggenburger G, Kessler M, Semenza G. Phlorizin as a probe of the small-intestinal Na⁺, D-glucose cotransporter. A model. *Biochim Biophys Acta.* 1982;688:557–71.
44. Borhani D, Shaw D. The future of molecular dynamics simulations in drug discovery. *J Comput Aided Mol Des.* 2012;26:15–26.
45. Ramalho-Pinto F, Gazzinelli G, Howells R, Mota-Santos T, Figueiredo E, Pellegrino J. *Schistosoma mansoni*: defined system for stepwise transformation of cercaria to schistosomule in vitro. *Exp Parasitol.* 1974;36:360–72.
46. Pfaffl M. A new mathematical model for relative quantification in realtime RT-PCR. *Nucleic Acids Res.* 2001;29, e45.
47. Lancelot J, Caby S, Dubois-Abdesselem F, Vanderstraete M, Trolet J, Oliveira G, et al. *Schistosoma mansoni* Sirtuins: characterization and potential as chemotherapeutic targets. *PLoS Negl Trop Dis.* 2013;7(9):e2428.
48. Edgar R. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32:1792–7.
49. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 2000;17:540–52.
50. Delport W, Poon A, Frost S, Kosakovsky P. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics.* 2010;26:2455–7.
51. Le S, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol.* 2008;25:1307–20.
52. Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol.* 2013;30:2725–9.
53. Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst Biol.* 2006;55:539–52.
54. Zhang H, Gao S, Lercher M, Hu S, Chen W. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res.* 2012;40:W569–572.
55. Felsenstein J. Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet.* 1988;22:521–65.
56. Wilke T, Schultheiß R, Albrecht C. As time goes by: a simple fool's guide to molecular clock approaches in invertebrates. *Amer Malac Bull.* 2009;27:25–45.
57. Nei M, Kumar S. *Molecular evolution and phylogenetics*. New York: Oxford University Press; 2000.
58. Shapiro S, Wilk M. An analysis of variance test for normality (complete samples). *Biometrika.* 1965;52:591–611.
59. Protasio A, Tsai I, Babbage A, Nichol S, Hunt M, Aslett M, et al. A systematically improved high quality genome and transcriptome of the human blood fluke *Schistosoma mansoni*. *PLoS Negl Trop Dis.* 2012;6, e1455.
60. Cheng J, Baldi P. A machine learning information retrieval approach to protein fold recognition. *Bioinformatics.* 2006;22:1456–63.
61. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics.* 2008;9:40.
62. Bates P, Kelley L, MacCallum R, Sternberg M. Enhancement of protein modelling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins.* 2001;5:39–46.
63. Vallat B, Pillardy J, Májek P, Meller J, Blom T, Cao B, et al. Building and assessing atomic models of proteins from structural templates: learning and benchmarks. *Proteins.* 2009;76:930–45.
64. Kelley L, Sternberg M. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc.* 2009;4:363–71.
65. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL Workspace: a web-based environment for protein structure homology modelling. *Bioinformatics.* 2006;22:195–201.
66. Berjanskii M, Zhou J, Liang Y, Lin G, Wishart D. Resolution-by-proxy: a simple measure for assessing and comparing the overall quality of NMR protein structures. *J Biomol NMR.* 2012;53:167–80.
67. Benkert P, Künzli M, Schwede T. QMEAN server for protein model quality estimation. *Nucleic Acids Res.* 2009;37:510–4.
68. McGuffin L, Buenavista M, Roche D. The ModFOLD4 server for the quality assessment of 3D protein models. *Nucleic Acids Res.* 2013;41:368–72.
69. Li X, Jacobson M, Zhu K, Zhao S, Friesner R. Assignment of polar states for protein amino acid residues using an interaction cluster decomposition algorithm and its application to high resolution protein structure modeling. *Proteins.* 2007;66:824–37.
70. Madadkar-Sobhani A, Guallar V. PELE web server: atomistic study of biomolecular systems at your fingertips. *Nucleic Acids Res.* 2013;41:322–33.
71. Borrelli K, Vitalis A, Alcantara R, Guallar V. PELE: Protein energy landscape exploration. A novel Monte Carlo based technique. *J Chem Theory Comput.* 2005;1:1304–11.
72. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J.* 2001;80:505–15.
73. Jacobson M, Friesner R, Xiang Z, Honig B. On the role of the crystal environment in determining protein side-chain conformations. *J Mol Biol.* 2002;320:597–608.

74. Still W, Tempczyk A, Hawley R, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J Am Chem Soc.* 1990;112:6127–9.
75. Jorgensen W, Tirado-Rives J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc.* 1988;110:1657–66.
76. Humphrey W, Dalke A, Schulten K. VMD - Visual Molecular Dynamics. *J Molec Graphics.* 1996;14:33–8.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

