J Mol Model (2009) 15:369–382
DOI 10.1007/s00894-008-0404-y

ORIGINAL PAPER

EGF domain II of protein Pb28 from *Plasmodium berghei* interacts with monoclonal transmission blocking antibody 13.1

Babita Sharma · Manoj Kumar Jaiswal

Received: 3 June 2008 / Accepted: 21 August 2008 / Published online: 9 December 2008
© Springer-Verlag 2008

Abstract Development of a vaccine against malaria is a major global health concern. The P28 proteins expressed on the surface of ookinetes of *Plasmodium* are the targets of transmission blocking antibodies. Injection of P28 proteins in vertebrate hosts induces antibodies that inhibit oocyst formation, blocking transmission of the parasite from mosquitos to human hosts. P28 proteins are crucial for parasite protection inside the mosquito midgut. Despite their importance, structural details of P28 family members have not been available to date. The purpose of this study was to structurally characterise a member of the P28 family, viz. Pb28 protein from *Plasmodium berghei*, and to study the interaction of Pb28 protein with the scFv (single chain variable fragment) of TBmAb (transmission blocking monoclonal antibody) 13.1 which blocks malaria transmission effectively. Pb28 protein and the TBmAb 13.1 scFv were modelled separately. To decipher the antigen–antibody interaction, ZDOCK and RDOCK programs were used. Our results suggest that, as compared to the template Pvs25, Pb28 protein has four EGF (epidermal growth factor)-like domains arranged in a triangular form with maximum root mean square deviations (RMSDs) present in the loop regions of EGF domains II and III. With the help of docking we were able to show that the B loop of EGF domain II of Pb28 protein interacts with the scFv of TBmAb 13.1. The predicted probable complex of Pb28 protein and 13.1 TBmAb suggests a mechanism for transmission blocking and may help in designing vaccine candidates in the absence of experimentally determined structures of these proteins.

B. Sharma (✉) · M. K. Jaiswal
School of Life Sciences, Jawaharlal Nehru University,
New Delhi 110067, India
e-mail: babita2005@gmail.com

Keywords Epidermal growth factor · Glycosylphosphatidylinositol anchor · Homology modeling · P28 family · *Plasmodium berghei* 28 protein · Protein-protein docking · Transmission blocking monoclonal antibody 13.1

Introduction

Every year, 300–600 million people across the globe are affected by malaria, leading to 1–3 million deaths [1, 2]. The malaria-causing parasite *Plasmodium* undergoes several stages of development inside its human host and forms gametocytes. When taken up by the mosquito vector in a blood meal, these gametocytes form micro and mega gametes inside the mosquito midgut. The gametes fuse to form zygotes, which develop into motile ookinetes that cross the mosquito midgut epithelium, forming oocysts. Each mature oocyst releases thousands of sporozoites. Sporozoites migrate to the salivary glands of the mosquito and are transferred to another human host during the mosquito bite that causes infection.

Proteins of the P28 family are present on the ookinete surface of all known *Plasmodium* species [3–5]. These proteins are expressed immediately after fertilisation and are present on the zygote, ookinete and young oocyst stages of *Plasmodium* [6]. P28 proteins are distributed evenly and abundantly over the entire ookinete surface as shown by immunofluorescent antibody staining [7] and immunogold electron microscopy [8, 9]. Gene knock-out experiments suggest that these proteins play an important role in parasite survival inside the mosquito midgut [10]. Proteins of the P28 family contain a signal sequence; four EGF (epidermal growth factor)-like domains and a C-terminal GPI (glycosylphosphatidylinositol) anchor [11, 12]. The presence of EGF-domains in surface proteins

participating in processes such as recognition, adhesion and signalling is well known [13], indicating that P28 proteins may play an important role in host–parasite interactions. Also, P28 proteins exhibit very few sequence polymorphisms. This is presumably because these proteins are not expressed in the vertebrate host and therefore not exposed to selection pressure from the vertebrate immune system [14].

The process by which transmission-blocking antibodies inhibit parasite development is not clear. Results show that transmission-blocking antibodies arrest or slow down the movement of the ookinete in the mosquito midgut [15]. The levels of antibody correlate directly with transmission-blocking activity [16]. It has been shown that an ookinete delayed in forming oocysts is unable to survive in the harsh proteolytic environment inside the mosquito midgut [17]. Recombinant versions of transmission-blocking antibodies also block transmission [18, 19]. Elucidating the molecular structure of transmission-blocking monoclonal antibody (TBmAb) 13.1 bound to antigen Pb28 (Swissprot id Q04620) from *Plasmodium berghei* may help in understanding mechanism of transmission blocking.

The popularity of protein–protein docking as a powerful means of predicting protein complex structures, particularly in cases where the crystal complex proves difficult to obtain, is growing almost daily [20, 21]. In some cases, even atomic level accuracy is within reach. Here, we used the programs ZDOCK and RDOCK (http://zlab.bu.edu/zdock/) to dock Pb28 to the single chain variable fragment (scFv) of TBmAb 13.1. These programs have successfully recapitulated structures of many known protein–protein complexes, and produced high accuracy predictions for multiple protein–protein targets in the CAPRI challenge [22]. Docking programs prove extremely helpful if some biochemical information is incorporated. The homology modelling of protein Pb28 presented here suggests that the mature Pb28 protein is probably a triangular flat molecule. Compared with the template Pvs25 (PBD id 1z27), Pb28 has four EGF domains with maximum variations present in EGF domain III and the B

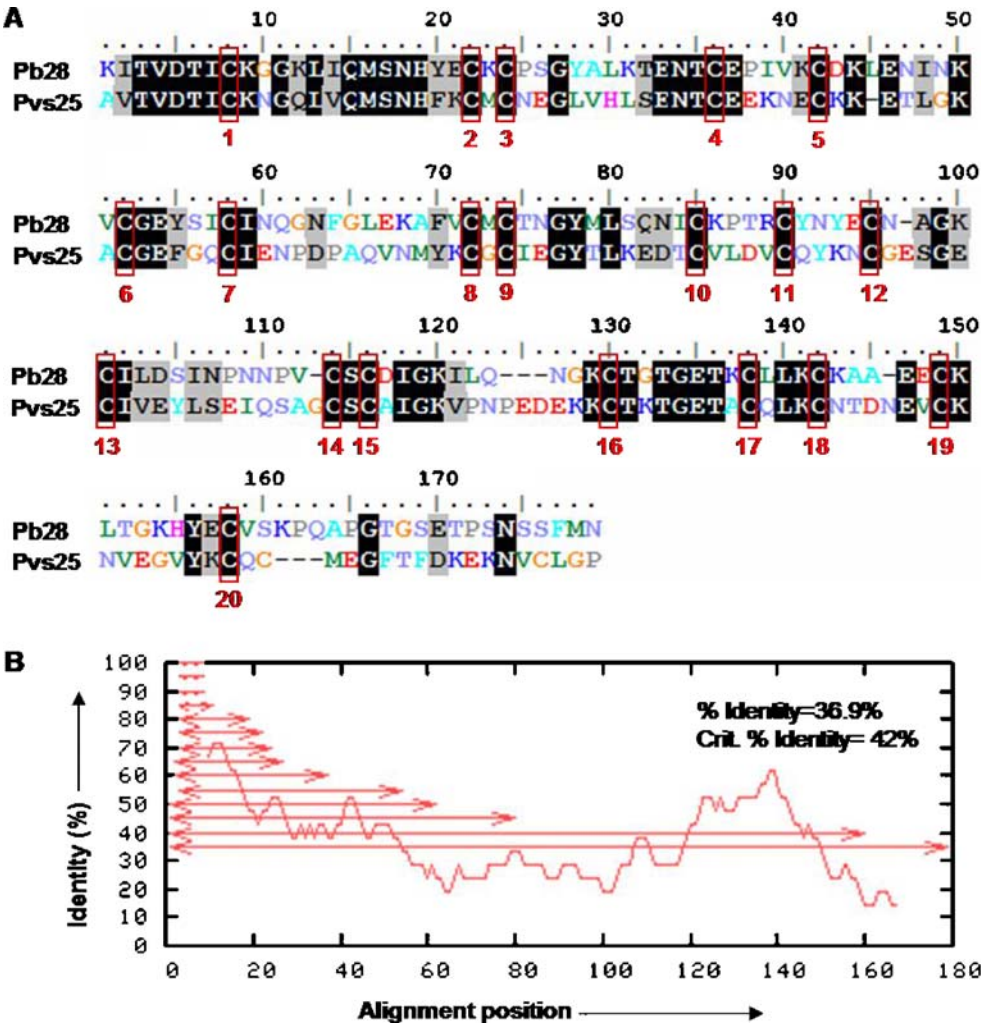

Fig. 1 **a** Sequence alignment of query sequence Pb28 from the P28 family of *Plasmodium berghei* ookinete surface antigens with template sequence Pvs25 (PDB id 1z27) protein from *Plasmodium vivax*. Completely conserved residues between the two sequences are shaded *black* whereas semi-conserved residues are shaded *grey*. The 20 conserved cysteines are *boxed in red*. Amino acids: *Red* Negatively charged (D, E), *blue* positively charged (K, R), *bluish-purple* polar (N, Q, S, T), *green* aliphatic (I, L, M, V), *cyan* aromatic (F, Y), *pink* histidine, *buff* glycine, *grey* proline. **b** Numerical and graphical representation of the percent identity between the target sequence (Pb28 starting from KITVDT…) and template sequence (Pvs25)

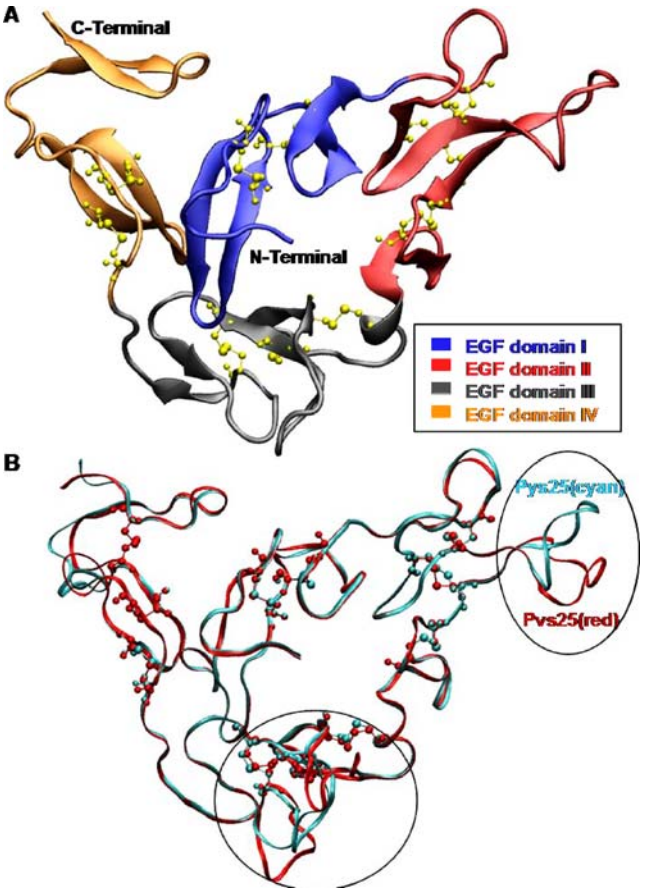

Fig. 2 **a** Cartoon representation of the Pb28 model. The model was generated assuming that the disulphide connectivity of 20 cysteines (shown in *yellow CPK* representation) in Pb28 is the same as that in the template Pvs25 (PDB code: 1Z27). The Pb28 model shows four EGF domains as discussed in the text: *blue* EGF-domain I, *red* EGF-domain II, *grey* EGF-domain III, *orange* EGF-domain IV. **b** Superimposed structures of Pb28 and Pvs25(C-alpha atoms only) after structural alignment using the program STAMP. Superimposition between the theoretical three-dimensional (3D) model of Pb28 and the Pvs25 template reveals the similarity in the overall fold of the two proteins. EGF domain III and the B-loop of EGF domain II show the most variation

loop of EGF domain II. The scFv of TBmAb 13.1 was modelled with the help of the WAM server (http://www.bath.ac.uk/cpad/) and refined with the help of the ModLoop program (http://alto.compbio.ucsf.edu/modloop//modloop.html). In *P. berghei*, TBMab 13.1 was mapped using deletions and peptides overlapping the sequence GLEKAFVC on the B loop of EGF domain II of the Pb28 protein [23]. We incorporated the available biochemical information while docking the two proteins. Docking of the Pb28 protein model with scFv of TBmAb 13.1 showed that EGF domain II of Pb28 interacts with the scFv of TBmAb 13.1.

Methods

Template selection

The single letter amino acid sequence of Pb28 protein was taken from Swiss-Prot (id=Q04620) [24], release 50.4 (http://www.expasy.org/uniprot/Q04620). In order to select a suitable template for the Pb28 protein, PSI-BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) [25] was performed against the PDB database (http://www.rcsb.org/pdb/Welcome.do, 51,977 structures) [26], which resulted in a single significant hit (E value=0.001), i.e. Pvs25 protein from *Plasmodium vivax*, which was found to be the only suitable template in the PDB database for modelling Pb28 protein. In both template and target proteins there are 20 absolutely conserved cysteines, which, in the oxidised state, will form ten disulphide bonds. These disulphide bonds are formed between cysteines 1 and 3, 2 and 4 ,and 5 and 6 of EGF domains II and III. In EGF domain I, four cysteines are conserved forming two disulphide bonds (1–3 and 2–4), whereas in EGF domain IV only four out of six cysteines are conserved. Sequence alignment was performed with the

Table 1 Conserved domain interactions between *Plasmodium berghei* protein Pb28 and the Pvs25 protein from *Plasmodium vivax*

<—— A T O M 1 ——>				EGF–domain involved	<—— A T O M 2 ——>				EGF–domain involved	H-bond distance (Å)
Atom no.	Atom name	Res name	Res no.		Atom no.	Atom name	Res name	Res no.		
109	NE2	Gln	15	I	<–> 997	O	Lys	132	III	2.88
120	N	Ser	17	I	<–> 869	O	Cys	114	III	2.68
123	OG	Ser	17	I	<–> 864	N	Cys	114	III	2.66
126	N	Asn	18	I	<–> 861	OG	Ser	113	III	2.96
136	ND1	His	19	I	<–> 861	OG	Ser	113	III	2.73
36	OD1	Asp	5	I	<–>1026	NZ	Lys	136	IV	2.70
103	N	Gln	15	I	<–>1136	O	Tyr	150	IV	3.08
111	O	Gln	15	I	<–>1125	N	Tyr	150	IV	2.95
108	OE1	Gln	15	I	<–>1004	N	Leu	134	IV	2.88
220	N	Leu	30	I	<–> 409	OE1	Glu	54	II	3.06
227	O	Leu	30	I	<–> 413	N	Tyr	55	II	2.71

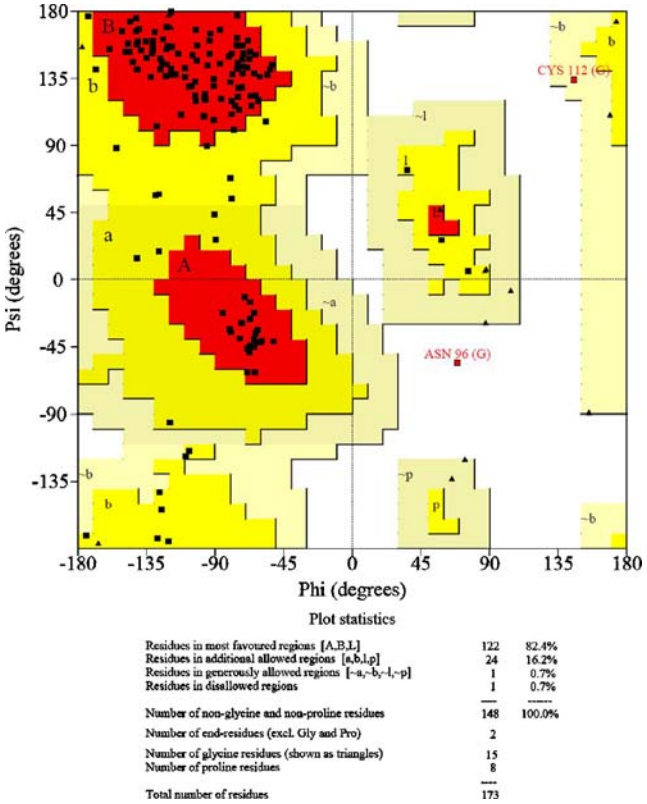

Fig. 3 Ramachandran plot of the theoretical 3D model of Pb28 using the X-ray crystallographic structure of native Pvs25 (PDB code 1Z27) as a template

help of ClustalX (http://www.embl.de/~chenna/clustal/darwin/index.html) [27]. Sequence identity between the two proteins was calculated with the help of the Graph Align (http://darwin.nmsu.edu/cgi-bin/graph_align.cgi) [28] and Ident and Sim programs (http://www.bioinformatics.org/sms2/ident_sim.html) [29].

Homology modelling

The Pb28 sequence obtained from the Swissprot database was modelled using MODELLER version 9v1. This software implements homology modelling of proteins by satisfaction of spatial restraints [30]. Hundred models were obtained for the Pb28 protein with the help of Modeller. Modeller automatically derives restraints from known related structures. The restraints include distances, angles, dihedral angles, pairs of dihedral angles and some other spatial restraints. Bond and angle values are taken from CHARMM-22 force field. Three-dimensional (3D) models are generated by molecular probability density function optimisation. The X-ray crystallographic structure of Pvs25 from *P. vivax* (PDB id 1Z27) [31–33] was used as a template. To model the variable loops, the program ModLoop (http://alto.compbio.ucsf.edu/modloop//modloop.html) was used [34, 35]. Loop B of EGF domain II was modelled with particular care. Modelling of the scFv of TBmAb 13.1 (NCBI accession no. JC5810) was done using the Web Antibody Modeling server: WAM (http://www.bath.ac.uk/cpad/) [36, 37], and was refined with the help of Modeller and ModLoop.

Model evaluation

Evaluation of the Pb28 protein model was carried out using the ProSA-web (https://prosa.services.came.sbg.ac.

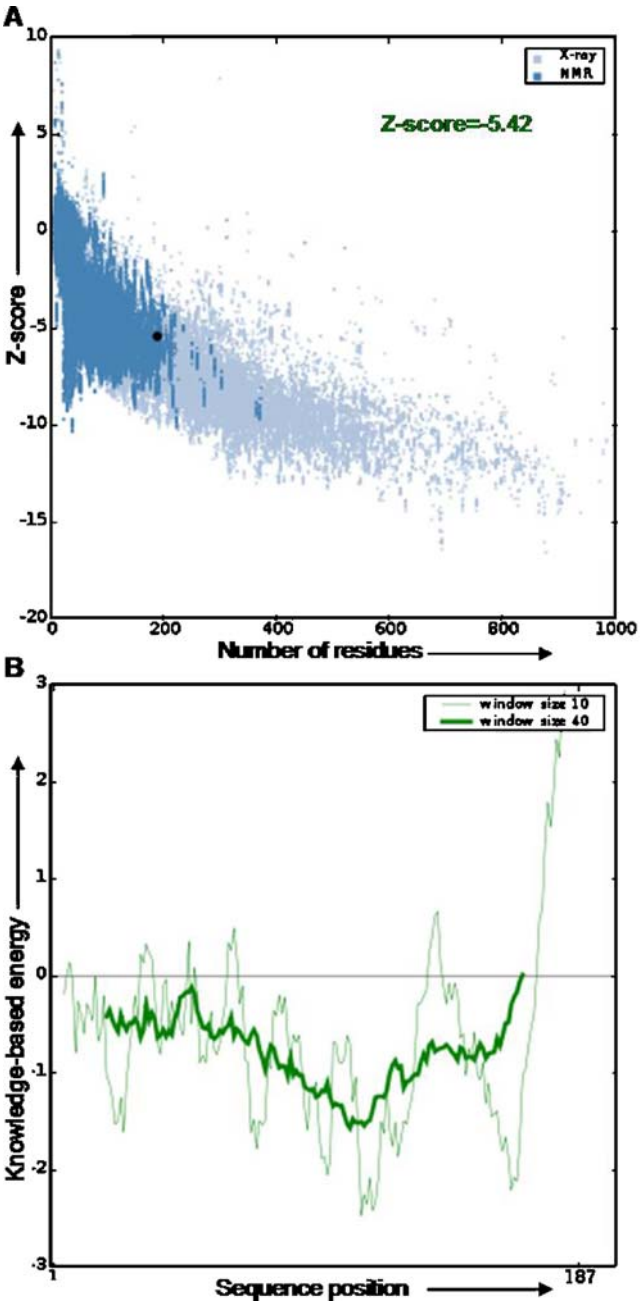

Fig. 4a,b ProsaWeb analysis of theoretical 3D model of Pb28. **a** Zscore plot of the Pb28 model where all the values lie within the normal range. **b** Energy plot of the Pb28 model. Energy for all residues is below zero or negative, indicating stability of the molecule

at/prosa.php) server [38, 39], PROCHECK (http://nihserver.mbi.ucla.edu/SAVS/) program [40] and WHATIF server (http://swift.cmbi.kun.nl/WIWWWI/) [41]. Only those models that showed a satisfactory PROSA, PROCHECK and WHATIF profile [42, 43] were selected. To model the flexible loops, ModLoop (http://alto.compbio.ucsf.edu/modloop//modloop.html) was used. 3D structural superimposition for comparing the structures was achieved using the program STAMP, which is part of VMD version 1.8.4 [44].

Docking

ZDOCK version 2.1 was used to predict the bound antigen–antibody complex structure from its unbound monomer protein components. ZDOCK (http://zlab.bu.edu/zdock/) is an initial stage protein–protein docking program that employs a fast Fourier transform (FFT) algorithm to translate and rotate the antigen around the surface of the antibody. ZDOCK evaluates pairwise shape complementarity, desolvation and electrostatic energies [45]. The script "Mark_sur" was used to mark the model PDB files in order to make the files readable with the ZDOCK program. Rigid body docking was done using ZDOCK, and 2,000 model complexes were generated. Refinement and ranking of the models obtained was done with the help of RDOCK (http://zlab.bu.edu/zdock/) program [46]. RDOCK is a program that improves significantly upon the top ZDOCK predictions by a three-stage energy minimisation using the program CHARMM. RDOCK refines complex structures, evaluates binding free energies, and re-ranks complexes generated by the ZDOCK program. The script "rdock.pl", which uses the CHARMM (v32) program [47] for minimisation, was used. RDOCK minimises the complexes generated by ZDOCK in a three-step process: step 1, small clashes are removed to allow small conformational changes; step 2, polar interactions are optimised; and step 3, optimisation of charged interactions. Note that the charges for ionic residues are turned off in steps 1 and 2. To re-rank the top RDOCK predictions, the script rerank.pl was used. A re-ranked list of complexes was obtained based on electrostatic, van der Waals, and atomic contact energy (ACE), and the top 30 complexes were selected out of the 2,000 solutions obtained. These top 30 complexes were grouped into different clusters based on their interaction with the EGF domains. Complexes showing binding to EGF domain II were selected and examined for an antigen–antibody type

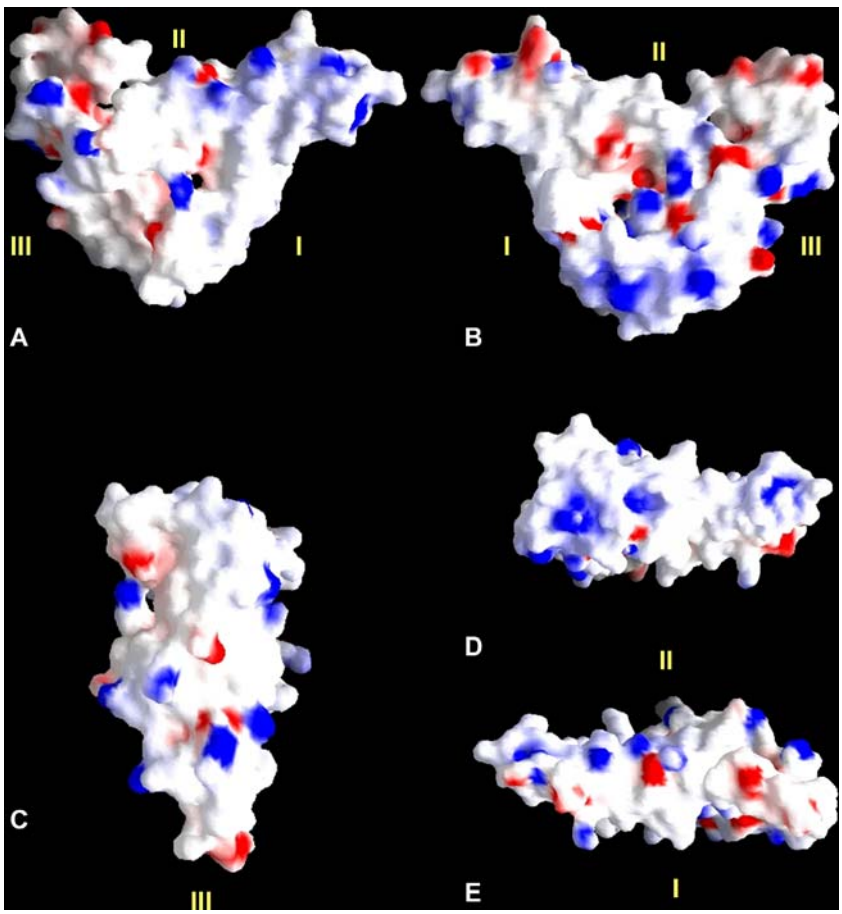

Fig. 5a–e Electrostatic representation of the Pb28 model generated by the GRASP program. Surface potential was taken from −10kT (*red*) to +10kT (*blue*). **a** Front view of Pb28, (**b**) opposite face after 180° rotation. **c–e** Edge-on views of the molecule. ProsaWeb, PROCHECK and WhatIF programs were used for statistical evaluation of the final model. As represented by the different views of the molecule, Pb28 model shows very few negatively charged residues in comparison to Pvs25 and, unlike Pvs25, the molecule is positively charged overall (total charge=+3). EGF domain II is more positively charged in the front view, whereas EGF domain III is more positively charged in the opposite view of the molecule. The orientation in **a** is same as that shown in Fig. 2

of interaction. Only those complexes showing interaction with complementarity determining regions (CDRs) of the scFv of antibody 13.1 were selected. The remaining complexes were discarded. The buried surface area was calculated for the selected complexes showing antigen–antibody type interaction using the POPS comp server (http://mathbio.nimr.mrc.ac.uk/tool/popscomp/) [48]. The complex showing maximum buried surface area was considered as the best possible interaction between the two proteins. Residue to residue interactions for this selected complex were calculated using the PDBsum server (http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/upload.html) [49].

Results and discussion

Sequence conservation between Pb28 protein and Pvs25

P28 proteins located on the surface of the ookinete form of the parasite are promising malaria transmission-blocking vaccine candidate proteins. Figure 1 depicts a sequence alignment of Pb28 protein (Swiss-Prot id=Q04620) with the template protein Pvs25 (1z27) [50, 51]. Along with several other residues, 20 cysteine residues are absolutely conserved in the P28 family. These conserved cysteines provide a structural scaffold to Pb28 protein, forming ten disulfide bonds, and this scaffold is conserved throughout the P28 family. Residues that make contacts among domains 1, 3 and 4 of Pb28 protein to form a triangular shape are also conserved, as shown in Table 1. Due to the presence of the many conserved cysteines, it has been proposed that proteins of the P25 and P28 families have evolved as a result of gene duplication and hence should contain similar folds and domains. This belief is further supported by the results of our Pb28 model, which show that the P28 proteins contain four EGF domains arranged in triangular fashion similar to Pvs25 protein of the P25 family as shown in Fig. 2a. EGF-domains have a high tolerance to polymorphism and mutations, and are tolerant to insertions and deletions; therefore, a high level of similarity between the template and query protein was expected. The structure of Pb28 protein was thus predicted to be similar to the template protein Pvs25 [31–33] because of the presence of 20 conserved cysteines, 37% sequence identity and 52% sequence similarity (residues taken as similar: GAVLI, FYW, CM, ST, KRH, DENQ, P) between the template protein Pvs25 and target protein Pb28.

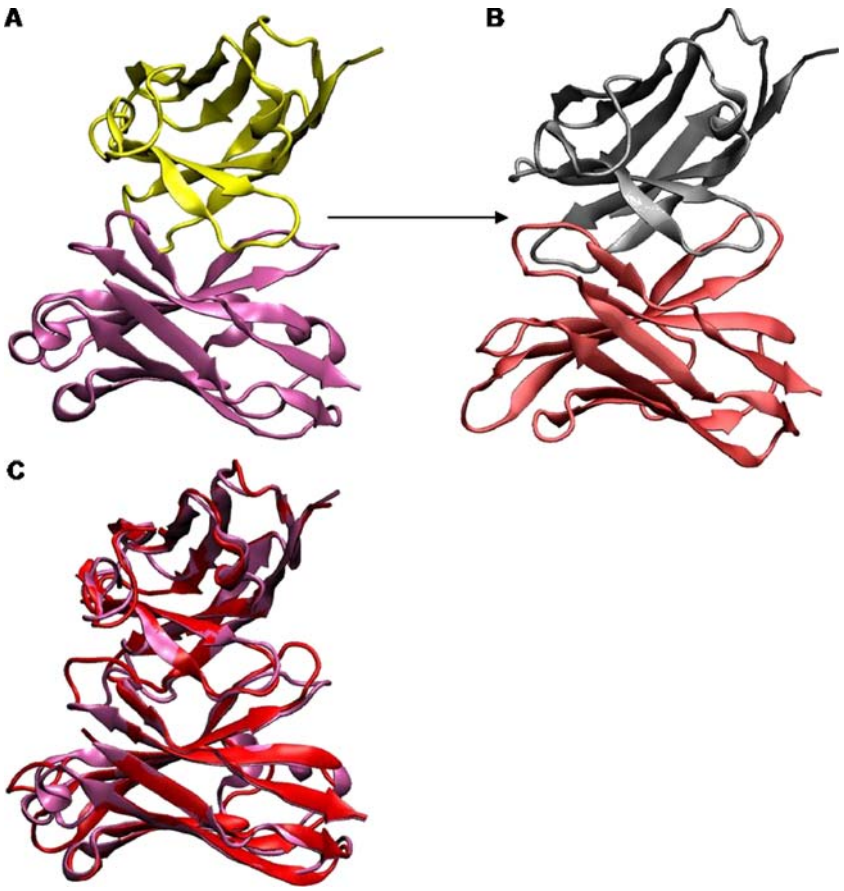

Fig. 6 **a** Cartoon representation of the model of the single chain variable fragment (scFv) region of transmission blocking antibody (TBmAb) 13.1. The model was generated with the help of the web antibody modeling (WAM) server. **b** Final model obtained after refining the scFv of 13.1 WAM model with the help of Modeller and ModLoop. **c** Superimposed WAM model and final model of scFv of 13.1 to show the changes due to refinement of the model

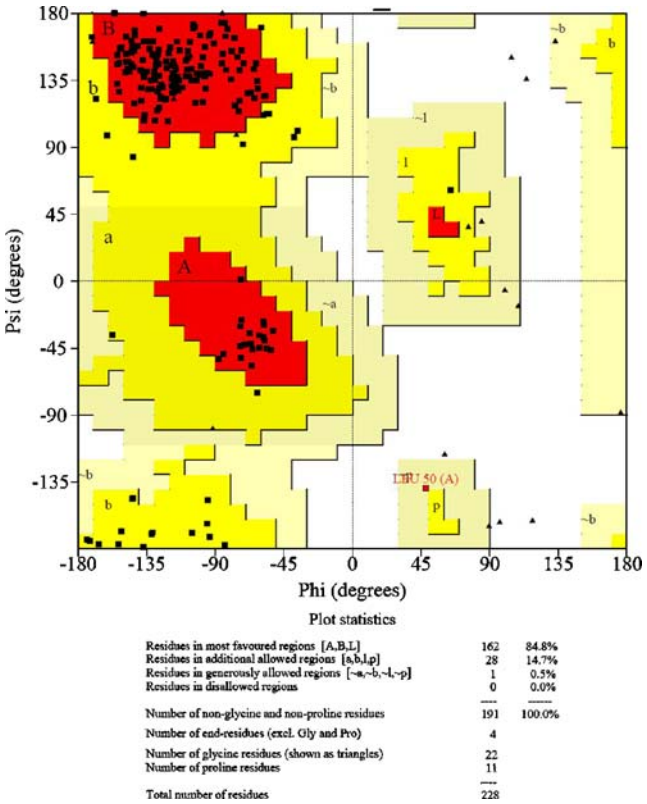

Fig. 7 Ramachandran plot of the theoretical 3D model of scFv of TBmAb 13.1

Template

The mature sequence of Pb28 protein from *Plasmodium* was taken from the Swissprot database (Swiss-Prot id = Q04620). The sequence of Pb28 protein is as follows:

- >Q04620|OS24_PLABA 24 kDa ookinete surface protein - *Plasmodium berghei* (strain Anka).
- MNFKYSFIFLFFIQLAIRYNNAKITVDTICKG
- GKLIQMSNHYECKCPSGYALKTENTCEP
- IVKCDKLENINKVCGEYSICINQGNFGLE
- KAFVCMCTNGYMLSQNICKPTRCYNYECNAG
- KCILDSINPNNPVCSCDIGKILQNGKCTGT
- GETKCLLKCKAAEECKLTGKHYECVSKPQA
- PGTGSETPSNSSFMNGMSIISIIALLVIYVIVM

Only the mature protein was modelled, i.e. the signal sequence MNFKYSFIFLFFIQLAIRYNNA was removed from the N-terminus and the GPI anchor sequence GMSII SIIALLVIYVIVM was removed from the C-terminus before modelling the protein. As already mentioned, the arrangement of cysteines in Pb28 is similar to that in Pvs25 (PDB code: 1Z27) protein, and there is significant sequence identity (52%; Fig. 1). Pvs25 protein was the only structure in the PDB database showing significant similarity to Pb28 protein, and therefore Pvs25 was used as a template for modelling of Pb28 protein.

Pb28 model

The model of Pb28 protein indicates three EGF domains and a fourth, truncated, EGF domain arranged in the form of a triangle, as represented in Fig. 2a. Structural superimposition of the Pb28 model with template Pvs25 reveals similarity in the overall fold of the two proteins. In comparison to the template protein Pvs25, Pb28 protein

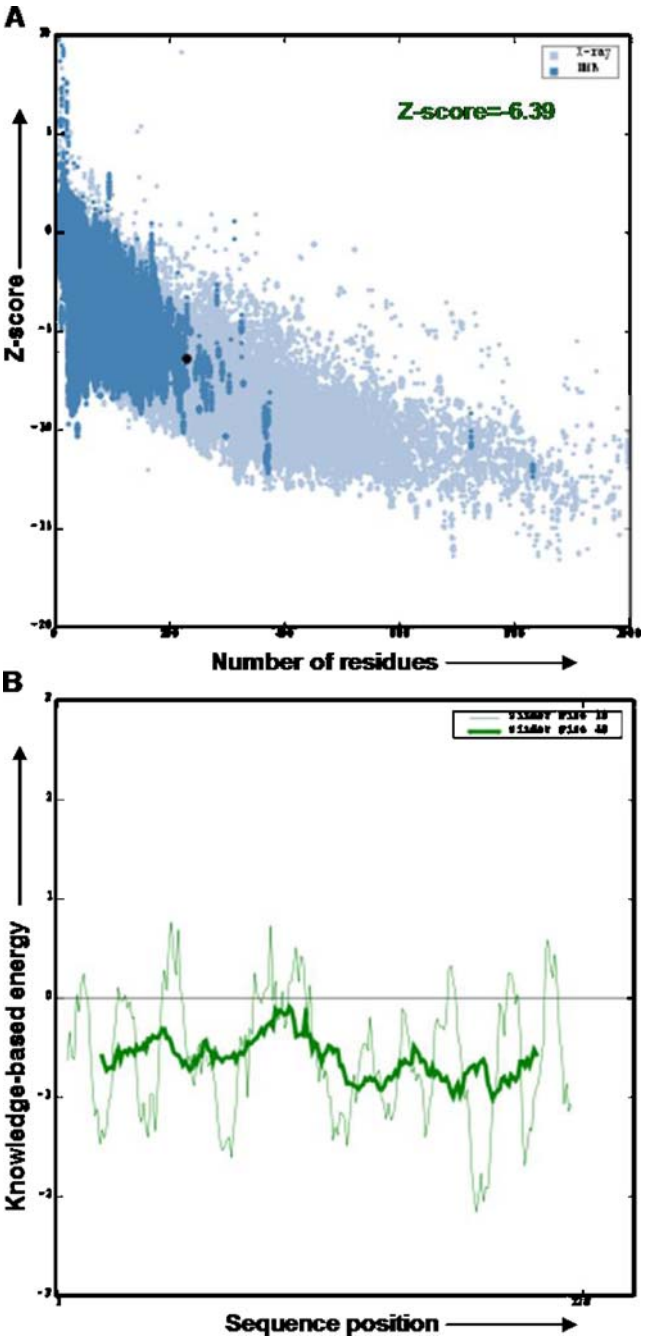

Fig. 8a,b ProsaWeb analysis of theoretical 3D model of TBmAb 13.1. **a** Zscore plot of the 13.1 model where all the values lie within the normal range. **b** Energy plot of the 13.1 model. Energy for all the residues is below zero or negative indicating stability of the molecule

contains most variation in the B loop of EGF domain II and in EGF domain III (Fig. 2b). Comparison of the Pb28 and Pvs25 proteins showed that, out of the four EGF-domains of Pb28, EGF domains I and IV had four cysteines whereas the other two domains contained six cysteines each. EGF domain I of Pb28 shows maximum similarity to EGF domain I of Pvs25, with an average root mean square deviation (RMSD) of 1.21 Å and maximum RMSD of 2.24 Å, whereas EGF domain III of Pb28 showed maximum deviation to EGF domain III of Pvs25, with a maximum RMSD difference of 5.58 Å and average RMSD of 2.79 Å. EGF domain II had a maximum RMSD of 5.07 Å, whereas the average RMSD for EGF domain II was 2.54 Å. In the case of EGF domain IV, the average RMSD was 3.26 Å. In Pb28, EGF domain IV was larger in comparison to Pvs25, hence calculating a maximum RMSD was not possible in the case of EGF domain IV. Ramachandran plot statistics of the Pb28 protein models show that only one residue, Asn96, is present in the disallowed regions (Fig. 3). This residue is present just before a disulphide-bonded cysteine, and a change in its conformation would lead to breakage of the disulphide bond; it was therefore decided to keep this residue as it is and no rotamers were selected. ProsaWeb analysis revealed that the Pb28 model fell within the permitted regions, with a Z-score of −5.42 (Fig. 4a), and the model obtained was found to be energetically favourable, as represented by the energy plot in Fig. 4b.

Electrostatic representation of Pb28

The two faces of the Pb28 molecule expose most of the positively charged residues on the surface (Fig. 5). In contrast to the template protein Pvs25, Pb28 molecule revealed an overall positively charged surface with a total charge of +3.00. The positively charged EGF domain II (Fig. 5a) interacts with the negatively charged patch present in the scFv of 13.1 TBmAb (see Fig. 9). Edge II was found to be significantly positively charged in the front view surface as shown in the Fig. 5a, whereas, on the opposite surface, EGF domain III showed maximum positive residues on the surface (Fig. 5b). Negatively charged residues were few and were concentrated primarily near the two N-terminal and C-terminal poles of the molecule. Pb28 is positively charged (total charge=+3.00) whereas the template the Pvs25 was negatively charged (total charge= −10.00). There was a minimal hydrophobic core with comparatively few residues that are buried; most residues are thus solvent-accessible. The positively charged residues were situated mainly on the surface of the molecule, with a relatively very highly accessible molecular surface, indicating

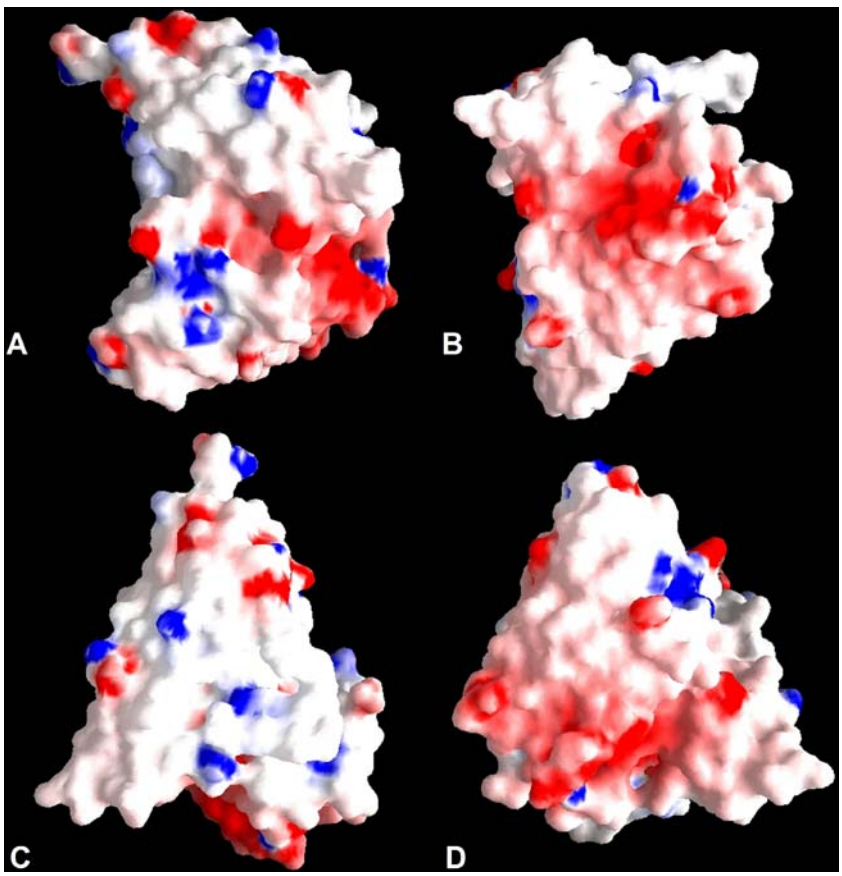

Fig. 9a–d Electrostatic representation of the theoretical structural model of scFv of TBmAb 13.1 prepared with the help of GRASP. Surface potential was taken from −10kT (*red*) to +10kT (*blue*). **a** Front view of Pb28, **b** opposite face after 180° rotation. **c**, **d** Top and bottom views of the molecule. ProsaWeb, PROCHECK and WhatIF programs were used for statistical evaluation of the final model. As represented by the different views of the molecule, negatively charged residues are present at the junction of the heavy and light chains. The total charge on the surface of the molecule= −5.00

the probable significance of these residues in the interaction of the parasite with the host cell receptors. Cysteines were present in the core region of the Pb28 molecule with the least relative surface exposure, which explains why cysteines are conserved in all members of the P28 family of ookinete surface proteins. In order to identify the residues important in interaction of ookinetes with the midgut cells of the mosquito, we compared the theoretical model of Pb28 with other functionally similar proteins of the P28 family. A similar cysteine arrangement and high sequence and structural identity between members of this family have been observed. Despite the fact that P28 proteins have similar functions, and all of them are thought to have four EGF domains, they differed in the loop regions of their EGF domains, and it is probably this feature that explains why they are able to recognise different antibodies and molecules. In addition, it was striking that, despite the high similarity in sequence, area and even volume of the molecules, the molecular surface charge distribution differed significantly. These results indicate the possible variable functions of these molecules inside the mosquito midgut.

The 13.1 antibody model

A model of TBmAb 13.1 was developed with the help of the WAM server (Fig. 6a) [36] and was further refined with the help of Modeller and ModLoop programs as shown in Fig. 6b. The refinement led to major changes in loop conformations as indicated in Fig. 6c. This antibody has been reported as an efficient transmission-blocking antibody against Pb28 protein [52, 53]. The final model obtained was evaluated with the help of Procheck, WhatIF and ProsaWeb servers. Ramachandran plot statistics of the antibody model showed that no residue was present in the disallowed regions (Fig. 7). ProsaWeb analysis also revealed that the Z-score (−6.39) of the 13.1 antibody model falls within the permitted regions (Fig. 8a) and the model obtained is energetically favourable (Fig. 8b). To show the charge distribution on the surface of the antibody molecule, an electrostatic presentation of the molecule was made with the help of the GRASP program [54] as shown in Fig. 9. The antibody clearly has a negatively charged surface (total charge=−5.0) in contrast to the positively charged Pb28

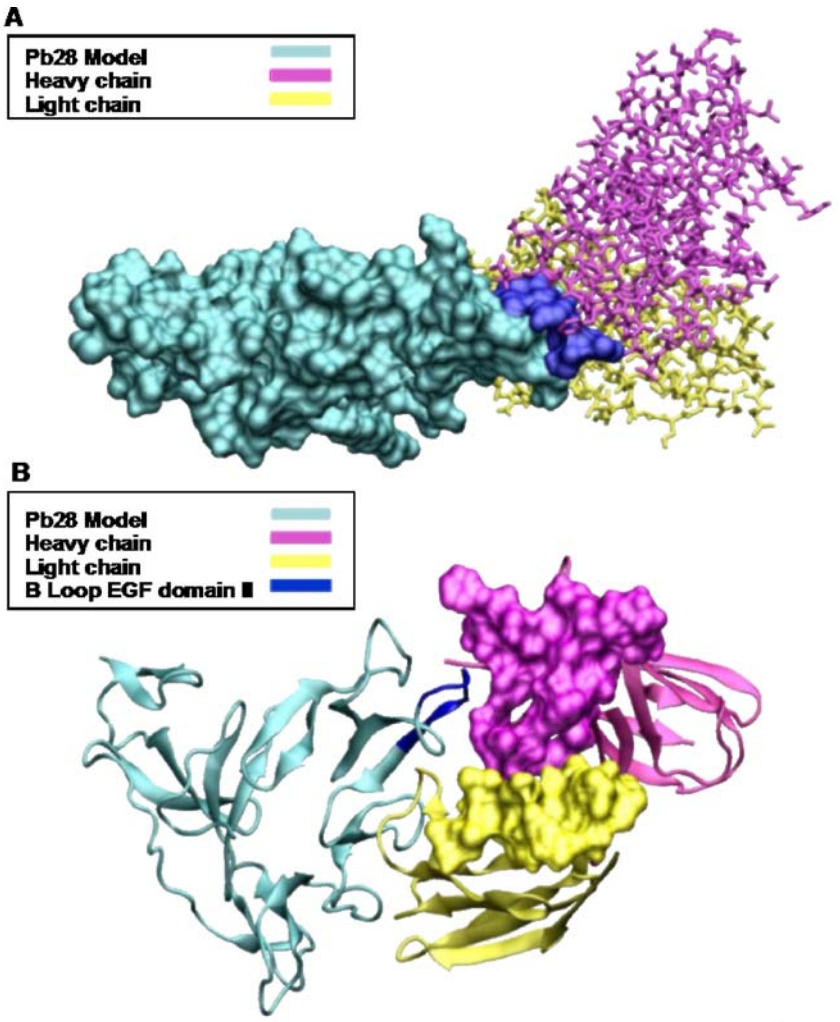

Fig. 10a,b Model of scFv region of TBmAb 13.1 (*yellow* light chain, *magenta* heavy chain) and the interaction of this region with the Pb28 model (*cyan*). The B loop of EGF domain II (as specified by Spano et al. [23]), i.e. GLEKAFVC is shown in *dark blue*. The EGF domain II of Pb28 interacts with both the light and heavy chains of scFv 13.1, forming two hydrogen bonds each. The light chain makes 136 non-bonded contacts whereas the heavy chain forms 80 non-bonded contacts. **a** ScFv of TBmAb 13.1 shown in liquorice presentation and Pb28 model shown in surface presentation. **b** Complementarity determining regions (CDRs) of TBmAb 13.1 shown in surface representation, with the Pb28 model and rest of the scFv shown in cartoon representation. The B loop of EGF domain II, which interacts with TBmAb 13.1, is shown in *dark blue* cartoon representation

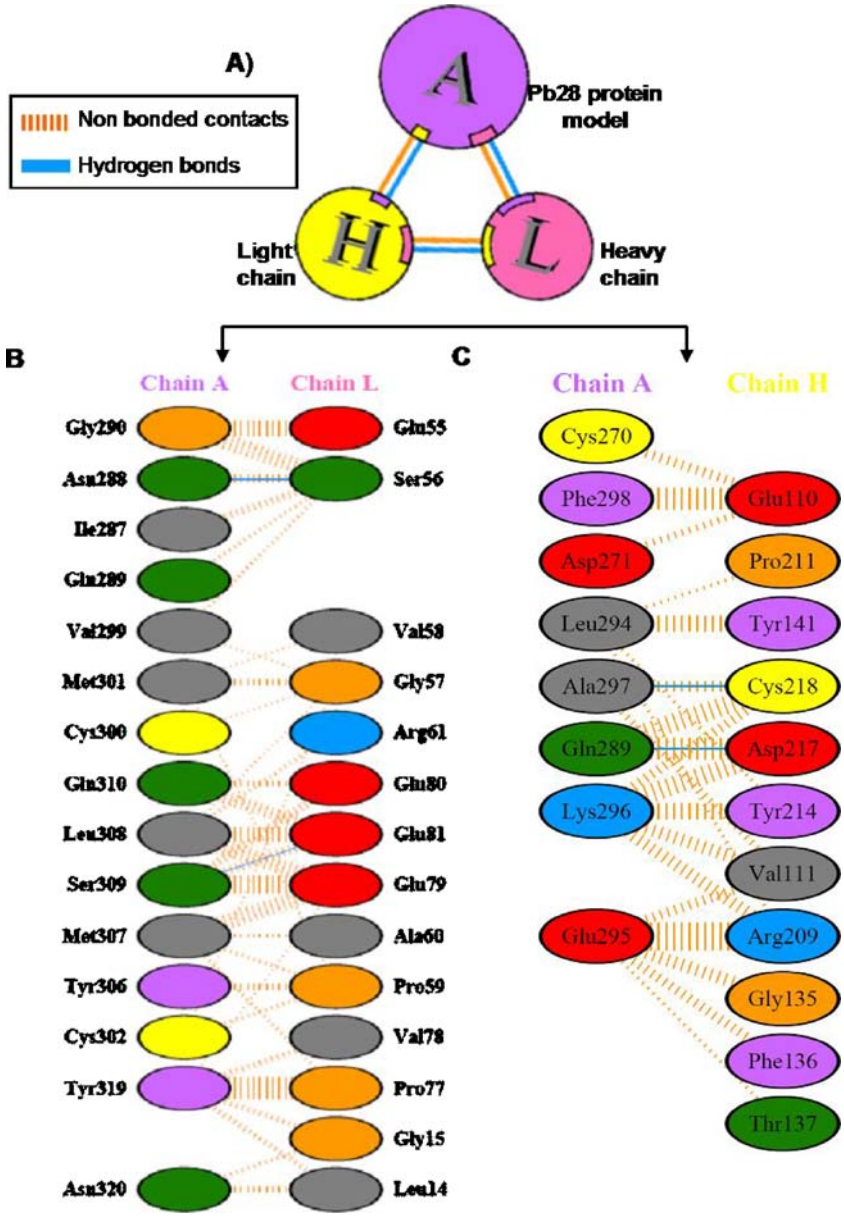

Fig. 11 **a** Schematic representation of the Pb28 protein (*A*) interaction with heavy (*H*) and light chain (*L*) of TBmAb 13.1. **b** Residue–residue interaction details of Pb28 protein and the light chain of scFv 13.1 TBmAb. **c** Residue–residue interaction details of the heavy chain of scFv 13.1 with Pb28 protein. *Vertical orange lines* Non-bonded interactions, *blue lines* hydrogen bonds

Table 2 Pb28 and transmission-blocking monoclonal antibody (TBmAb) 13.1 interaction interface statistics

Chain	No. of interface residues	Interface area (Å^2)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts				
A-H	8:11	386:368	-	-	2	80				
A-L	15:14	665:718	-	-	2	136				
Hydrogen bond formation between Pb28 protein and scFv TBmAb 13.1										
S. No	<— A T O M 1 —>				<— A T O M 2 —>					
	Atom no.	Atom name	Res. name	Res. no. & Chain	Atom no.	Atom name	Res. name	Res. no. and chain	Bond length	
Hydrogen bonds formed by light chain										
1	2,737	O	Asn	288(A)	491	OG	Ser	56(L)	2.72 Å	
2	2,928	O	Ser	309(A)	704	N	Glu	81(L)	2.67 Å	
Hydrogen bonds formed by heavy chain										
1	2,745	NE2	Gln	289(A)	2,074	O	Asp	217(H)	2.73 Å	
2	2,815	N	Ala	297(A)	2,079	SG	Cys	218(H)	2.82 Å	

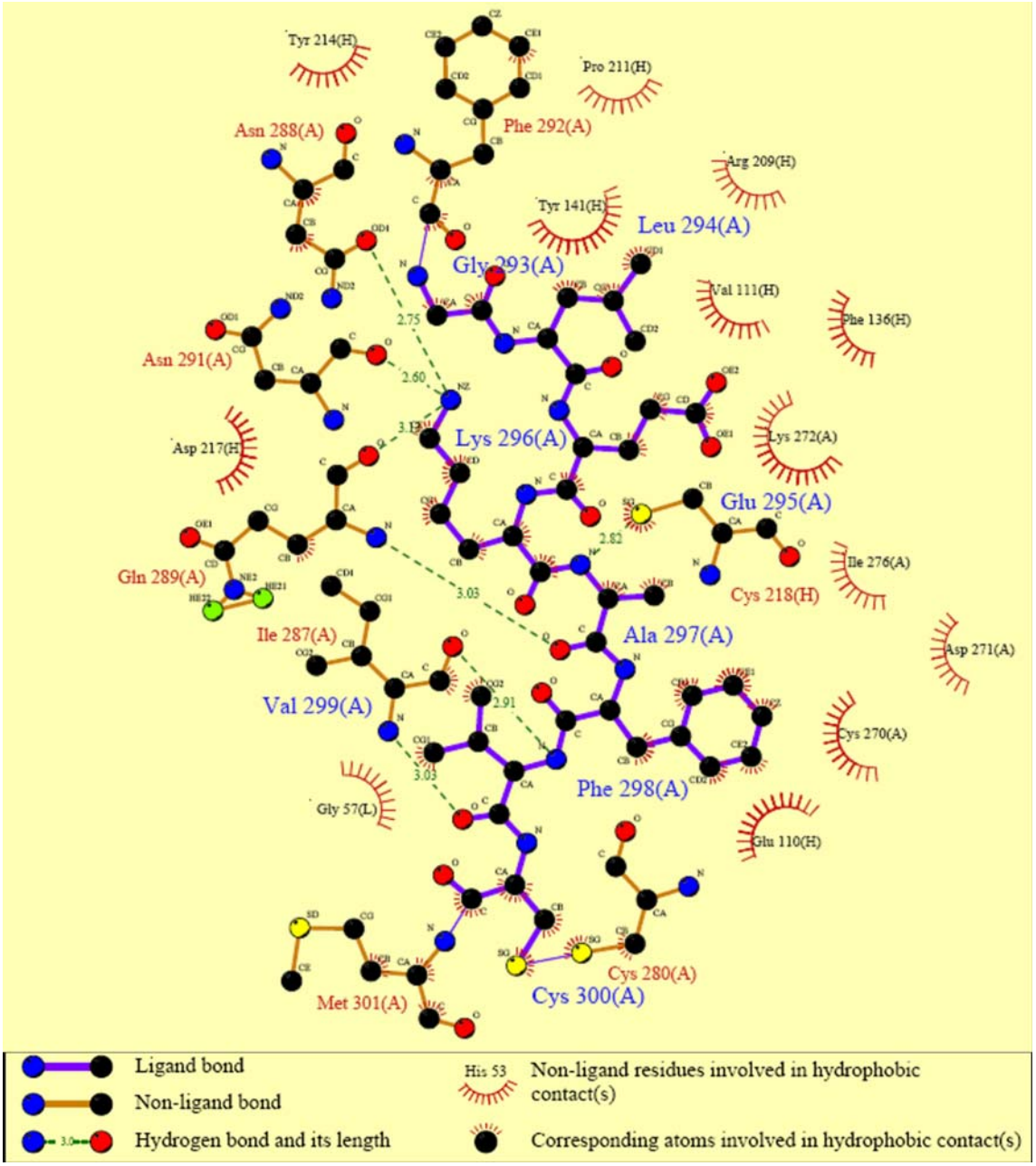

Fig. 12 Schematic representation of the interaction of the B loop (GLEKAFVC) of EGF domain II of Pb28 protein (*A*) with the heavy (*H*) and light (*L*) chains of TBmAb 13.1. The figure was generated using the program Ligplot

protein. A large negatively charged patch is present on the surface of the model of scFv of TBmAb 13.1 (Fig. 9), indicating its potential affinity towards positively charged molecules.

Docking studies

We were further interested in checking whether scFv of TBmAb 13.1 interacts with the Pb28 model. To gain insight into this important interaction, which has been shown to block the transmission of malaria [23, 52, 53], we docked the Pb28 model and the model of the TBmAb 13.1 scFv with the help of the ZDOCK program. The models obtained were refined and re-ranked with the help of the RDOCK program. The top 30 complexes obtained were grouped into clusters according to their interactions. Non-specific interactions were eliminated and only those complexes that showed the interaction of Pb28 protein to the CDRs of the scFv of TBmAb 13.1 were selected. In 1996, Spano et al. mapped the immunogenic loop of this transmission-blocking antibody to the B loop (GLEKAFVC) of EGF domain II of Pb28 protein (then named Pbs21 protein) [23]. Other

important biochemical studies in this field were done by Yoshida et al. in 1999 and 2001 [52, 53]. Taking this information into account, complexes showing interaction with EGF domain II were selected. All the complexes interacting with EGF domain II were taken and checked for an antigen–antibody type of interaction. Complexes representing interactions with CDRs were selected, and the best probable complex was then chosen on the basis of maximum buried surface area. It was found that EGF domain II of the Pb28 model interacted with the light and heavy chains of the scFv region of TBmAb 13.1 (Fig. 10), forming two hydrogen bonds each. The heavy chain formed 80 whereas the light chain formed 136 non-bonded contacts along with the four hydrogen bonds already mentioned (Fig. 11, Table 2). Buried surface area between Pb28 protein and scFv of TBmAb 13.1 was calculated as 1,940.3 $Å^2$. Figures 11 and 12 represent the detailed interactions between the two proteins at the residue level. Figure 12 was created with the help of LIGPLOT software [55]. A similar complex of template Pvs25 protein with its TBmAb 2A8 is available, where only the light chain of the 2A8 TBmAb participates in complex formation, whereas in the case of Pb28 and TBmAb 13.1 both light and heavy chains of the scFv participate equally in complex formation. This study clearly demonstrates that variability in loop regions leads to specificity of the monoclonal antibodies and therefore each ookinete surface protein is able to recognise its specific antibody.

Conclusion

In this study, we have carried out a detailed bioinformatics analysis of the sequence and structural features of ookinete surface protein Pb28 in order to gain an understanding of the functional aspects of this protein, which belongs to the recently characterized P28 family of ookinete surface proteins. Comparative studies between the theoretical model of Pb28 (modelled here) and the functionally similar ookinete surface protein Pvs25, suggest the importance of the variations in the loop regions. These loop regions are thought to be responsible for the varied molecular recognition among the P25 and P28 proteins. These studies indicate the probable differences in the active sites of these proteins, and may help explain why EGF domain II of Pb28 interacts with TBmAb 13.1. Our study has drawn attention to an functionally important region, in particular of Pb28—a typical member of the P28 family, and provides a pointer to experimental work to validate the observations based on in silico analysis made here. It has been shown earlier by Spano et al. [23] that the GLEKAFVC sequence on the B loop of EGF Domain II of Pb28 interacts with TBmAb 13.1, which provides experimental support for the results of our computational study. The present study also demonstrates that the interaction of ookinete surface proteins with their respective TBmAb is likely to differ in different P28 proteins, and that loops are the major areas where the specificity of transmission-blocking antigen–antibody interaction probably lies.

Acknowledgements Sharma B. thanks Council of Scientific and Industrial Research for providing fellowship. We thank Dr. Subha Rao for helping in protein–protein interaction studies, Dr. Andrew Lynn for help regarding CHARMM, H. R. Kushwaha, Centre for Computational Biology and Bioinformatics, and J. Ramanathan, School of Life Sciences, Jawaharlal Nehru University, for suggestions and discussions. We thank Dr. Zhiping Weng, Associate Professor, Bioinformatics program, Department of Biomedical Engineering, Boston University for help regarding ZDOCK and RDOCK programs. Sharma B. thanks her supervisor Dr. A. K. Saxena for allowing me to design and implement the protocols for this study independently.

References

1. Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI (2005) The global distribution of clinical episodes of *Plasmodium falciparum* malaria. Nature 434:214–217 doi:10.1038/nature03342
2. World Malaria Report (2005). Geneva. RBM/WHO/UNICEF
3. Kumar N, Carter R (1985) Biosynthesis of two stage-specific membrane proteins during transformation of *Plasmodium gallinaceum* zygotes into ookinetes. Mol Biochem Parasitol 14:127–139 doi:10.1016/0166-6851(85)90032-5
4. Vermeulen AN, Ponnudurai T, Beckers PJA, Verhave JP, Smits MA et al (1985) Sequential expression of antigens on sexual stages of *Plasmodium falciparum* accessible to transmission blocking antibodies in the mosquito. J Exp Med 162:1460–1476 doi:10.1084/jem.162.5.1460
5. Paton MG, Barker GC, Matsuoka H, Ramesar J, Janse CJ et al (1993) Structure and expression of a post-transcriptionally regulated malaria gene encoding a surface protein from the sexual stages of *Plasmodium berghei*. Mol Biochem Parasitol 59:263–275 doi:10.1016/0166-6851(93)90224-L
6. Fries HC, Lamers MB, Van Deursen J, Ponnudurai T, Meuwissen JH (1990) Biosynthesis of the 25-kDa protein in the macrogametes/zygotes of *Plasmodium falciparum*. Exp Parasitol 71:229–235 doi:10.1016/0014-4894(90)90025-8
7. Winger LA, Tirawanchai N, Nicholas NJ, Carter HE, Smith JE et al (1988) Ookinete antigens of *Plasmodium berghei.* Appearance on the zygote surface of an Mr 21 kD determinant identified by transmission-blocking monoclonal antibodies. Parasite Immunol 10:193–207 doi:10.1111/j.1365-3024.1988.tb00214.x
8. Sinden RE, Winger L, Carter EH, Hartley RH, Tirawanchai N et al (1987) Ookinete antigens of *Plasmodium berghei*: a light and electron-microscope immunogold study of expression of the 21 kDa determinant recognized by a transmission-blocking antibody. Proc R Soc Lond B Biol Sci 230:443–458 doi:10.1098/rspb.1987.0028
9. Duffy PE, Pimenta P, Kaslow DC (1993) Pgs28 belongs to a family of epidermal growth factor-like antigens that are targets of malaria transmission-blocking antibodies. J Exp Med 177:505–510 doi:10.1084/jem.177.2.505
10. Tomas AM, Margos G, Dimopoulos G, Van Lin LH, Koning-Ward TF et al (2001) P25 and P28 proteins of the malaria ookinete surface have multiple and partially redundant functions. EMBO J 20:3975–3983 doi:10.1093/emboj/20.15.3975
11. Kaslow DC, Quakyi IA, Syin C, Raum MG, Keister DB et al (1988) A vaccine candidate from the sexual stage of human

malaria that contains EGF-like domains. Nature 333:74–76 doi:10.1038/333074a0
12. Tsuboi T, Kaslow DC, Gozar MM, Tachibana M, Cao YM et al (1998) Sequence polymorphism in two novel *Plasmodium vivax* ookinete surface proteins, Pvs25 and Pvs28 that are malaria transmission-blocking vaccine candidates. Mol Med 4:772–782
13. Appella E, Weber IT, Blasi F (1998) Structure and function of epidermal growth factor like regions in proteins. FEBS Lett 231:1–4 doi:10.1016/0014-5793(88)80690-2
14. Carter R, Graves PM, Quakyi IA, Good MF (1989) Restricted or absent immune responses in human populations to *Plasmodium falciparum* gamete antigens that are targets of malaria transmission blocking antibodies. J Exp Med 169:135–147 doi:10.1084/jem.169.1.135
15. Yoshida S, Matsuoka H, Luo E, Iwai K, Arai M et al (1999) A single-chain antibody fragment specific for the *Plasmodium berghei* ookinete protein Pbs21 confers transmission blockade in the mosquito midgut. Mol Biochem Parasitol 104:195–204 doi:10.1016/S0166-6851(99)00158-9
16. Kaslow DC (1996) Transmission-blocking vaccine. In: Hoffman SL (ed) Malaria vaccine development: a multi-immune response approach. American Society for Microbiology, Washington, DC, pp 181–227
17. Malkin EM, Durbin AP, Diemert DJ, Sattabongkot J, Wu Y et al (2005) Phase-I vaccine trial of Pvs25: a transmission blocking vaccine for *Plasmodium vivax* malaria. Vaccine 23:3131–3138 doi:10.1016/j.vaccine.2004.12.019
18. Han YS, Thompson J, Kafatos FC, Barillas-Mury C (2000) Molecular interactions between *Anopheles stephensi* midgut cells and *Plasmodium berghei*: the time bomb theory of ookinete invasion of mosquitos. EMBO J 19:6030–6040 doi:10.1093/emboj/19.22.6030
19. Tsuboi T, Cao YM, Hitsumoto Y, Yanagi T, Kanbara H et al (1997) Two antigens on zygotes and ookinetes of *Plasmodium yoelii* and *Plasmodium berghei* that are distinct targets of transmission blocking immunity. Infect Immun 65:2260–2264
20. Sivasubramanian A, Chao G, Pressler HM, Witturp KD, Gray JJ (2006) Structural model of the mab 806-EGFR complex using computational docking followed by computational and experimental mutagenesis. Structure 14:401–414 doi:10.1016/j.str.2005.11.022
21. Gray JJ (2006) High resolution protein–protein docking. Curr Opin Struct Biol 16:183–193 doi:10.1016/j.sbi.2006.03.003
22. Wiehe K, Pierce B, Mintseris J, Tong WW, Anderson R et al (2005) ZDOCK and RDOCK performance in CAPRI rounds 3, 4, and 5. Proteins 60:207–213 doi:10.1002/prot.20559
23. Spano F, Matsuoka H, Ozawa R, Chinzei Y, Sinden RE (1996) Epitope mapping on the ookinete surface antigen Pbs21 of *Plasmodium berghei*: identification of the site of binding of transmission-blocking monoclonal antibody 13.1. Parassitologia 38:559–563
24. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD et al (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784–3788 doi:10.1093/nar/gkg563
25. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
26. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN et al (2000) The protein data bank. Nucleic Acids Res 28:235–242 doi:10.1093/nar/28.1.235
27. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 31:3497–3500 doi:10.1093/nar/gkg500
28. Spalding JB, Lammers PJ (2004) BLAST Filter and GraphAlign: rule-based formation and analysis of sets of related DNA and protein sequences. Nucleic Acids Res 32:W26–W32 doi:10.1093/nar/gkh459
29. Stothard P (2000) The sequence manipulation suite: javascript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28:1102–1104
30. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815 doi:10.1006/jmbi.1993.1626
31. Saxena AK, Singh K, Long C, Garboczi DN (2004) Preparation, crystallization and preliminary X-ray analysis of complex between antibody and *Plasmodium vivax* sexual stage 25 kDa protein Pvs25 and malaria transmission-blocking antibody Fab fragment. Acta Crystallogr D Biol Crystallogr D60:2054–2057 doi:10.1107/S0907444904021584
32. Saxena AK, Saul A, Garboczi DN (2004) Crystallization and preliminary X-ray analysis of the *Plasmodium vivax* sexual stage 25 kD protein, Pvs25, a transmission-blocking vaccine candidate for malaria. Acta Crystallogr D Biol Crystallogr D60:706–708 doi:10.1107/S0907444904001398
33. Saxena AK, Singh K, Su HP, Klein MM, Stower AW et al (2006) Plasmodium P25 and P28, essential proteins for survival of the malaria parasite in the mosquito, are tile-like triangular prism. Nat Struct Mol Biol 13:90–91 doi:10.1038/nsmb1024
34. Fiser A, Sali A (2003) ModLoop: automated modeling of loops in protein structures. Bioinformatics 19:2500–2501 doi:10.1093/bioinformatics/btg362
35. Fiser A, Do RK, Sali A (2000) Modeling of loops in protein structures. Protein Sci 9:1753–1773
36. Whitelegg N, Rees AR (2000) WAM: an improved algorithm for modelling antibodies on the WEB. Protein Eng 13(12):819–824 doi:10.1093/protein/13.12.819
37. Whitelegg N, Rees AR (2004) Antibody variable regions: toward a unified modeling method. Methods Mol Biol 248:51–91
38. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 10:1093
39. Sippl MJ (1993) Recognition of errors in three-dimensional structures of proteins. Proteins 17:355–362 doi:10.1002/prot.340170404
40. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291 doi:10.1107/S0021889892009944
41. Vriend G (1990) WHAT IF: A molecular modeling and drug design program. J Mol Graph 8:52–56 doi:10.1016/0263-7855(90)80070-V
42. Hooft RW, Vriend G, Sander C, Abola EE (1996) Errors in protein structures. Nature 381:272 doi:10.1038/381272a0
43. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) Stereochemical quality of protein structure coordinates. Proteins 12:345–364 doi:10.1002/prot.340120407
44. Humphrey W, Dalke A, Schulten K (1996) VMD—visual molecular dynamics. J Mol Graph 14:33–38 doi:10.1016/0263-7855(96)00018-5
45. Chen R, Weng Z (2003) A novel shape complementarirty scoring function for protein-protein docking. Proteins 51:397–408 doi:10.1002/prot.10334
46. Li L, Chen R, Weng Z (2003) RDOCK: Refinement of rigid-body protein docking predictions. Proteins 53:693–707
47. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S et al (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4:187–217 doi:10.1002/jcc.540040211
48. Cavallo L, Kleinjung J, Fraternali F (2003) POPS: A fast algorithm for solvent accessible surface areas at atomic and residue level. Nucleic Acids Res 3:3364–3366 doi:10.1093/nar/gkg601
49. Laskowski RA (2001) PDBsum: summaries and analyses of PDB structures. Nucleic Acids Res 29:221–222 doi:10.1093/nar/29.1.221
50. Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. Nucleic Acids Res 18:6097–6100 doi:10.1093/nar/18.20.6097

51. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) Weblogo: a sequence logo generator. Genome Res 14:1188–1190 doi:10.1101/gr.849004
52. Yoshida S, Ioka D, Matsuoka H, Endo H, Ishii A (2001) Bacteria expressing single-chain immunotoxin inhibit malaria parasite development in mosquitoes. Mol Biochem Parasitol 113:89–96 doi:10.1016/S0166-6851(00)00387-X
53. Yoshida S, Matsuoka H, Luo E, Iwai K, Arai M, Sinden RE et al (1999) A single-chain antibody fragment specific for the *Plasmodium berghei* ookinete protein Pbs21 confers transmission blockade in the mosquito midgut. Mol Biochem Parasitol 104:195–204 doi:10.1016/S0166-6851(99)00158-9
54. Nicholls A, Sharp KA, Honig B (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins: Struct Funct Genet 11:281–296 doi:10.1002/prot.340110407
55. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. Protein Eng 8:127–134 doi:10.1093/protein/8.2.127