

# *Plasmodium vivax* Tryptophan-Rich Antigen PvTRAg33.5 Contains Alpha Helical Structure and Multidomain Architecture

Hema Bora<sup>1</sup>, Sheena Garg<sup>1</sup>, Priyankar Sen<sup>2</sup>, Deepak Kumar<sup>1,3</sup>, Punit Kaur<sup>3</sup>, Rizwan Hasan Khan<sup>2\*</sup>, Yagya D. Sharma<sup>1\*</sup>

**1** Department of Biotechnology, All India Institute of Medical Sciences, New Delhi, India, **2** Interdisciplinary Biotechnology Unit, Aligarh Muslim University, Aligarh, India, **3** Department of Biophysics, All India Institute of Medical Sciences, New Delhi, India

## Abstract

Tryptophan-rich proteins from several malarial parasites have been identified where they play an important role in host-parasite interaction. Structural characterization of these proteins is needed to develop them as therapeutic targets. Here, we describe a novel *Plasmodium vivax* tryptophan-rich protein named PvTRAg33.5. It is expressed by blood stage(s) of the parasite and its gene contains two exons. The exon 1 encodes for a 23 amino acids long putative signal peptide which is likely to be cleaved off whereas the exon 2 encodes for the mature protein of 252 amino acids. The mature protein contains B-cell epitopes which were recognized by the human immune system during *P. vivax* infection. The PvTRAg33.5 contains 24 (9.5%) tryptophan residues and six motifs whose patterns were similar among tryptophan-rich proteins. The modeled structure of the PvTRAg33.5 consists of a multidomain architecture which is stabilized by the presence of large number of tryptophan residues. The recombinant PvTRAg33.5 showed predominantly  $\alpha$  helical structure and alpha helix to beta sheet transition at pH below 4.5. Protein acquires an irreversible non-native state at temperature more than 50°C at neutral pH. Its secondary and tertiary structures remain stable in the presence of 35% alcohol but these structures are destabilized at higher alcohol concentrations due to the disturbance of hydrophobic interactions between tryptophanyl residues. These structural changes in the protein might occur during its translocation to interact with other proteins at its final destination for biological function such as erythrocyte invasion.

**Citation:** Bora H, Garg S, Sen P, Kumar D, Kaur P, et al. (2011) *Plasmodium vivax* Tryptophan-Rich Antigen PvTRAg33.5 Contains Alpha Helical Structure and Multidomain Architecture. PLoS ONE 6(1): e16294. doi:10.1371/journal.pone.0016294

**Editor:** Georges Snounou, Université Pierre et Marie Curie, France

**Received:** August 17, 2010; **Accepted:** December 19, 2010; **Published:** January 20, 2011

**Copyright:** © 2011 Bora et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by grants from Department of Biotechnology (to YDS and RHK), Indian Council of Medical Research (to YDS and PK), Department of Science and Technology (to RHK) and Council for Scientific and Industrial Research (to RHK). Junior Research Fellowship (to HB) and Senior Research Fellowships (to PS and SG) from Department of Biotechnology are gratefully acknowledged. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: ydsharma\_aaims@yahoo.com (YDS); rizwanhkhan@hotmail.com (RHK)

## Introduction

Malaria remains uncontrolled to date and requires effective antimalarial drugs and vaccines [1]. Both, vaccine and drug development strategies require identification and characterization of the target molecules. The parasite molecules which interact with the host molecules are good drug or vaccine targets. This is because any molecule or antibody which can disrupt this host - parasite molecular interaction can be used as therapeutic reagent to interrupt the disease progression. Enormous efforts are therefore being made to identify such parasite molecules. However, in comparison to the *Plasmodium falciparum*, only fewer vaccine/drug target molecules of *P. vivax* have been identified because of its non-cultivable nature although this parasite affects large number of people in tropical countries and sometime can cause severe complications in humans [1,2,3].

Tryptophan-rich proteins from several *Plasmodium* species have been identified. They play an important role in the host-parasite interaction [4,5,6,7,8]. Some of them may be involved in reinvasion of the host erythrocytes by the parasite

[5]. The synthetic peptides derived from the *P. falciparum* tryptophan-threonine rich antigen (PfTryThrA) were able to inhibit the invasion of the erythrocyte by the merozoite [9]. These proteins across the malarial species contain positional conservation of the tryptophan residues which could be functionally important for the parasite. The *P. vivax* genome encodes for more number of such tryptophan-rich proteins than any other malarial species (<http://www.plasmodb.org>). These proteins need to be characterized so as to develop them as the drug or vaccine targets. Recently, we have described immunological characterization of some of these *P. vivax* proteins and reported that they were highly immunogenic in humans [10,11,12,13,14]. The tryptophan-rich domains of all these proteins were also conserved in the parasite population. Unfortunately, so far none of the tryptophan - rich antigen of any of the *Plasmodium* species has been structurally characterized. Here, we describe cloning, expression, human humoral immune response, physico-chemical characterization and molecular modeling of a novel 33.5 kDa *P. vivax* tryptophan-rich protein called PvTRAg33.5 which showed a significant homology to the PfTryThrA of *P. falciparum*.

## Materials and Methods

### Reverse transcription PCR

Previously isolated RNA from the *P.vivax* - infected blood was used for a two step reverse transcription PCR with oligodT primers and random decamers using RETROScript® (Ambion, Inc. Austin, TX USA) as per the manufacturer's instructions. The exon 1 and part of the exon 2 was amplified from the cDNA as well as the genomic DNA (to detect intron) using the gene specific primers F1 (forward) (5'-ATGGTTGCCTTATTACCAATTTCA-3') and R1 (reverse) (5'-AGTTTTCCAGTCAGCATCATTC-3'). A 2 µL of 1:5 dilution of the primary PCR product was used as a template for the semi-nested PCR using primers F2 (5'-TCTGCGGCTTACCTTTTAAGCAAC-3') and R1 (mentioned above). A 50 µL PCR reaction mixture included the Taq DNA polymerase (1U) with 1X PCR buffer, 0.5 mM dNTPs, 1.5 mM MgCl<sub>2</sub>, 0.5 µM of each primer and 10 ng of the template DNA. A control reaction was also included in which no reverse transcriptase was added to rule out the possibility of the DNA contamination after DNase treatment in the cDNA synthesis. The PCR cycling conditions included initial denaturation at 94°C for 5 min followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 56°C (for F1 and R1) or 58°C (for F2 and R1) for 30 sec and extension at 72°C for 30 sec. Final extension was carried out at 72°C for 5 minutes. The amplified products were purified from the gel using AccuPrep® gel purification kit (Bioneer Corporation, Korea) and sequenced from both the strands using F2 and R1 primers on the ABI Prism A310 Genetic Analyzer (PE Applied Biosystems, CA, USA) as described before [15].

### Cloning, expression and purification of the recombinant PvTRAg33.5

The exon 2 of *pvtrag33.5* gene was PCR amplified from the *P.vivax* DNA using primers, 5'-TGTAGTCGACTCAAAGCGCAGTAG-3' (forward) and 5'TTGTTCTAATTGAGTCTAGAATTC3' (reverse) having *SaI* and *XbaI* sites (underlined) respectively. The cycling parameters for PCR included initial denaturation at 94°C for 5 minutes followed by 35 cycles of denaturation at 94°C for 15 sec, annealing at 55°C for 30 sec, extension at 68°C for 1 minute and final extension at 68°C for 15 minutes. The PCR product was cloned into the pGEMT vector and sequenced with the universal primers. The fragment was transferred to the expression vector pPROExHT b (Invitrogen Life Technologies, Carlsbad, CA, USA). The recombinant PvTRAg 33.5 was expressed in *E.coli* BL21 (DE3) codon plus- RP and induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 22°C for 4 hours.

The recombinant protein from the culture pellet was purified using the Immobilized Metal Affinity Chromatography on the Ni<sup>2+</sup> NTA agarose column according to the manufacturer's instructions (Qiagen, GmbH, Hilden, Germany). It was further purified by the Anion Exchange High Performance Liquid Chromatography (Biologic Duo flow, Biorad laboratories, Hercules, CA, USA). The dialyzed Ni<sup>2+</sup> NTA purified protein was injected in 50 µl sample loop from where it was loaded on to the top of 1.3 ml UNO Q1 Anion exchange column (Biorad laboratories, Hercules, CA, USA) which had been pre-equilibrated at a flow rate of 2 ml/minute with ten column volumes of buffer A (20 mM Tris, pH 8.0). The bound proteins were then eluted by the NaCl gradient (0–500 mM) by employing buffer B (20 mM Tris pH 8.0 and 0.5 M NaCl) at a flow rate of 2 ml/minute over a period of ten minutes. One ml fractions each were collected and detected at a wavelength of 280 nm by the deuterium lamp in the Quad Tec UV-Vis detector integrated with the HPLC unit. The

homogeneity of the purified recombinant PvTRAg33.5 protein was confirmed by the SDS-PAGE as described earlier [15].

### Direct binding Enzyme-Linked Immunosorbent Assay (ELISA)

The ELISA was performed as described earlier [15,16] using 50 *P. vivax*-infected and 39 uninfected sera samples at a 1:100 dilution and 100 ng of the purified recombinant PvTRAg33.5 protein. The goat anti-human IgG horseradish peroxidase conjugate at a 1:400 dilution (Pierce Chemical Company, Rockford, IL, USA) was used as secondary antibody. The mean OD plus 3 standard deviations of the normal healthy controls was used as the cutoff value for a positive response.

### Physico-chemical characterization of recombinant PvTRAg33.5

All the measurements were carried out at the room temperature, unless stated. The concentration of protein samples were determined spectrophotometrically from the extinction coefficient reported at 280 nm.

**Circular dichroism measurements.** The circular dichroism (CD) was measured with a JASCO J-815 spectropolarimeter calibrated with ammonium d-10 camphorsulfonate. A cell of path length 0.1 cm was used for scanning between 250–195 nm and a cell of path length 1.0 cm was used for scanning between 300–250 nm. The results were expressed as the mean residue ellipticity (MRE) in deg.cm<sup>2</sup>.dmol<sup>-1</sup>, which is defined as:

$$\text{MRE} = \theta_{\text{obs}}(\text{mdeg}) / (10 \times n \times C_p \times l) \quad (\text{i})$$

Where  $\theta_{\text{obs}}$  is the observed ellipticity in degrees,  $n$  is the number of peptide bonds per molecule,  $C_p$  is the molar concentration, and  $l$  is the length of light path in cm. The  $\alpha$ -helical content of proteins was calculated from the MRE value at 222 nm (MRE<sub>222</sub>) using the following equation [17]:

$$\text{Percent } \alpha\text{-helix} = [(MRE_{222} - 2340) / 30300] \times 100 \quad (\text{ii})$$

CD data have also been analyzed by online available software, K2d [18].

**Fluorescence measurements.** The fluorescence spectra were recorded with a Hitachi F-3500 spectrofluorophotometer in a 1.0 cm path length quartz cell. Samples containing different concentrations of the organic solvent were equilibrated at room temperature for 30 minutes before recording of the fluorescence measurements. The excitation wavelength was 295 nm and the emission was recorded from 300 to 400 nm. The ANS (8-anilino-1-naphthalene-sulfonate) binding was measured by fluorescence emission with excitation at 380 nm and emission was recorded from 400 to 600 nm. Typically, the ANS concentration was 50 times excess of protein concentration (7 µM).

### Sequence analysis

The PvTRAg33.5 full length sequence was analyzed using signal prediction servers, Ipsort [19], and SOSUI [20] utilizing different prediction methods to determine the presence of the signal peptide region. The transmembrane helical region analysis of the PvTRAg33.5 (exon2) was carried out using TMHMM [21] and TM Pred [22] servers. InterPro scan and ProScan [23] analysis was carried out for the motif and/or the pattern prediction. The 91 similar sequences obtained from the PlasmoDB

database (<http://www.plasmodb.org>) were input into MEME [24] to identify conserved domains, aligned in MAST [25] and fed into meta-MEME [26]. These motifs were then combined into a single, motif-based Hidden Markov Model (HMM) and used for searching homologs. The complete sequence of the PvTRAg33.5 was used for the identification of motifs.

### Secondary structure prediction

The prediction servers incorporating different algorithms were used for the secondary structure prediction to minimize the error in prediction (supplementary material S1). The consensus secondary structure was derived employing the probability of the predicted sequence and the type of secondary structure prediction from these methods. This was compared with the data derived from the experimental CD studies.

### Model development and verification

The initial step for the model development is the identification of a structural homology. Threading technique or fold recognition that assign folds to target sequences with very low sequence identity to known structures can also be employed for the template generation. The threading servers utilized to build initial models included PHYRE [27], Fugue [28], PredictProtein [29], SAM-T08 [30] and iTASSER [31,32,33]. These servers exploit various machine learning fold recognition techniques like recognizing distant homologues by sequence-structure comparison, profile-profile alignments and protein folding potentials. The final templates so obtained were used to build a homology model with Discovery Studio v1.7 ([www.accelrys.com](http://www.accelrys.com)). Energy minimization of generated model was by CHARMM [34] force field, a implicit distance-dependent dielectric constant and a non-bonded atom cut-off 12 angstrom, by the Adopted Basis Newton-Raphson energy minimization. Two thousand steps of energy minimization of r.m.s. gradient of 0.05 Kcal/(mol × Angstrom) was performed.

The model was checked for its accuracy and correctness by the Ramachandran plot to determine the amount of misplaced phi and psi angles. The stereochemical environment, the conformation of the backbone and side chains was analysed using PROCHECK [35].

### Molecular Dynamics

The overall stability of the PvTRAg33.5 model was checked by a molecular dynamics simulation using the CHARMM [34] module on a fully hydrated model. In the first step of energy minimization, the backbone of model was kept flexible. The minimized, hydrated complex was then subjected to a molecular dynamics simulation in three stages. In the first stage, the temperature of the system was raised from 0 to 300 K for 2 pico sec. The system was then equilibrated for 20 sec, and the final production run was carried out over another 100 picoseconds.

### Accession numbers

The nucleotide sequences described in the paper have been deposited in the Genbank under the accession numbers FJ481116 for the genomic DNA and FJ481117 for the cDNA derived from the *pvtrag33.5* gene of the Indian *P. vivax* isolate.

## Results

### Sequence analysis of the *pvtrag33.5* gene

Homology searches for the tryptophan-rich antigen of *P. vivax* identified a number of related protein molecules in the *Plasmodium* database (<http://www.plasmodb.org>). We selected a protein named here as PvTRAg33.5 which is of 33.5 kDa mol.

wt and rich in tryptophan residues. This protein showed maximum homology of 56% (33% identity) with the PvTRAg35.2 of *P. vivax* (Fig. S1). Its sequence homology with the *P. falciparum* and *P. yoelii* tryptophan rich proteins PfTryThrA and PypAg3 was 55% (30% identity over a 43-275 aa overlap) and 51% (29% identity over a 20-270 aa overlap), respectively. Twenty out of 24 tryptophan residues of PvTRAg33.5 were positionally conserved among these proteins. Several tryptophan residues were positionally conserved across the tryptophan-rich proteins of *Plasmodium* species (Fig. S1). The *pvtrag33.5* gene contains two exons of 69 bp (exon-1) and 756 bp (exon-2). The exon 1 encodes for a putative signal peptide which is likely to be cleaved away and thus may not be part of the mature folded protein. Hence, only the exon2 encoded protein was used for further characterization and model determination. The ProScan analysis found one putative glycosylation site (Asn 81) and eight putative phosphorylation sites (Ser 4, Ser 98, Thr 129, Ser 177, Thr 179, Ser 217, Ser 243 and Thr 232). The sequences input to MEME yielded the presence of seven different motifs in the full length sequence (Table 1). The tryptophan-rich proteins having more than 30% identity contained all 7 motifs in the order 'M7 M3 M6 M1 M4 M5 M2'. A database search using these clusters with MAST package also did not yield any significant match indicating the absence of orthologues for PvTRAg33.5.

### Expression and purification of the recombinant PvTRAg33.5

We were unable to express the complete cDNA of the *pvtrag 33.5* gene in *E. coli* and thus decided to express the exon 2 coding region and discarded exon 1 as it encodes only for the putative signal peptide which may not be part of the mature protein. The exon 2 encoded protein was found in the inclusion bodies (90%) as well as in the soluble fraction (10%). In order to avoid the in-vitro refolding process, we purified the recombinant protein from the soluble fraction. The recombinant protein was purified on the Ni<sup>2+</sup> NTA agarose column since the expressed protein contained the His tag. Although the protein was purified to a great extent by this column, the minor contaminants were still present in the preparation (lane3 in Fig. 1). These contaminants were removed by the HPLC using an anion exchange column where this protein eluted as a single peak between 440 and 500 mM NaCl concentration (data not shown). This peak was found to contain only a single band of 33.5 kDa on 12% SDS-PAGE (lane 4 in Fig. 1).

### In vivo expression of the PvTRAg33.5 by *P. vivax*

Since the *pvtrag33.5* gene was not annotated in the PlasmoDB sequence database at the time of initiating this work, we wanted to know if this gene was actively expressed by the parasite. The RT-PCR results revealed the presence of the *pvtrag33.5* gene transcript in the parasite (Fig. 2a). The size difference between the cDNA and genomic DNA PCR products not only confirmed the specificity of the RT-PCR but also established the presence of an intron in the gene (Fig. 2a). The sequence analysis of the RT-PCR product and its comparison with the genomic DNA sequence revealed that the size of the intron was 184 bp and the consensus splice site sequence GU/AG at the intron-exon junction (Fig. 2b). That this *pvtrag33.5* gene transcript was translated in to the protein product in the parasite was evident from the fact that majority of the *P. vivax* patients produced antibodies against PvTRAg33.5 during the acute phase of infection (described below).

**Table 1.** Motifs identified from MEME present in the various *Plasmodium* species.

Motif No	Motif regular expression	Motif in PvTRAg33.5
M1	YKS[ND]ILKKSSTW[DN][DE]S[QE]W[EK]EW[IM]KTEGKE[LF][MLI]	YKNYLLKKSEKWNDAWENWANTEMVAHL
M2	[FY]NEWME[ED]SF[IV]NKKW[K]EK[QK]WNVW	YSTWRNDFINRWSEKKKNSI
M3	WKNNEW[KN][NK]W[MK]KKLEx[DE]WkxFN	WKDNEWHNWKLEEDWDSFS
M4	WI[QK]WKNxKix[ES]WL[MS]S[DE]WkxEEDEYWSKW	WNQWQHDKMSSWLSDDWKKVGMAYWDLQ
M5	xKW[KN]ERIN[RK]ExE[EQ]WxNWWKxKENx[YF]	IKWNRNARENIEWSKVVQNKKEYFI
M6	KEK[ED][WL][EN]EW[L]Kx[ML][EQ]NKWMH[YF]NE	KTDELNGWLNLEENKWNFFSG
M7	V5xLxLFLSSAF	VALLPISFFSLSAAY

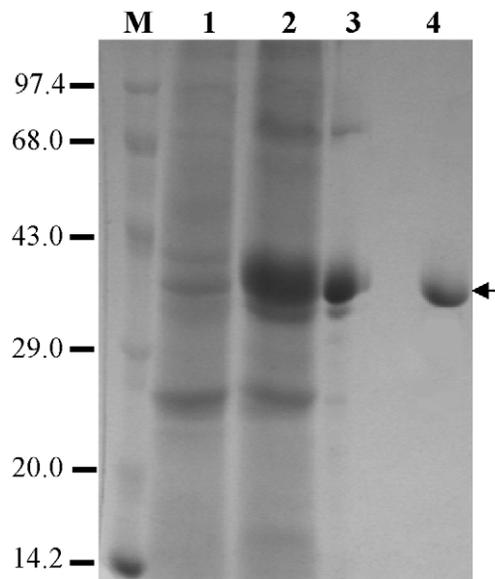
doi:10.1371/journal.pone.0016294.t001

### Humoral immune response against the exon-2 encoded polypeptide of PvTRAg33.5 among a *P.vivax* infected individuals

The natural humoral immune response to the PvTRAg33.5 was confirmed by assessing the reactivity of the exon-2 derived recombinant protein with the antibodies present in each sera sample by the ELISA. The positive responders were defined as those who exhibited an OD<sub>495</sub> higher than 0.55 (mean +3 SD of control sera). Forty seven of the 50 serum samples (seropositivity 94%) from the *P.vivax* infected individuals gave an OD above this cutoff (Fig. 3). The mean  $\pm$  SD values of OD for the *P.vivax* infected and naive individuals were  $1.04 \pm 0.39$  and  $0.25 \pm 0.10$ , respectively. All individuals had different levels of serum antibodies to the PvTRAg33.5.

### Physico-chemical characterization of the PvTRAg33.5

**Effect of pH.** Studies of the far UV CD can be used to quantitatively assess the overall secondary structure content of the protein, as the absorbing group is principally the peptide bond.



**Figure 1. Purification of recombinant PvTRAg33.5.** SDS-PAGE profile of the purification steps. Uninduced and induced culture pellets of the recombinant clone were solubilized in the loading buffer and loaded in lane 1 and 2, respectively. The Ni<sup>2+</sup> NTA column purified preparation was loaded in lane 3. Lane 4 contains the HPLC purified protein. Size of protein bands in the marker lane (M) is indicated on left hand side. The recombinant PvTRAg33.5 band is indicated by an arrow. doi:10.1371/journal.pone.0016294.g001

Figure 4a shows the far UV CD of the PvTRAg33.5 at neutral to low pH. The UV CD spectrum at pH 7.2 shows minima between 208 and 222 nm, a characteristic feature of the  $\alpha$ -helical proteins [36]. Interestingly, the protein did not show any loss of the secondary structure till pH 4.5, but further decrease in pH resulted in the alpha helix to beta sheet transition with a single negative minima near 215 nm and a positive ellipticity at 197 nm [37]. Further, the formation of an isobestic point at 197 nm also indicates the same [38]. Similar results were also observed with increased ammonium sulfate (see section 3.5.3) and ethanol (section 3.5.4) concentrations.

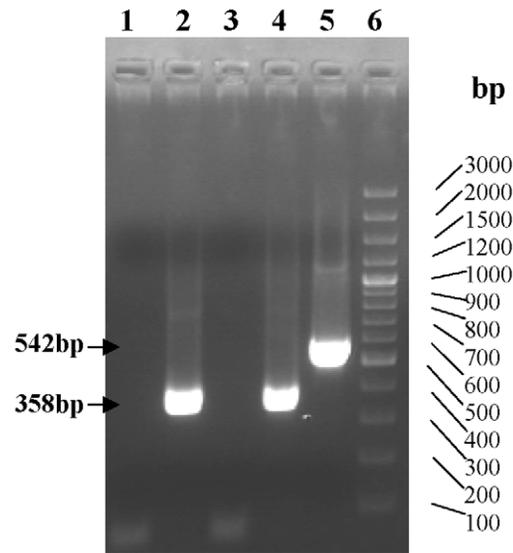
The tryptophan residues excite exclusively above 295 nm and emit in the range of 300 to 400 nm [39]. Figure 4b shows the shift in wavelength of maximum intensity ( $\lambda_{max}$ ) as a probe for the tertiary structure of the protein against pH. It shows an increase in  $\lambda_{max}$  as pH decreases, indicating an increase in the polarity of the tryptophan microenvironment, and thus increases in solvent accessibility to the core of the protein.

The hydrophobic dye, 1-anilino-8-naphthalene sulphonate (ANS), is used as a probe to determine the concentration of solvent accessible hydrophobic patches [40]. The fluorescence of protein-ANS complex showed no significant change as pH decreased from 7.3 to 3.6, but it was followed by an increase in the ANS binding with decrease in  $\lambda_{max}$  (Fig. 4c; inset) as pH decreases from 3.6 to 1.6 (Fig. 4c). The results show that the PvTRAg33.5 has lost its secondary and tertiary contacts to a significant level below pH 3.6.

**Effect of temperature.** To understand the stability of the tertiary structure of PvTRAg33.5, we have observed change in tryptophanyl fluorescence maxima (at 340 nm) as a probe for tertiary structure, against increase in temperature (Fig. 4d). The fluorescence intensity has decreased with increase in temperature from 20–45°C, followed by a sudden increase up to 55°C, and then a decrease till 80°C. But on cooling, the increase in fluorescence intensity did not crawl back in the same track below 50°C, indicating the protein might have been trapped in some non-native conformation. The change in  $\lambda_{max}$  seems to support the fact that on heating the  $\lambda_{max}$  increased from 339 to 344 nm as temperature increased from 20 to 80°C, showing increased accessibility of the hydrophobic tryptophan residues in solvent due to unfolding. But on reversing the temperature below 80°C, no significant change was observed, indicating entrapment of the protein molecules in local minima of folding.

**Effect of ammonium sulfate.** Effect of ammonium sulfate (0–90 mM) on PvTRAg33.5 has been observed by the far UV CD and the tryptophanyl fluorescence at different concentration of ammonium sulfate at pH 7.0 (Fig. 4e). The far UV CD spectra of the PvTRAg33.5 have shown minima at 208 and 222 nm, a characteristic feature of the  $\alpha$ -helical structure. Increase in the  $\alpha$ -helical structure was observed as salt concentration increased from

a)



b)

gDNA : TCTGCGGCTTACCTTTTAAGCAACATTTCTCCA**GT**AAGTCAGATTTAATCATTCAAAAAAAAAATTACCAT : 70  
 cDNA : TCTGCGGCTTACCTTTTAAGCAACATTTCTCCA----- : 33

gDNA : TTCATTTTTATCTACTTGATTGAATTTGCATTCATAGTAGAGAGATATGAATAAACATGTGTTGTTATTA : 140  
 cDNA : ----- : -

gDNA : CATATATACTAACTTGATAGATGCGGTTTAACTGACGCAAATTCATCCCCCAATGTGTACCTCATTTTT : 210  
 cDNA : ----- : -

gDNA : TGTGT**AG**TCCTCTCAAAGCGCAGTAGATTACATTGAACAGGAACCACTAGATATACTAAACCTAGAAGAA : 280  
 cDNA : -----TCCTCTCAAAGCGCAGTAGATTACATTGAACAGGAACCACTAGATATACTAAACCTAGAAGAA : 96

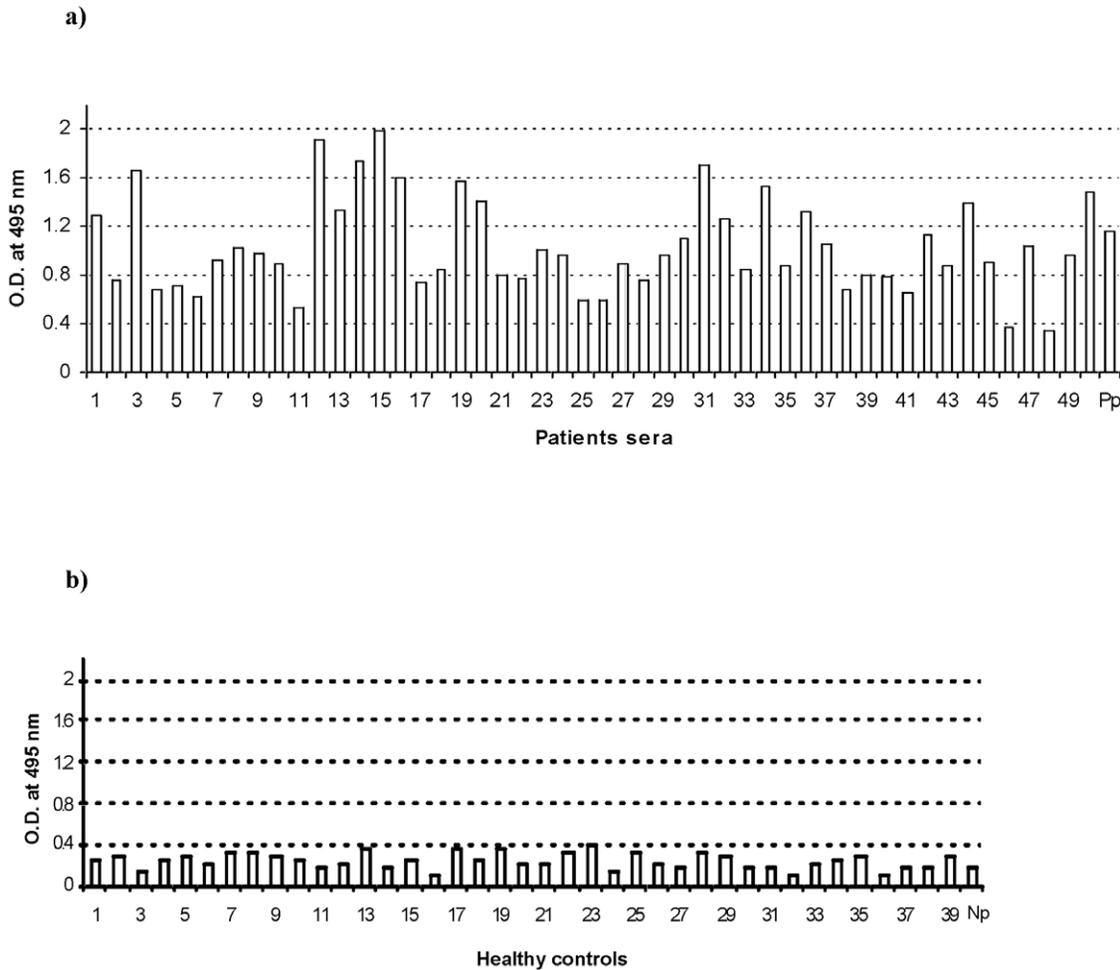
gDNA : GGTGATTTAGAAGTAACCGAACAAATGGAAAGATAATGAATGGCACAACCTGGAAGTTGAAATGGAGGAAG : 350  
 cDNA : GGTGATTTAGAAGTAACCGAACAAATGGAAAGATAATGAATGGCACAACCTGGAAGTTGAAATGGAGGAAG : 166

gDNA : ATTGGGATTCCTTTTCTACATCTTTGATTAGGGATAAGAAAGATTTTCATGAAAATAAAAAACGGATGAATT : 420  
 cDNA : ATTGGGATTCCTTTTCTACATCTTTGATTAGGGATAAGAAAGATTTTCATGAAAATAAAAAACGGATGAATT : 236

gDNA : AAATGGATGGCTTAATCTCGAGGAAAATAAATGGAATAATTTTTCTGGATATTTGAGTGTGGATACAAA : 490  
 cDNA : AAATGGATGGCTTAATCTCGAGGAAAATAAATGGAATAATTTTTCTGGATATTTGAGTGTGGATACAAA : 306

gDNA : AATTATCTATTAAAGAAATCCGAAAAATGGAATGATGCTGACTGGGAAAAC : 542  
 cDNA : AATTATCTATTAAAGAAATCCGAAAAATGGAATGATGCTGACTGGGAAAAC : 358

**Figure 2. In vivo expression of the *pvtrag33.5* gene.** a) Detection of the *pvtrag33.5* gene transcript by Reverse Transcription PCR. PCR amplification of the *P.vivax* cDNA (Lanes 1–4) and genomic DNA (Lane 5) using *pvtrag33.5* gene specific primers. RT- PCR product of *pvtrag33.5* specific cDNA synthesized using oligodT (Lane 2) or random decamers (Lane 4) with gene specific primers. Reverse transcription negative control using oligodT (Lane 1) or random decamers (Lane 3) as primers for cDNA synthesis. Lane 5: PCR product of genomic DNA amplification, and Lane 6: 100 bp ladder. PCR product sizes are indicated by arrows. b) Sequence alignment of PvTRAg33.5 genomic DNA and cDNA. Splice site sequence is shown in bold letters. Dashes indicate the absence of nucleotides. Numbers on the right-hand side indicate the number of nucleotides.  
 doi:10.1371/journal.pone.0016294.g002



**Figure 3. Seroreactivity of recombinant PvTRAg33.5 among *P. vivax* exposed individuals.** ELISA was performed with 50 sera from *P. vivax* patients (a) and 39 from unexposed healthy controls (b) by coating 96-well microtitre ELISA plates with the purified recombinant PvTRAg33.5. Each bar represents the OD value for the individual serum. Pp: pooled patient sera; Np: pooled sera from healthy controls. doi:10.1371/journal.pone.0016294.g003

0 to 90 mM. Figure 4e (inset) shows the emission spectra of the tryptophanyl residues of the PvTRAg33.5 at different salt concentrations. The increase in fluorescence on addition of the salt indicates opening of the hydrophobic tryptophanyl residues towards the polar solvent. From the far UV CD and fluorescence experiments, we may conclude that increase in salt concentration has disrupted the tertiary structure of the protein without destabilizing its secondary structure.

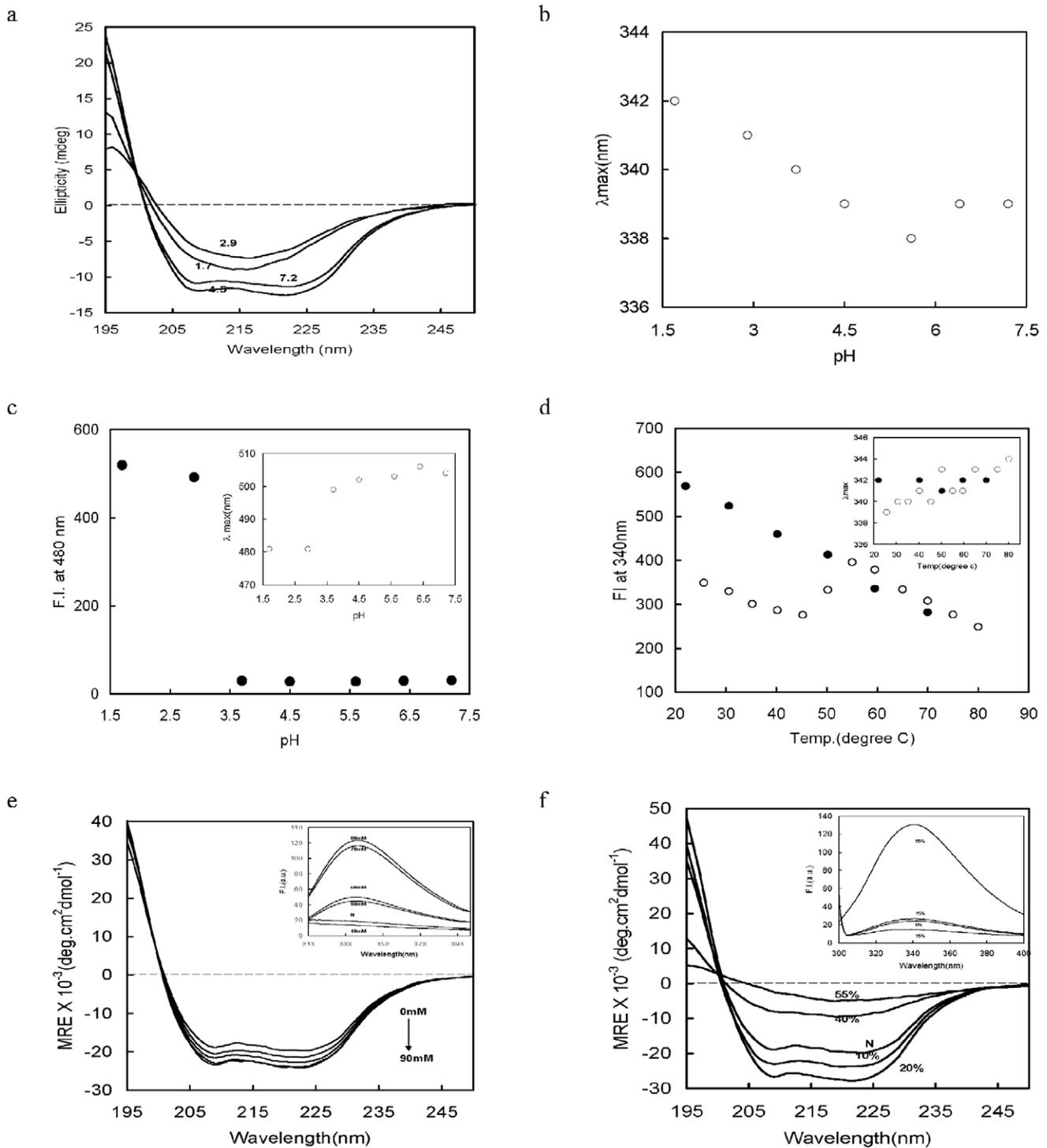
**Effect of ethanol.** Alcohols can be used as a model for plasma membrane [41]. Some of the tryptophan-rich proteins are found to get anchored to the RBC membrane [4,5]. We have checked the effect of ethanol on the PvTRAg33.5 by the far UV CD and tryptophanyl fluorescence spectroscopy to study the effect of hydrophobic environment on the protein conformation. Fig. 4f shows stabilization of the secondary and tertiary structure (inset) in the presence of 0 to 40% and 0 to 35% ethanol respectively. But further increase in the alcohol concentration disrupts the tertiary as well as secondary structure.

### Secondary structure and molecular model of the PvTRAg33.5

The secondary structure prediction studies show that the exon 2 encoded polypeptide predominantly comprises of helix and coil

regions. A consensus secondary structure prediction sequence was drawn from the results of the different secondary structure prediction servers (Fig. S2). In this, the sequences of each and every amino acid position was assigned either helix or coil or extended sheets only when at least 50% of the above web servers predicted the similar secondary structure. The consensus secondary structure prediction showed that the PvTRAg33.5 has 76.2% helix and 23.8% coil with no extended sheets. This helical content derived from the analysis was more than that of the above mentioned experimental CD data.

In the absence of a clear structural homologue from PDB or pattern prediction, knowledge-based approach to structure prediction employing the fold recognition and protein threading methods was undertaken to determine computationally a suitable template model for the exon 2 encoded polypeptide. The results from the diverse threading servers were compiled and analyzed to derive possible templates based on a three-fold criterion. Models predicted with less than 60% accuracy, comprising short segments of fewer than 90 amino acid residues and structures consisting mainly of  $\beta$ -sheets were not included as these contradicted both experimental and secondary structure consensus data. Several models were built using the templates so obtained (PDB Ids: 2c5k, 1is1, 2a2f, 2lis, 1zro, 2bs5, 2hsb, 2pe4, 2c6j, 1u4q, 2j68 and



**Figure 4. Physico-chemical characterization of PvTRAg33.5.** Far UV CD spectra at different pH (a), emission wavelength at maximum intensity after exciting it at 295 nm as a function of pH (b), fluorescence intensity (filled circle) and emission wavelength at maximum intensity (empty circle) (inset) of ANS-protein complex as a function of pH, after exciting it at 380 nm (c), fluorescence intensities and emission wavelength maxima (inset) after exciting at 295 nm as a function of increase in temperature (empty circle) and consecutive decrease in temperature (filled circle) at pH 7.0 (d), far UV CD spectra and fluorescence intensities (inset) in the absence and presence of different ammonium sulfate concentrations at pH 7.0 after excited at 295 nm (e), and far UV CD spectra and fluorescence intensities (inset) in the absence (N) and presence of different ethanol concentrations after exciting at 295 nm (f) of PvTRAg33.5.  
doi:10.1371/journal.pone.0016294.g004

models from SAM-T08 and iTASSER servers) in the Discovery Studio v1.7. The loops were built using the LOOPREFINE-MENT module. A high quality model was selected based on the

probability density function (PDF) calculated by the program. The PDF is derived from spatial restraints while building the model and identifies regions of high energy. The final model showed

similarity to template models predicted via SAM-T08 and PHYRE. The stereochemistry of the final model was checked with the program PROCHECK [35].

The refinement of the model with special emphasis on the loops was carried out using molecular dynamics simulation. The total energy and simulation temperature were found to remain steady with little fluctuation (Table 2). In addition to this the maximum r.m.s. deviation was observed to be 0.6 Å. Hence, the results from the molecular dynamics simulation indicated that the modeled structure is stable.

## Discussion

For the first time, we describe here the structural characteristic features of a tryptophan-rich antigen (PvTRAg33.5). This protein is not a membrane integrated protein, it lacks cysteine residues, and it is soluble in nature. The tryptophan residues are distributed over the entire length of the mature protein (encoded by exon 2) which may influence the folding of the protein in the absence of the cysteine residues. Thus, the three dimensional conformation of the protein is most probably constrained by the number and hydrophobic nature of these tryptophan residues. The tryptophan residue is conserved in the motifs, however, a great deal of variability is observed for adjacent polar and hydrophobic residues. Despite the absence of orthologues, the profile HMM analysis clearly indicates the presence of similar motifs in the *Plasmodium* species. In non malarial proteins, the tryptophan rich domains have been reported in a variety of the transmembrane surface proteins which play an important role in the folding and assembly [42].

The modeled tertiary structure of the protein consists of three domains (Fig. 5). The subdomain 1 comprising residues 1–28 adopts a random coil structure except for a two turn helix (H1), the subdomain 2 spanning residues 30–113 has 2 long (H2 and H3) and 2 short helices and the subdomain 3 stretching from residues 126–252 consists of three long helices (H6, H7: and H8). A long loop of 16 residues connects helices 6 and 7 whereas helices 7 and 8 are anti-parallel and are joined by a short loop of 5 residues. The subdomains 2 and 3 are associated by a 12 amino acid long flexible linker region. The linker region adopts a random coil structure and permits different orientations of the subdomains 1 and 2 as regards to the third subdomain. The motifs, M2-M7, identified from MEME have helical structures. The subdomain 3 participates in numerous hydrogen bonds and van der Waals interactions with subdomain 2 while these are limited to the tail residues to subdomain 1.

Thus the PvTRAg33.5 protein belongs to the proteins from the all-alpha class with no identical fold in Structural Classification of Proteins (SCOP) database [43]. The closest structural similar fold in SCOP was observed to be the Duffy binding like domain (DBL)

fold which has earlier been observed in *P. knowlesi* (PDB Id: 2C6J) [44] and *P. falciparum* (PDB Ids: 3CML [44] 1ZRO and 1ZRL [45]). The DBL domain possesses two subdomains, each with a three-helical bundle. The subdomain 3 of PvTRAg33.5 has a 3 helix bundle comparable to the DBL domain but a structural superimposition of the long helices of this subdomain revealed that the position of the helices with respect to each other are dissimilar. The main difference is observed in the subdomain 2. A structural superimposition along the long helix H6 of the subdomain 3 indicated that the orientation and positioning of the subdomains 2 and 3 was different with respect to each other in these proteins. These differences indicate that the linker region connecting the subdomain 2 and 3 is flexible and a conformational change can alter their relative orientations.

The CD data indicates that the PvTRAg33.5 has nearly 50%  $\alpha$ -helix structure which was also predicted by the secondary structure prediction analysis. However, the protein has shown a total loss of the alpha helical structure if pH was decreased below 4.5. Such structural change from the alpha helical to the beta sheet and vice-versa has been reported in various systems, for example, the beta sheet structure of the HIV's gp120 gets converted into the alpha helix when virus binds to the host membrane protein CD4 [45–49]. The PvTRAg33.5 protein, however, remains stable upto pH 4.5 or till 50°C at pH 7.0. Even 40% ethanol could not disrupt its secondary structure. These evidences suggest that the protein is stable enough to withstand some environmental stresses up to certain extent.

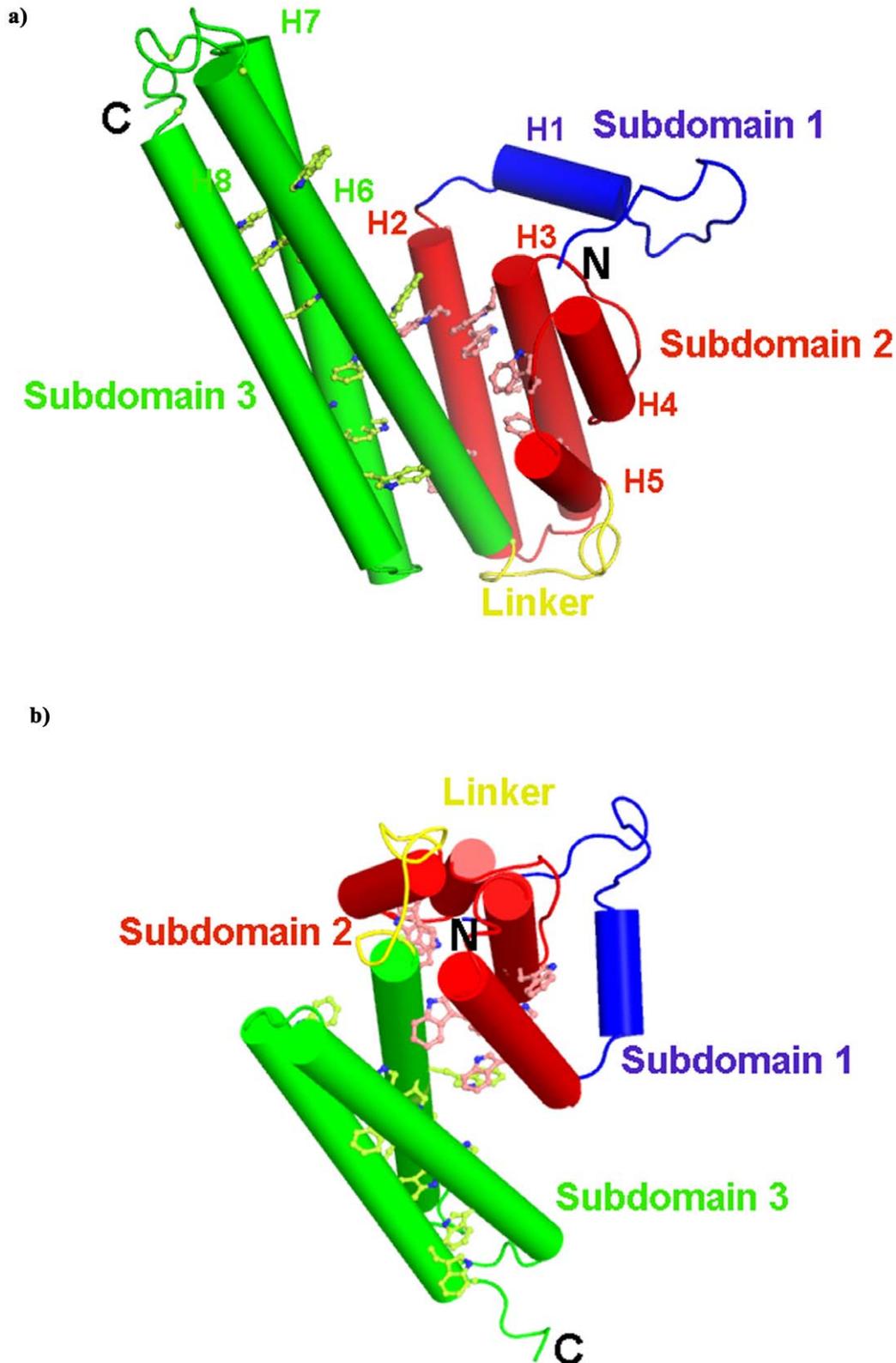
The PvTRAg33.5, probably being a small polypeptide, has two state transitions in its secondary structure under the denaturing conditions. The PvTRAg55, probably being a longer polypeptide [12], seems to be more resistant to the effect of low pH than the PvTRAg33.5 (Data not shown). In short, the PvTRAg33.5 is a predominantly  $\alpha$ -helical but a loss of the structure occurs under severe environmental stressed conditions. Under stress, either more number of the tryptophan residues gets exposed or the similar number of tryptophan residues exposed with larger extents, as indicated by the tryptophan fluorescence under low pH, high temperature and high concentration of the salt or the non polar solvent. The tryptophan rich regions appear to be critical for the cell-cell fusion activity of the trans-membrane glycoproteins [46]. In nature, certain external stimuli may also cause the structural changes in the proteins leading to the diseased conditions [47]. For example, binding of the ligands could trigger the conformational changes in some of the Trypanosomal enzymes to modulate their activity leading to the disease pathogenesis [48,49]. The structural perturbation of the PvTRAg33.5 may, therefore, connect with its biological function such as its probable role in the erythrocyte invasion. The Tryptophan residues have also been shown to play an important role in protein-protein interactions and signal transduction [50,51,52,53,54,55]. There are rapidly mounting evidence that the proteins must at least partly unfold in order to be translocated across a biological membrane. The precursor proteins are found to unfold before or at the time of their transport across the membranes [56].

In conclusion, for the first time the structural characteristic features of a *Plasmodium* tryptophan-rich protein are described here. The PvTRAg33.5, a novel tryptophan-rich antigen from *P. vivax*, is expressed during the blood stage of the parasite, has a predominantly  $\alpha$ -helical structure, and its secondary structure is lost upto 50% under the stressed conditions which may also be faced by the parasite in its host. The modeled structure has three domain architecture and this conformation is stabilized by the presence of a large number of tryptophan residues which in the absence of disulphide bonds influence the folding due

**Table 2.** Molecular Dynamics simulation.

Time (ps)	rmsd w.r.t to first conformer (Å)	Potential Energy (kcal/mol)	Temp (K)
0	0	–570.36	303
25	0.56	–566.30	306
50	0.59	–568.31	302
75	0.59	–569.15	301
100	0.60	–570.04	302

doi:10.1371/journal.pone.0016294.t002



**Figure 5. Model of PvTRAg33.5 structure.** a) Model of PvTRAg33.5 structure which has three subdomains: subdomain 1 (blue), subdomain 2 (red), subdomain 3 (green) and a linker (yellow) region. The helices are represented as cylinders and the coiled structure as loops. Subdomain 1 adopts a random coil structure except for a two turn helix from residues 16–25 (H1). The second subdomain has 2 long (H2: 29–54 and H3: 60–80) and 2 short helices (H4: 89–98 and H5: 104–113) connected by three loops. The subdomain 3 consists of three long helices (H6: 126–162, H7: 177–213 and H8: 217–244). The tryptophan (Trp) residues in the helices are drawn in ball-and-stick and assigned the helix color. b) The three dimensional model of

PvTRAg33.5 indicating the orientation of the subdomains relative to each other at a rotation of 90° to Panel a. The linker region connecting subdomains 2 and 3 is flexible. The tryptophan (Trp) residues in the helices are drawn in ball-and-stick and assigned the helix color. doi:10.1371/journal.pone.0016294.g005

to their predominantly hydrophobic nature. Furthermore, the PvTRAg33.5 contains B-cell epitopes which are highly conserved in the parasite population and high antibody response is being generated in humans during *P. vivax* infection. Further studies are required to define the biological function of this protein in the parasite.

## Supporting Information

**Figure S1 Multiple sequence alignment of PvTRAg33.5 (Plasmodb ID: PVX\_121897) with close tryptophan rich homologs present in *Plasmodium* species.** The sequences of these proteins were retrieved from the *Plasmodium* database (www.plasmodb.org) and aligned using ClustalW software at <http://www.ebi.ac.uk/clustalW>. The plasmodb ID of *P. vivax* homologs are PVX\_090250 for PvTRAg39.8, PVX\_109280 for PvTRAg35.2, PVX\_090265 for PvTRAg, PVX\_101515 for PvTRAg40, PVX\_096995 for PvTRAg55, PVX\_112655 for PvTRAg80.6 and PVX\_101510 for PvATRAg74. The plasmodb ID of *P. falciparum* homologs are PFA0135w for PMATrA, PF08\_0003 for PfTryThrA, PF10\_0026 for TrpA-3, and MAL13P1.269 for LysTrpA. The plasmodb ID of *P. yoelii* homologs are PY06023 for PypAg1, and PY03625 for PypAg3. Stars (“\*”) indicate identical amino acids while double (“:”) and single (“.”) dots indicate the conserved and semi-conserved substitutions, respectively. All tryptophan residues are in boldface. The positionally conserved tryptophan residues are shaded grey. Dashes indicate the absence of amino acids. Numbers on the right-

hand side indicate the number of amino acid residue. The secondary structure elements for PvTRAg33.5 are indicated on the top of the sequences. The coils are indicated as bold lines and helices as tubes. The helices are numbered H1 to H8. Alignment is according to the PvTRAg33.5 amino acids sequence (complete sequences of all proteins are not shown). (TIF)

**Figure S2 Secondary structure prediction of exon 2 encoded PvTRAg33.5 by webserver and the consensus secondary structure prediction.**

(TIF)

**Supplementary Material S1** Secondary Structure Prediction Servers.

(DOC)

## Acknowledgments

We acknowledge the facility of Biotechnology Information System (BTIS) and Bio-Medical Informatics Centre of the Indian council of Medical Research to analyze the data.

## Author Contributions

Conceived and designed the experiments: YDS RHK PK. Performed the experiments: HB SG PS DK. Wrote the paper: YDS RHK PK.

## References

- Reed ZH, Friede M, Kieny MP (2006) Malaria vaccine development: progress and challenges. *Curr Mol Med* 6: 231–245.
- Malkin EM, Durbin AP, Diemert DJ, Sattabongkot J, Wu Y, et al. (2005) Phase 1 vaccine trial of Pvs25H: a transmission blocking vaccine for *Plasmodium vivax* malaria. *Vaccine* 23: 3131–3138.
- King CL, Michon P, Shakri AR, Marcotty A, Stanisic D, et al. (2008) Naturally acquired Duffy-binding protein-specific binding inhibitory antibodies confer protection from blood-stage *Plasmodium vivax* infection. *Proc Natl Acad Sci USA* 105: 8363–8368.
- Burns JM, Jr., Adeeku EK, Dunn PD (1999) Protective immunization with a novel membrane protein of *Plasmodium yoelii*-infected erythrocytes. *Infect Immun* 67: 675–680.
- Burns JM, Adeeku EK, Belk CC, Dunn PD (2000) An unusual tryptophan-rich domain characterizes two secreted antigens of *Plasmodium yoelii*-infected erythrocytes. *Mol Biochem Parasitol* 110: 11–21.
- Ntumngia FB, Bahamontes-Rosa N, Kun JF (2005) Genes coding for tryptophan-rich proteins are transcribed throughout the asexual cycle of *Plasmodium falciparum*. *Parasitol Res* 96: 347–353.
- Ntumngia FB, Bouyou-Akotet MK, Uhlemann AC, Mordmuller B, Kremsner PG, et al. (2004) Characterisation of a tryptophan-rich *Plasmodium falciparum* antigen associated with merozoites. *Mol Biochem Parasitol* 137: 349–353.
- Uhlemann AC, Oguariri RM, McColl DJ, Coppel RL, Kremsner PG, et al. (2001) Properties of the *Plasmodium falciparum* homologue of a protective vaccine candidate of *Plasmodium yoelii*. *Mol Biochem Parasitol* 118: 41–48.
- Curtidor H, Ocampo M, Rodriguez LE, Lopez R, Garcia JE, et al. (2006) *Plasmodium falciparum* TryThrA antigen synthetic peptides block in vitro merozoite invasion to erythrocytes. *Biochem Biophys Res Commun* 339: 888–896.
- Alam MT, Bora H, Mitra P, Singh N, Sharma YD (2008) Cellular immune responses to recombinant *Plasmodium vivax* tryptophan-rich antigen (PvTRAg) among individuals exposed to vivax malaria. *Parasite Immunol* 30: 379–383.
- Alam MT, Bora H, Singh N, Sharma YD (2008) High immunogenicity and erythrocyte-binding activity in the tryptophan-rich domain (TRD) of the 74-kDa *Plasmodium vivax* alanine-tryptophan-rich antigen (PvATRAg74). *Vaccine* 26: 3787–3794.
- Siddiqui AA, Singh N, Sharma YD (2007) Expression and purification of a *Plasmodium vivax* antigen - PvTRAg55 tryptophan- and alanine-rich antigen and its immunological responses in human subjects. *Vaccine* 26: 96–107.
- Siddiqui AA, Bora H, Singh N, Dash AP, Sharma YD (2008) Expression, purification, and characterization of the immunological response to a 40-kilodalton *Plasmodium vivax* tryptophan-rich antigen. *Infect Immun* 76: 2576–2586.
- Garg S, Chauhan SS, Singh N, Sharma YD (2008) Immunological responses to a 39.8 kDa *Plasmodium vivax* tryptophan-rich antigen (PvTRAg39.8) among humans. *Microbes Infect* 10: 1097–1105.
- Jalah R, Sarin R, Sud N, Alam MT, Parikh N, et al. (2005) Identification, expression, localization and serological characterization of a tryptophan-rich antigen from the human malaria parasite *Plasmodium vivax*. *Mol Biochem Parasitol* 142: 158–169.
- Ray P, Ansari MA, Sharma YD (1994) *Plasmodium vivax*: immune responses in a cross-section of the population in the Delhi area of India. *Am J Trop Med Hyg* 51: 436–443.
- Chen YH, Yang JT, Martinez HM (1972) Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion. *Biochemistry* 11: 4120–4131.
- Andrade MA, Chacon P, Merelo JJ, Moran F (1993) Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng* 6: 383–390.
- Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S (2002) Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* 18: 298–305.
- Gomi M, Sonoyama M, Mitaku S (2004) High performance system for signal peptide prediction: SOSUsignal. *Chem-bio informatics journal* 4: 142–147.
- Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305: 567–580.
- Hofmann K, Stoffel W (1993) TMbase - A database of membrane spanning proteins segments. *biol chem Hoppe-Seyler* 374: 166.
- Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, et al. (2005) InterProScan: protein domains identifier. *Nucleic Acids Res* 33: W116–120.
- Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Menlo Park, California: Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology AAAI Press. pp 28–36.
- Bailey TL, Gribskov M (1998) Combining evidence using p-values: application to sequence homology searches. *Bioinformatics* 14: 48–54.

26. Grundy WN, Bailey TL (1999) Family pairwise search with embedded motif models. *Bioinformatics* 15: 463–470.
27. Bennett-Lovsey RM, Herbert AD, Sternberg MJ, Kelley LA (2008) Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins* 70: 611–625.
28. Shi J, Blundell TL, Mizuguchi K (2001) FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J Mol Biol* 310: 243–257.
29. Rost B, Liu J (2003) The PredictProtein server. *Nucleic Acids Res* 31: 3300–3304.
30. Karplus K, Karchin R, Barrett C, Tu S, Cline M, et al. (2001) What is the value added by human intervention in protein structure prediction? *Proteins: Structure Function and Genetics* 45: 86–91.
31. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9: 40.
32. Zhang Y (2007) Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins* 69(Suppl 8): 108–117.
33. Wu S, Skolnick J, Zhang Y (2007) Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biology* 5: 17.
34. Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, et al. (1983) CHARMM: A Program for macromolecular Energy, Minimization, and Dynamics calculation. *Journal Comput Chem* 4: 187–217.
35. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK - A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26: 283–291.
36. Ahmad B, Kamal MZ, Khan RH (2004) Alkali-induced conformational transition in different domains of bovine serum albumin. *Protein Pept Lett* 11: 307–315.
37. Khan F, Khan RH, Muzammil S (2000) Alcohol-induced versus anion-induced states of alpha-chymotrypsinogen A at low pH. *Biochim Biophys Acta* 1481: 229–236.
38. Zsila F, Bikadi Z, Simonyi M (2004) Circular dichroism spectroscopic studies reveal pH dependent binding of curcumin in the minor groove of natural and synthetic nucleic acids. *Org Biomol Chem* 2: 2902–2910.
39. Bhakuni V (1998) Alcohol-induced molten globule intermediates of proteins: are they real folding intermediates or off pathway products? *Arch Biochem Biophys* 357: 274–284.
40. Matulis D, Lovrien R (1998) 1-Anilino-8-naphthalene sulfonate anion-protein binding depends primarily on ion pair formation. *Biophys J* 74: 422–429.
41. Konno T, Iwashita J, Nagayama K (2000) Fluorinated alcohol, the third group of cosolvents that stabilize the molten-globule state relative to a highly denatured state of cytochrome c. *Protein Sci* 9: 564–569.
42. Schiffer M, Chang C, Stevens FJ (1992) The functions of tryptophan residues in membrane proteins. *Protein Eng* 5: 213–214.
43. Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* 247: 536–540.
44. Singh K, Gitis AG, Nguyen P, Gowda DC, Miller LH, et al. (2008) Structure of the DBL3x domain of pregnancy-associated malaria protein VAR2CSA complexed with chondroitin sulfate A. *Nat Struct Mol Biol* 15: 932–938.
45. Tolia NH, Enemark EJ, Sim BK, Joshua-Tor L (2005) Structural basis for the EBA-175 erythrocyte invasion pathway of the malaria parasite *Plasmodium falciparum*. *Cell* 122: 183–193.
46. Salzwedel K, West JT, Hunter E (1999) A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity. *J Virol* 73: 2469–2480.
47. Pan KM, Baldwin M, Nguyen J, Gasset M, Serban A, et al. (1993) Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. *Proc Natl Acad Sci U S A* 90: 10962–10966.
48. Buschiazzo A, Amaya MF, Cremona ML, Frasch AC, Alzari PM (2002) The crystal structure and mode of action of trans-sialidase, a key enzyme in *Trypanosoma cruzi* pathogenesis. *Mol Cell* 10: 757–768.
49. McNae IW, Martinez-Oyanedel J, Keillor JW, Michels PA, Fothergill-Gilmore LA, et al. (2009) The crystal structure of ATP-bound phosphofructokinase from *Trypanosoma brucei* reveals conformational transitions different from those of other phosphofructokinases. *J Mol Biol* 385: 1519–1533.
50. Li B, Wu X, Zhou H, Chen Q, Luo Y (2004) Acid-induced unfolding mechanism of recombinant human endostatin. *Biochemistry* 43: 2550–2557.
51. Yeliseev AA, Kaplan S (2000) TspO of *rhodobacter sphaeroides*. A structural and functional model for the mammalian peripheral benzodiazepine receptor. *J Biol Chem* 275: 5657–5667.
52. Divita G, Baillon JG, Rittinger K, Chermann JC, Goody RS (1995) Interface peptides as structure-based human immunodeficiency virus reverse transcriptase inhibitors. *J Biol Chem* 270: 28642–28646.
53. Nishio M, Garcin D, Simonet V, Kolakofsky D (2002) The carboxyl segment of the mumps virus V protein associates with Stat proteins in vitro via a tryptophan-rich motif. *Virology* 300: 92–99.
54. Schulz H, Pelliccioli EC, Thony-Meyer L (2000) New insights into the role of CcmC, CcmD and CcmE in the haem delivery pathway during cytochrome c maturation by a complete mutational analysis of the conserved tryptophan-rich motif of CcmC. *Mol Microbiol* 37: 1379–1388.
55. Xue H, Wong JT (1995) Interferon induction of human tryptophanyl-tRNA synthetase safeguards the synthesis of tryptophan-rich immune-system proteins: a hypothesis. *Gene* 165: 335–339.
56. Matouschek A, Azem A, Ratliff K, Glick BS, Schmid K, et al. (1997) Active unfolding of precursor proteins during mitochondrial protein import. *Embo J* 16: 6727–6736.