



Computational Design of Protein-Based Inhibitors of Plasmodium vivax Subtilisin-Like 1 Protease

Giacomo Bastianelli, Anthony Bouillon, Christophe Nguyen, Dung Le-Nguyen, Michael Nilges, Jean-Christophe Barale

► To cite this version:

Giacomo Bastianelli, Anthony Bouillon, Christophe Nguyen, Dung Le-Nguyen, Michael Nilges, et al.. Computational Design of Protein-Based Inhibitors of Plasmodium vivax Subtilisin-Like 1 Protease. PLoS ONE, 2014, 9 (10), 10.1371/journal.pone.0109269 . pasteur-01414431

HAL Id: pasteur-01414431

<https://pasteur.hal.science/pasteur-01414431>

Submitted on 12 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Computational Design of Protein-Based Inhibitors of *Plasmodium vivax* Subtilisin-Like 1 Protease

Giacomo Bastianelli^{1,2}, Anthony Bouillon^{3,4*}, Christophe Nguyen⁵, Dung Le-Nguyen⁵, Michael Nilges^{1,2*}, Jean-Christophe Barale^{3,4*}

1 Institut Pasteur, Unité de Bioinformatique Structurale, Département de Biologie Structurale et Chimie, Paris, France, **2** CNRS UMR 3528, Paris, France, **3** Institut Pasteur, Unité d'Immunologie Moléculaires des Parasites, Département de Parasitologie et de Mycologie & CNRS URA 2581, Paris, France, **4** CNRS, URA2581, Paris, France, **5** SYSDIAG, CNRS UMR3145 CNRS-BioRad, Montpellier, France

Abstract

Background: Malaria remains a major global health concern. The development of novel therapeutic strategies is critical to overcome the selection of multiresistant parasites. The subtilisin-like protease (SUB1) involved in the egress of daughter *Plasmodium* parasites from infected erythrocytes and in their subsequent invasion into fresh erythrocytes has emerged as an interesting new drug target.

Findings: Using a computational approach based on homology modeling, protein-protein docking and mutation scoring, we designed protein-based inhibitors of *Plasmodium vivax* SUB1 (PvSUB1) and experimentally evaluated their inhibitory activity. The small peptidic trypsin inhibitor EETI-II was used as scaffold. We mutated residues at specific positions (P4 and P1) and calculated the change in free-energy of binding with PvSUB1. In agreement with our predictions, we identified a mutant of EETI-II (EETI-II-P4LP1W) with a *K_i* in the medium micromolar range.

Conclusions: Despite the challenges related to the lack of an experimental structure of PvSUB1, the computational protocol we developed in this study led to the design of protein-based inhibitors of PvSUB1. The approach we describe in this paper, together with other examples, demonstrates the capabilities of computational procedures to accelerate and guide the design of novel proteins with interesting therapeutic applications.

Citation: Bastianelli G, Bouillon A, Nguyen C, Le-Nguyen D, Nilges M, et al. (2014) Computational Design of Protein-Based Inhibitors of *Plasmodium vivax* Subtilisin-Like 1 Protease. PLoS ONE 9(10): e109269. doi:10.1371/journal.pone.0109269

Editor: Laurent Rénia, Agency for Science, Technology and Research - Singapore Immunology Network, Singapore

Received: April 18, 2014; **Accepted:** August 16, 2014; **Published:** October 24, 2014

Copyright: © 2014 Bastianelli 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.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. The 3D-model of PvSUB1 is available from the Dryad Digital Repository at <http://doi.org/10.5061/dryad.h8c79>.

Funding: This work was partly supported by MEST-CT-05-020311, a Marie Curie Early Stage Research Training Fellowship (EIMID) of the Framework Program 6 by the European Commission. AB is a fellow of the "Direction Generale pour l'Armement" from the French Ministry of Defense. This work was partly supported by the "Fond dédié: combattre les maladies parasitaires" granted by Sanofi-Aventis and the French Ministry of Research and the Institut Carnot-Pasteur Maladies Infectieuses. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. SYSDIAG, CNRS UMR3145 CNRS-BioRad provided support in the form of salaries and operating costs for authors CN and DLN, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the "author contributions" section.

Competing Interests: The authors have the following interests. This work was partly supported by the "Fond dédié: combattre les maladies parasitaires" granted by the French Ministry and Sanofi-Aventis. Christophe Nguyen and Dung Le-Nguyen are employed by SYSDIAG, CNRS UMR3145 CNRS-BioRad. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* Email: nilges@pasteur.fr (MN); jcb@pasteur.fr (JCB)

† These authors participated equally in the supervision of this work.

* Current address: Institut Pasteur, Unité de Biologie et Génétique du Paludisme, Team Malaria Targets and Drug Development, Département de Parasitologie et de Mycologie, Paris, France

Introduction

With more than 400 millions infections worldwide, malaria remains a major public health issue, principally in sub-Saharan Africa. An effective vaccine would help reduce disease burden, but the best candidates are still in development or evaluation phase [1,2]. The rapid development of multidrug-resistant *Plasmodium* [3] parasites necessitates accelerating the discovery of novel anti-malarial compounds to meet the needs of the agenda for malaria control and eradication [4].

In humans, *Plasmodium sp.* development comprises different stages, with the asexual intra-erythrocytic forms being responsible for the symptoms of the disease, such as fever, anemia, and cerebral malaria that can lead to death [5]. The erythrocyte invasion by *Plasmodium* merozoites critically depends on protease activities involved in both the daughter parasites egress from erythrocytes, and invasion into another erythrocyte. The parasite subtilisin-like protein 1 (SUB1) plays a critical role during both the hepatic and erythrocytic phases of *Plasmodium* biological cycle and is hence considered an interesting multi-stage target for developing a new class of anti-malarials [6] [7].

Most of the ancient therapies against *Plasmodium* are based on small molecules such as chloroquine, quinolones, antifolate, artemisinin derivatives, or atovaquone. The development of new classes of active molecules such as protein-based drugs or peptidomimetics [8,9] is an active and promising field of research. Among protein-based drugs, dermaseptin S4 (DS4) was shown to irreversibly inhibit the *in vitro* parasite growth through a cytotoxic hemolytic activity. Dermaseptin S3 acts in a similar manner as DS4 but did not present hemolytic activity through a cytotoxic hemolytic activity [10].

STRUCTURES MODELING

SUB1 (Homology Modeling)

EETI-II (Substrate Mutant)

DOCKING

Ensemble Docking

Conformational Sampling (MDs)

Docking Refinement

SCORING

Mutate

Conformational Sampling (MDs)

Free Energy Calculations (GBSA)

EXPERIMENTAL TESTING

Protein Production

Enzymatic testing

In the design of protein-based drugs, most approaches use combinatorial libraries based on different screening methods such as phage [11], ribosome [12] or mRNA display [13]. Their use is wide-spread, in particular for selecting high-affinity protein binders, despite their limitations due to the library size and the large quantities of the target protein needed to perform screening. Moreover, when the selection is not based on binding but on inhibiting a crucial enzyme of the biological cycle, a rather complex selection system has to be employed. Computational protein design can be used to reduce the sequence/structure space that needs to be explored and thus accelerate the process of screening and selection of target inhibitors.

Here, we present a strategy for the computational design of protein-based inhibitors targeting the subtilisin-like 1 protease of the human parasite *Plasmodium vivax* (PvSUB1). PvSUB1 can be expressed as a recombinant active enzyme [14] [15], and a specific enzymatic assay allows one to evaluate specific inhibitors. To search for potential inhibitors of PvSUB1, we used a computational design strategy, employing as scaffold the small protein EETI-II (*Ecballium elaterium* trypsin inhibitor II) [16], a trypsin inhibitor extracted from *Ecballium elaterium*. The family of cysteine-knot proteins, to which EETI-II belongs, and in particular the cyclotides [17], possesses interesting biochemical properties [18]. EETI-II is composed of 28 amino-acids and its three-dimensional structure is tightly constrained by 3 disulphide bridges that contribute to its rigidity and biological stability [19]. We opted for this scaffold because several studies showed the possibility to engineer this protein to obtain specific mutants [20], *via* the extension of the EETI bioactive loop [21] or by changing its sequence to change its specificity towards the targeted enzyme [22] [23] [24] [25].

Compared to studies using an iterative computational design procedure focused on electrostatic binding contributions and single mutants [26], or on re-designing a scaffold protein to bind to a specified region on a target protein [27], we here faced the additional challenge that the 3D structure of the target itself or a close sequence homologue was not known. Nonetheless, the use of state-of-the-art structure prediction, docking and scoring methods allowed us to successfully identify mutants of the scaffold EETI-II that inhibited the target PvSUB1 enzyme.

Results and Discussion

The computational protein design approach involved four steps (see Figure 1). The first step was the modeling of the structure of the enzyme (PvSUB1) and the scaffold (EETI-II). Because of the lack of an experimental PvSUB1 structure, we built structures based on sequence homology. We also generated the model of a mutant of EETI-II containing the substrate sequence of PvSUB1, which we called EETI-II-sub. The second step was the docking of EETI-II-sub to the target protein. We employed an ensemble docking procedure with several conformations obtained from molecular dynamics (MD) simulations for each protein partner to implicitly include flexibility in the docking, and refined the best docking solutions by molecular dynamics to obtain high-quality structures of the complex. The third step aimed at identifying mutants of EETI-II-sub that had higher binding affinity towards PvSUB1. In this step, we mutated residues in EETI-II-sub at the protein-protein interface of the complex, ran conformational sampling of the mutant with molecular dynamics, and calculated the free energy of binding via implicit solvent models based on the Generalized Born approximation (GBSA). The last step consisted in the experimental testing of the inhibitor by an enzymatic inhibitory assay specific for the PvSUB1 recombinant enzyme.

Figure 1. Computational Protein Design Strategy. Step 1: Prediction of the structure of the enzyme (PvSUB1) by comparative modelling and of the scaffold for mutational analysis (EETI-II-sub) by replacing one of the loops with a substrate sequence. Step 2: docking of EETI-II-sub to the target protein by ensemble docking procedure with several conformations from molecular dynamics simulations for each protein partner, and refinement of the best solutions. Step 3: mutation of the scaffold, conformational sampling and scoring of the mutants. Step 4: experimental testing by an enzymatic inhibitory assay on the recombinant enzyme of PvSUB1.

doi:10.1371/journal.pone.0109269.g001

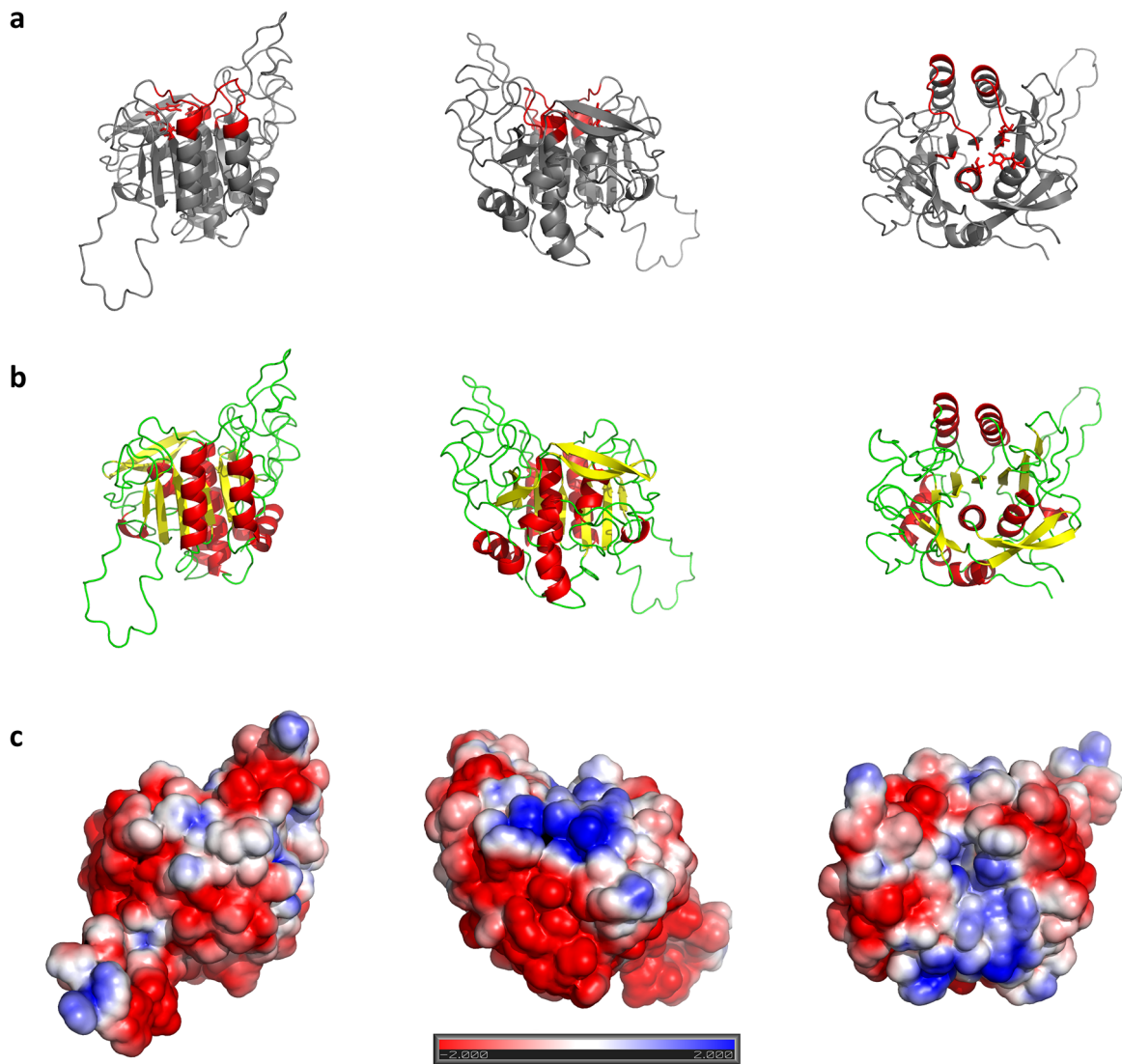


Figure 2. 3D model of PvSUB1 catalytic domain. **A:** Highlighted in red is the region forming the substrate binding pocket and red sticks correspond to the residues that form the catalytic triad; **B:** Cartoon representation of secondary structures; **C:** APBS surface electrostatic representation.

doi:10.1371/journal.pone.0109269.g002

Table 1. Catalytic site distances along MD simulations.

Distance	1R0R	1T02	PvSUB1
HIS@CA-SER@CA	8.52±0.34	9.38±0.36	9.89±0.24
HIS@CB-SER@CB	6.56±0.43	7.55±0.41	7.91±0.31
HIS@CA-ASP@CA	7.78±0.52	7.91±0.29	8.33±0.18
HIS@CB-ASP@CB	6.59±0.66	6.73±0.41	7.46±0.24
SER@CA-ASP@CA	10.24±0.35	9.18±0.32	10.35±0.25
SER@CB-ASP@CB	7.42±0.42	8.83±0.49	8.72±0.34
SER@OG-HIS@NE2*	7.54±1.11	4.69±1.19	5.11±0.71
ASP@CG-HIS@ND1	8.52±0.34	9.38±0.36	6.74±0.54
ASN@CG-SER@CB	6.70±0.71	7.5±0.91	6.52±0.41

Values are expressed in Å. The distance SER@OG-HIS@NE2 shows the largest fluctuation for both subtilisins with known structure and for our models, consistent with variations of this distance in subtilisin crystal structures (Table 2). Catalytic triad: Asp316, His372, Ser549; Asp137, His168, Ser325 and Asp139, His171, Ser328 for PvSUB1, 1R0R and 1T02 respectively.

doi:10.1371/journal.pone.0109269.t001

Table 2. Subtilisin catalytic site geometries.

Distance	lower-range	upper-range
HIS@CA-SER@CA	8.3	8.72
HIS@CB-SER@CB	6.44	6.89
HIS@CA-ASP@CA	7.22	7.46
HIS@CB-ASP@CB	5.83	6.61
SER@CA-ASP@CA	9.87	10.11
SER@CB-ASP@CB	8.15	8.53
SER@OG-HIS@NE2	2.57	3.36
ASP@CG-HIS@ND1	3.15	3.35
ASN@CG-SER@CB	6.34	6.74
SER@OH-surface	1.58	1.60

Smallest (lower-range) and longest (upper-range) measures from experimentally determined structures used as templates in the modeling of PvSUB1. The asparagine (Asn) is the residue forming the oxyanion hole. Catalytic triads of PvSUB1, bacterial subtilisins Carlsberg (1R0R) and BPN (1TO2): Asp316, His372, Ser549; Asp137, His168, Ser325 and Asp139, His171, Ser328, respectively.

doi:10.1371/journal.pone.0109269.t002

Modeling and molecular dynamics simulations of PvSUB1

In order to generate a reliable 3D-model of PvSUB1, we used the procedure described in our previous publication where we modeled the structure of PfSUB1, a close homologous of PvSUB1 [14]. A similar homology modeling strategy generated 3D-models of PfSUB1 used to identify small-molecule inhibitors of PfSUB1 with an *in silico* screening approach [15]. The particular challenge of obtaining a high quality 3D model of PvSUB1 was the low sequence identity with the available templates (only ~30%, just above the “twilight zone” for homology modeling). Using state of the art modeling methods, it is possible to generate homology models with a C α RMSD < 1.0 Å when the sequence identity with the template is >50%. With sequence identity below 25%, larger divergences from the target structure can appear, making the model less precise [28]. However, our previous analysis had shown that major divergences between the PvSUB1 sequence and the structural templates were localized in regions distant from the catalytic groove that binds the substrate, and that the sequence identity in the substrate binding area is >30% [14,15]. In addition, to evaluate the stability and the overall quality of the PvSUB1 model (Figure 2), we performed multiple molecular dynamics simulations of our model PvSUB1 and of two of the templates used in the modeling, subtilisins BPN (1TO2) and Carlsberg (1R0R).

Tables 1 and 2 show distance ranges among the residues of the catalytic triad, in the MD simulations of PvSUB1 and in the crystal structures of the templates. All distances fell within the ranges observed in the experimental X-ray crystal structures apart from the distances HIS372@CB-ASP316@CB and ASP316@CG-HIS372@ND1, which are slightly outside (less than 1 Å). This is consistent with the fact that this distance shows the highest variation in subtilisin X-ray crystal structures (Table 2). Stability can be also measured by analyzing the RMSD along the MD trajectories from the starting structure. In Figure 3.A, we plotted the average RMSDs obtained for the 5 MD trajectories of 10 ns each.

The trajectories of the PvSUB1 model diverge more than those of BPN (1TO2) and Carlsberg (1R0R). This is primarily due to the regions in the model for which there is no structural information in the templates. In the model, these regions are unstructured, solvent exposed and distant from the binding pocket. When we removed them from the analysis, the RMSD reduced to values

similar to those observed in the MD trajectories of subtilisin BPN (1TO2). The trend of the RMSD from the average structure shows a stabilization of the model along the MD trajectories (Figure 3.B). The per-residue fluctuation (RMSF) in Figure 3.C shows some very flexible regions, for example the region 50–80, where structural information in the templates is absent.

Most of residues forming the binding region (orange rectangles) were much less flexible than the rest, similar to what we observed for the other two subtilisins (data not shown). The fact that the model of the binding region showed similar stability in MD simulations as the X-ray crystal structures is an indication that there are no major errors in the model of this region. Obviously, despite the care we took in generating and validating our model, there may be structural errors with an effect on the success of our computational design procedure. For this reason, we included an additional step to refine the model of the structure of PvSUB1.

Refinement of PvSUB1 model

To obtain a refined structure of the PvSUB1 model, we performed MD simulations of the complex of PvSUB1 and its substrate hexapeptide (P4-VGADDV-P2'). The hexapeptide was docked according to the X-ray structure of Subtilisin E with its pro-peptide (1SCJ), and refined with MD, where we restrained a few distances between the protein and the peptide (see [14] for details). Even though subtilisins do not undergo major conformational changes upon binding [29], small rearrangements might take place at the interface at the level of side-chains for example. The refinement allowed us to obtain bound-like conformations of PvSUB1, which facilitated the subsequent docking step (Figure 4). Figure 5 shows the catalytic triad Ser-His-Asp of the PvSUB1 model.

Docking of EETI-II-sub

The wild type EETI-II did not inhibit PvSUB1 (Table 3). We then replaced the EETI-II residues involved in the binding to the protease catalytic groove with the sequence of the PvSUB1 natural hexapeptide substrate. This EETI-II-sub mutant inhibited PvSUB1 activity with a K_i > 0.75 mM (Table 3). We derived snapshots (50 for each protein) for the ensemble docking of EETI-II-sub onto PvSUB1 by a cluster analysis (based on the residues at the interface) of the 5×2 ns molecular dynamics trajectories to obtain the best representative structures from the simulations.

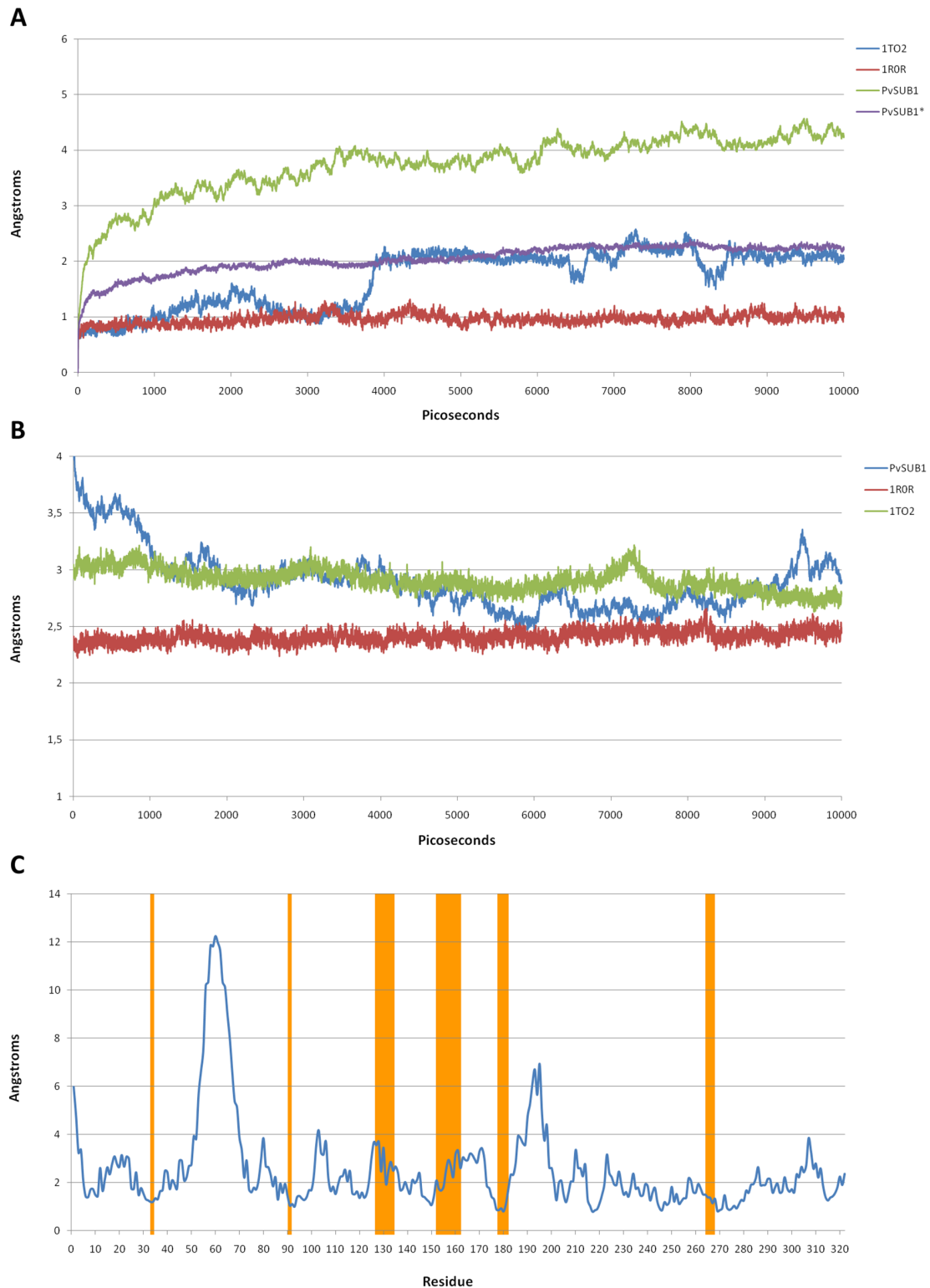


Figure 3. PvSUB1 molecular dynamics simulations. **A:** Average RMSD values for PvSUB1 and the 3D structure of two homologous bacterial subtilisins (1TO2, 1ROR). PvSUB1* shows the RMSD calculated without the regions missing template structural information; **B:** Fluctuation of the RMSD from the average structure. **C:** Root mean square fluctuation (RMSF) on a per-residue basis. In orange are highlighted PvSUB1 residues involved in the substrate-binding region.
doi:10.1371/journal.pone.0109269.g003

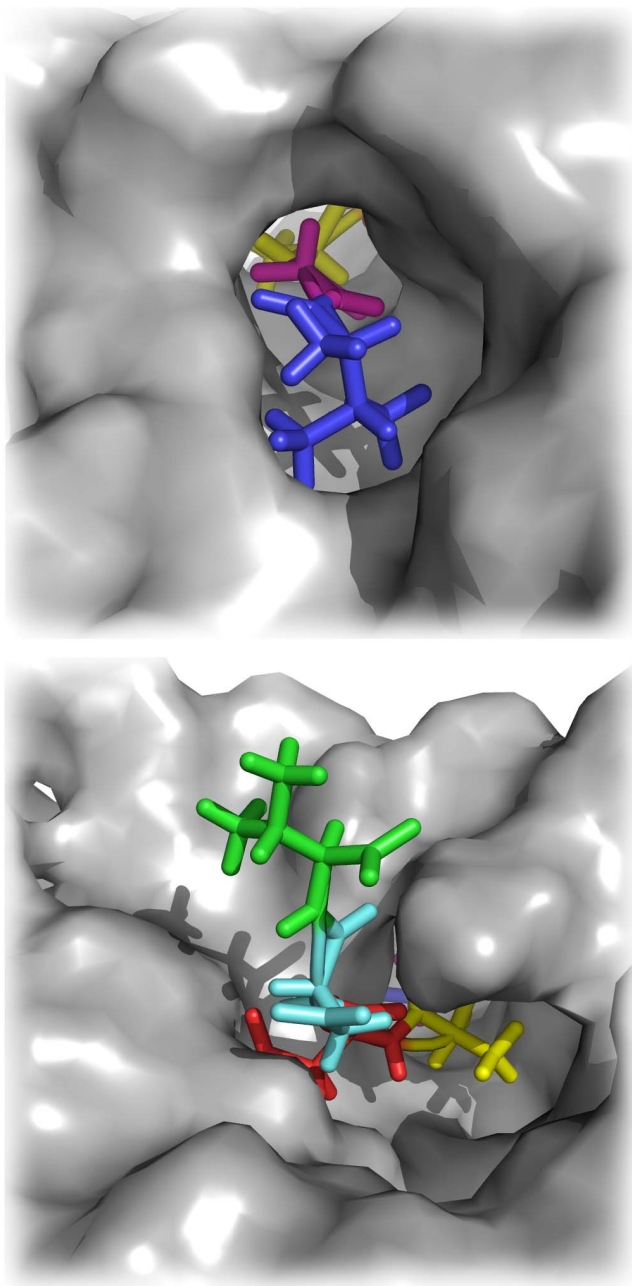


Figure 4. Docking of PvSUB1 hexapeptide substrate into PvSUB1 catalytic groove. Blue: P4, Violet: P3, Yellow: P2, Red: P1, Cyan: P1', Green: P2'.
doi:10.1371/journal.pone.0109269.g004

Figure 6 shows the score distribution among the solutions, with the 5 solutions with the highest geometric score (see Methods) highlighted with red circles. The relatively low geometric score indicated that most docking solutions were sub-optimal, although we had used an elaborate ensemble docking procedure. To optimize and refine the docking results we took the best 5 docking poses (red circles in 6) and refined them by a restrained molecular dynamics procedure similar to what we used for the refinement of the PvSUB1 model. This procedure helped re-establish native contacts when compared to regular unrestrained MD simulations. Subsequently we selected the solution that fulfilled all distances. Figure 7 shows the docked complex.

Scoring mutants

A preliminary free energy calculation was performed with snapshots from multiple MD simulations of EETI-II-sub docked onto PvSUB1 to obtain more consistent MM/GBSA results [30]. A free-energy decomposition (Figure 8) shows the contribution of each single residue to the total free energy of binding.

The biggest contribution to the free energy of binding came from the main-chain contacts of residues P4, P3, P2 and P1. This is in agreement with previous observations of important interactions between a protein-inhibitor and a serine protease active site, where important contacts are made by main-chain atoms [31]. For the case of EETI-II-sub the highest contribution originates from the cysteine in P3 and its main-chain, accounting for -4.34 kcal/mol.

We then tried to identify the most favorable mutations that could improve the binding affinity of EETI-II-sub to PvSUB1. The cysteine in P3 cannot be mutated because its side-chain is involved in a disulphide bridge that has an important function in stabilizing the EETI-II scaffold and maintaining the loop rigid, whereas the alanine in P2 already contributes with -4.17 kcal/mol to the total binding energy. We also looked at the parasite sequences that are natural substrates of PvSUB1 or PfSUB1 to suggest positions to introduce mutations in the EETI-II scaffold. Table 4 lists these sequences (experimentally derived or deduced from sequence alignments) for PfSUB1 and PvSUB1. While the sequences of several PfSUB1 substrates were experimentally determined [32,33], few were identified for PvSUB1 [15]. Considering the evolutionary proximity of *P. vivax* and *P. falciparum*, with active sites displaying $>60\%$ sequence identity [15], these predicted sequences can be considered reliable. Comparing the cleavage sites, we observed that only alanine and glycine appeared in P2, suggesting that only small residues are tolerated in this position. Position P1' has a negative contribution to the energy and therefore is an interesting position to mutate. However, considering the lack of a specific pocket for this residue (Figure 7) we can consider this position almost a secondary contact residue [31] and we decided to keep the wild-type residue. Finally, SUB1 cleavage site sequences have a fairly high similarity at the P1 and P4 positions (Table 4) and we therefore focused on these positions to mutate the EETI-II inhibitory loop. It is worth mentioning that the contribution of these P1 and P4 positions within the substrate-PfSUB1 interaction has recently been experimentally established [34].

We performed 10×100 ps MD simulations and MM/GBSA free energy calculations for all possible residues in position P4 and P1 independently, assuming that the effect of the two mutations would be additive. The free energy calculations for the mutants in P4 showed that hydrophobic and bulky residues were preferred for this position (Figure 9.A). This result fits with the fact that pocket S4 is composed of six hydrophobic residues (L131, M134, F153, I161, F162, P205) and seems to have enough space to accommodate larger hydrophobic side-chains than valine (Figure 10). Position P1 instead presents as favorable mutations aromatic residues with polar groups (Tyr, Trp), glutamate and positively charged residues (Lys, Arg). Surprisingly we found as favorable mutations some positively charged residues, whereas most of sequences recognized by the homologous PfSUB1 present negatively charged (Asp, Glu) or neutral polar (Gln, Asn) side-chains at P1. This might be explained by either the low substrate specificity common to some subtilisins or imprecisions in the structure of the complex.

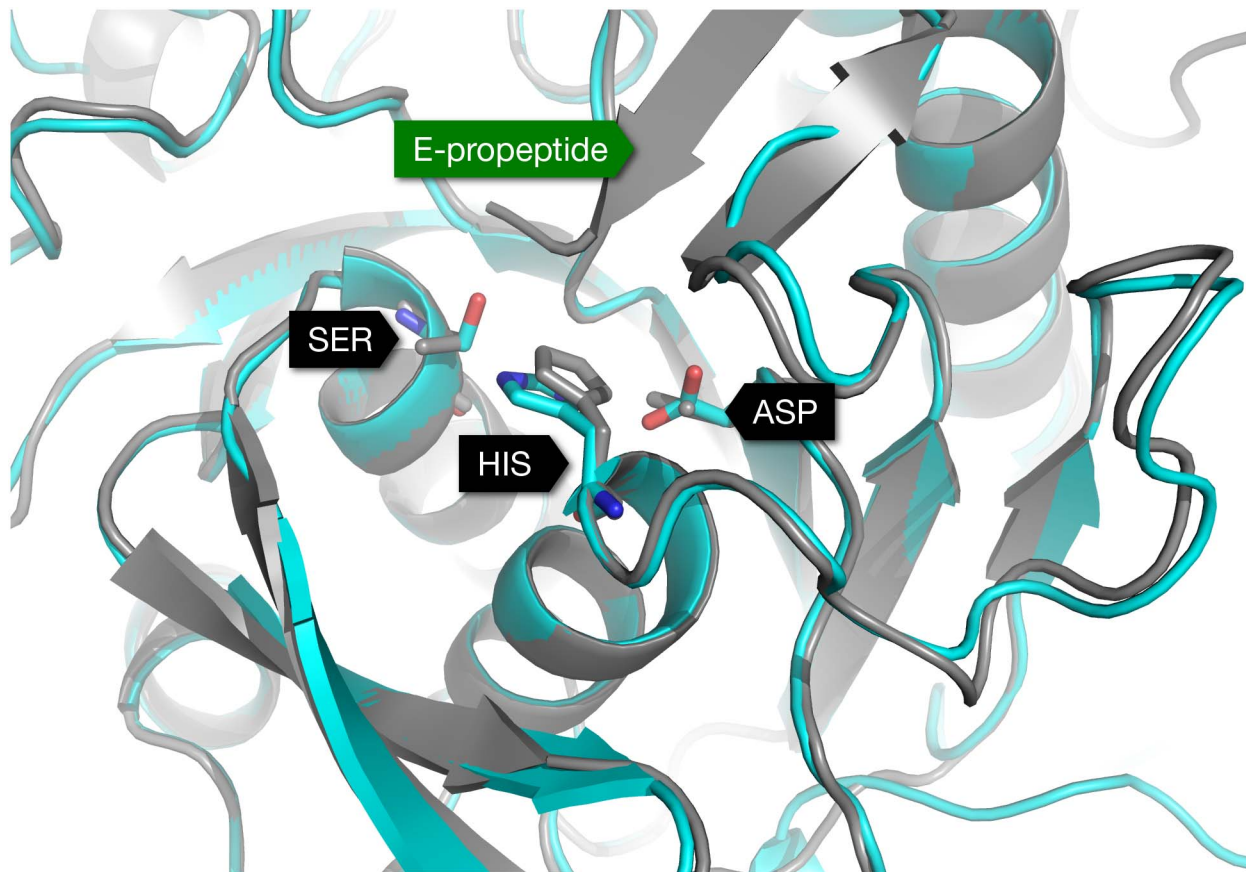


Figure 5. Structural alignment of the obtained PvSUB1 model (cyan) with the 3D-structure of Subtilisin E (gray, PDB 1SCJ) that was used as a template in the homology modelling. The catalytic triads in both proteins are highlighted with a stick representation. PvSUB1 catalytic triad: Asp 316, His 372 and Ser 549. Subtilisin E catalytic triad: Asp 32, His 64, Ser 221.
doi:10.1371/journal.pone.0109269.g005

Evaluation of EETI-II mutants on PvSUB1 enzymatic activity

All mutants of EETI-II were produced by chemical synthesis, folded and purified by reverse-phase HPLC. In Table 3 we

present the results of PvSUB1 inhibition of the different synthesized EETI-II mutants. We initially tested mutants in position P4 according to our scoring results and found that the EETI-II with a leucine in P4 (EETI-II-P4L) inhibits PvSUB1 with a K_i of 147 μ M, i.e., about one order of magnitude higher than the

Table 3. Sequence and inhibitory activity of EETI-II mutants on PvSUB1.

Name of tested EETI-II	EETI-II active site sequences $P_4 P_3 P_2 P_1 P'_1 P'_2$	K_i on PvSUB1 (mM)
EETI-II-WT	G C P R I L	NI
PvS1-WT	V C A D D V	>0.75
PvS1- P_{4W}	W C A D D V	>0.75
PvS1- P_{4P}	P C A D D V	>0.75
PvS1- P_{4M}	M C A D D V	>0.75
PvS1- P_{4L}	L C A D D V	0.15 ± 0.03
PvS1- P_{4I}	I C A D D V	0.6 ± 0.01
PvS1- $P_{4L} P_{1E}$	L C A E D V	0.34 ± 0.05
PvS1- $P_{4L} P_{1K}$	L C A K D V	0.39 ± 0.14
PvS1- $P_{4L} P_{1R}$	L C A R D V	0.75 ± 0.035
PvS1- $P_{4L} P_{1Y}$	L C A Y D V	0.24 ± 0.02
PvS1- $P_{4L} P_{1W}$	L C A W D V	0.08 ± 0.01

Active site sequences of the tested EETI-II mutants and their K_i for PvSUB1. NI: No Inhibition.
doi:10.1371/journal.pone.0109269.t003

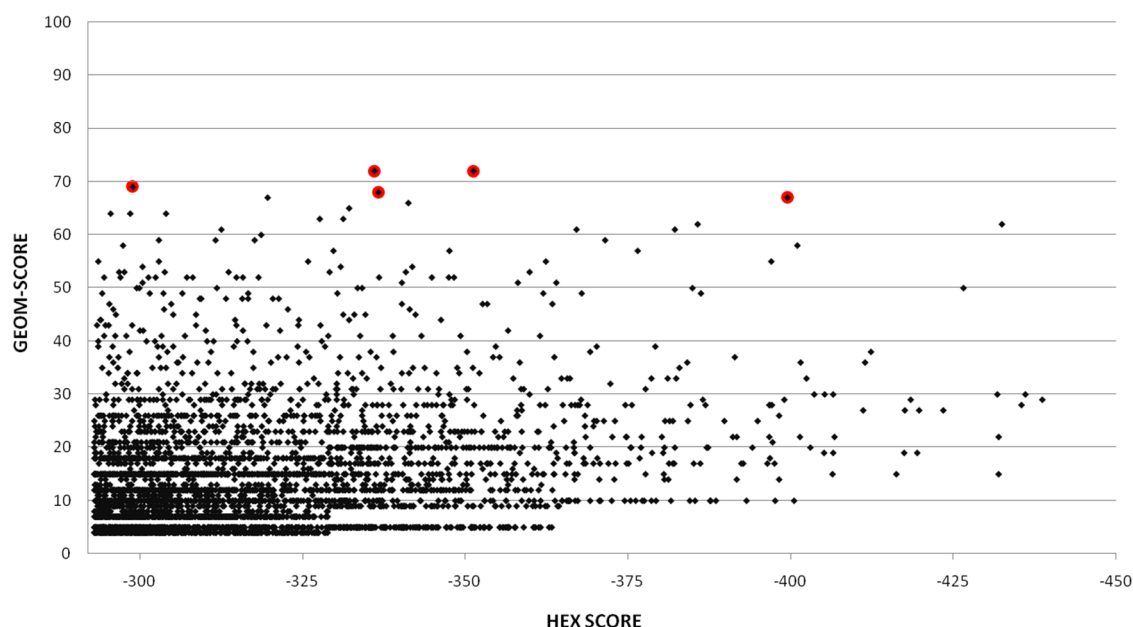


Figure 6. Docking results. The red circles indicate the docking poses that have been selected for refinement.
doi:10.1371/journal.pone.0109269.g006

valine (Table 3). This could be explained by the higher flexibility of the leucine and its bulkier side-chain, which could fit more tightly into the hydrophobic S4 pocket (Figure 10). The fact that the isoleucine mutant shows a K_i of 591 μ M, which is three times less than that of leucine, could be the effect of its beta-branched side-chain. We decided to keep the leucine mutant in P4 and test mutations in P1. Among all the EETI-II-P4L mutants for the P1 position, only one showed a better K_i of 86 μ M, the EETI-II-P4L-P1W, which contained a tryptophane in P1. All other mutants had higher K_i suggesting the importance of keeping the aspartate in this position. We found the comparison of the active mutants with the list of substrate sequences of PvSUB1 in Table 4 particularly intriguing. In position P4 of the substrate there is a preference for valine and threonine, while leucine is only present in 2 out of 18 sequences; the mutant with a leucine in P4 is more than 5x more active on PvSUB1 than the one with a valine (PvS1-WT). This might be explained by the presence of important structural constraints (disulphide bonds) that are present in EETI-II compared to a more flexible conformation of the substrate sequence [29]. According to the cleavage sites predicted by sequence alignment of the substrates, we observe a prevalence of negatively charged residues such as glutamate and aspartate while there was a preference for aromatic residues in the designed inhibitors, with the tryptophan having the best K_i . Even in this case, the explanation might lie in the structural constraints that are present in the inhibitors and/or a set of new interactions at position P1 (Figure 10).

Conclusions

Subtilisin-like proteins of *Plasmodium* are promising biological targets for developing novel therapeutics. One of these proteins, SUB1 plays an essential role in both the hepatic and erythrocytic stages of *Plasmodium*, making this enzyme a particularly interesting drug target. With the aim to develop an inhibitor of *Plasmodium vivax* SUB1 (PvSUB1), we redesigned a protein scaffold, the natural trypsin inhibitor EETI-II. Despite challenges in this project (the 3D structure of the target protein had to be

modeled from homologues with only 30% sequence identity), the computational procedure allowed us to predict mutants that proved to be inhibitors of PvSUB1 in experimental tests.

However, some predicted good inhibitors did not show any improvements in binding to PvSUB1 when experimentally tested. This might be caused by the flexibility in the binding as proteases and their S# pockets are inherently flexible and known to alter their shape to accommodate various substrates. Imprecisions in the homology model of PvSUB1 can also have obvious consequences for the precision of the docking and of the energetic analysis. The free energy analysis based on implicit solvent methods is in itself approximate and neglects important factors influencing binding. It is encouraging that, nonetheless, we have obtained protein-based inhibitors of PvSUB1 and this opens new ways to further improve our best mutants by computational or experimental protein engineering protocols. The protein design approach described in this work demonstrates the capabilities of computational procedures to accelerate and guide the design of novel proteins with potential future therapeutic applications.

Methods

Modeling PvSUB1

Models of PvSUB1 were generated and validated with the same protocol as in a previous study [14]. Binding pocket residues of PvSUB1 were defined according to the interaction of canonical inhibitors with subtilisins: 152–162, 127–134, 265–268, 179–182, 90 (His), 34 (Asp), 267 (Ser). Regions where structural information from template structures is lacking are 307–324, 1–9, 51–73, 190–199, 227–231. The models were further evaluated by performing molecular dynamic simulations and comparing their dynamic behavior to that of two subtilisins 3D structures used in the homology modeling (Carlsberg: 1R0R, BPN: 1TO2). The validated model of PvSUB1 was further refined in order to obtain bound-like conformations for the ensemble docking. We docked the hexapeptide of the sequence recognized by PvSUB1 to its structure, using 1SCJ (subtilisin E + pro-peptide) as template. We performed restrained molecular dynamic simulation to refine the

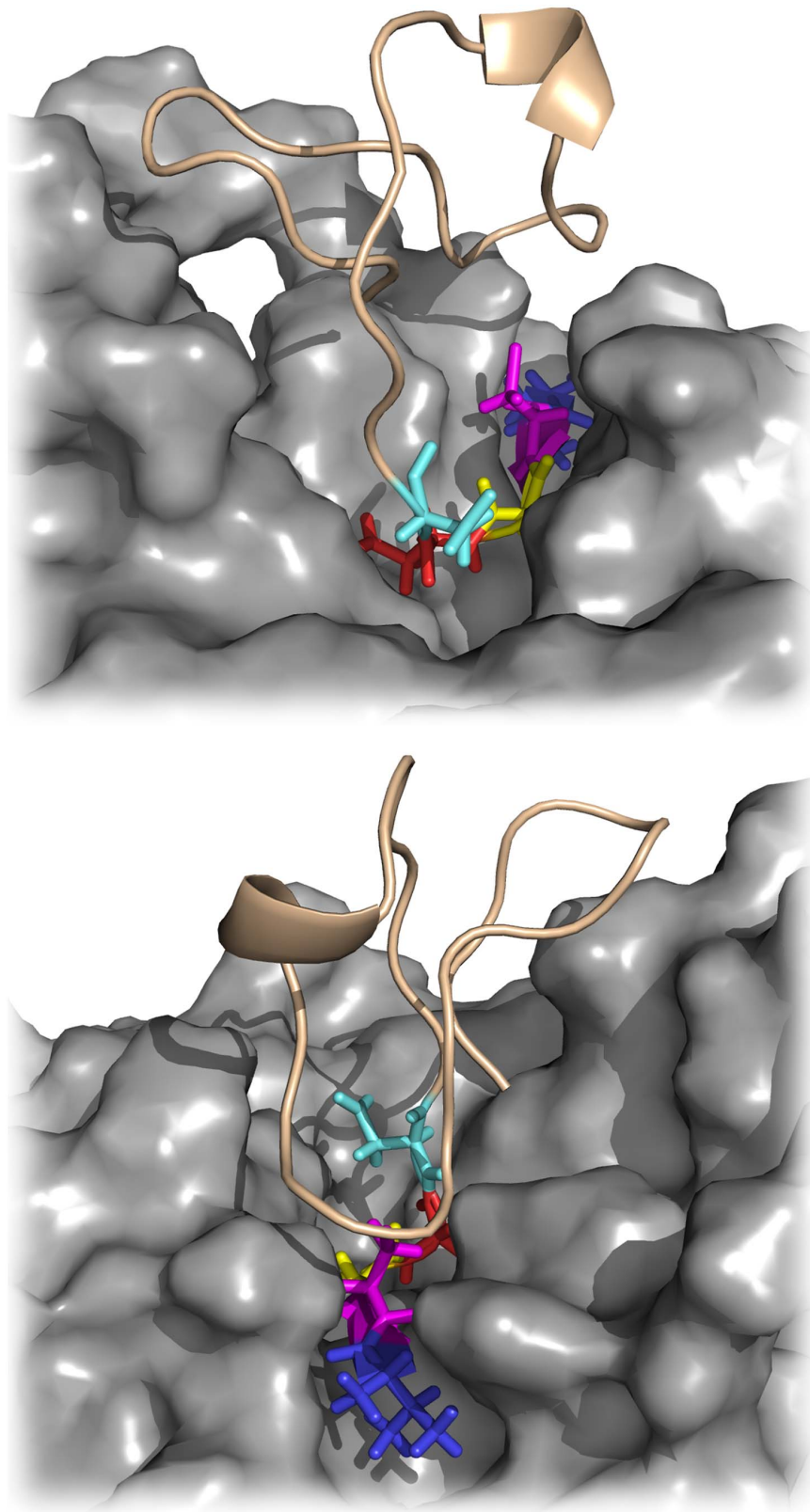


Figure 7. EETI-II-sub docked to PvSUB1. Blue: P4, Violet: P3, Yellow: P2, Red: P1, Cyan: P1'.
doi:10.1371/journal.pone.0109269.g007

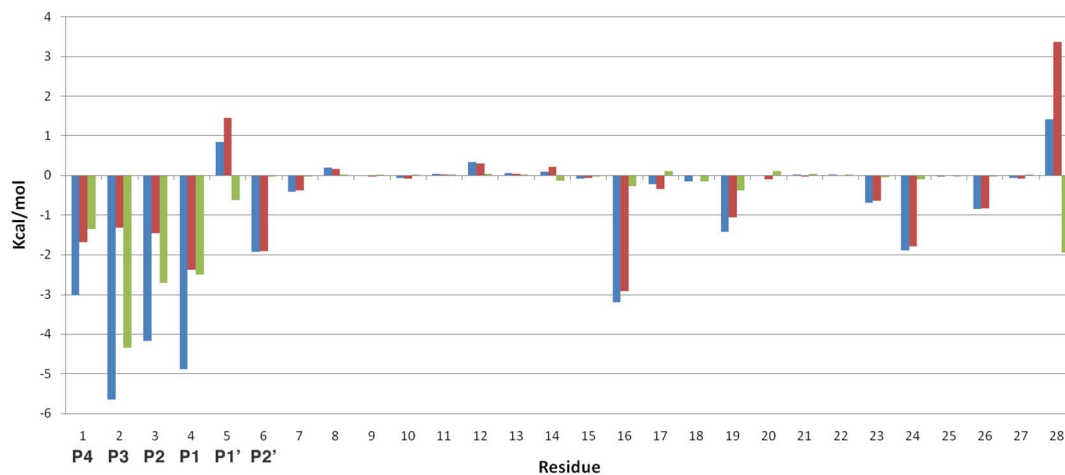


Figure 8. Free energy decomposition. Blue: All atoms, Red: Side chain atoms, Green: Backbone atoms. The largest contribution to the free energy of binding comes from the main-chain contacts of residues P4, P3, P2 and P1. The highest contribution comes from the cysteine in P3 and its main-chain, accounting for -4.34 kcal/mol.
doi:10.1371/journal.pone.0109269.g008

complex, using the protocol described in the Section “Docking Refinement”. Surface electrostatic distributions on models were calculated with the APBS [35] module implemented in Pymol.

EETI-II. The structure of EETI-II (2IT7) was retrieved from the PDB (<http://www.rcsb.org>) [36] and mutations at position P1 and P4 of the inhibitory loop were generated with the toolkit MMTSB [37]. We built the EETI-II substrate-like mutant (EETI-II-sub) by replacing its inhibitory loop (GCPRIL) with the sequence recognized by PvSUB1 (VGADDV). 3D images of the protein complexes were rendered with the molecular modeling software Pymol.

Molecular Dynamics (MD) Simulations

All molecular dynamics simulations were performed with the SANDER module from the AMBER9 [38] package and the force-field ff99SB [39]. After minimization *in vacuo* the complexes were hydrated with TIP3P [40] water molecules and neutralized by adding an appropriate number of monovalent counterions. The MD unit cell was a truncated octahedral box with a minimum distance of 10 Å between the solute and the cell boundary. To minimize the water molecules we ran a two stages minimization protocol in which we first applied positional restraints with an energy constant of 10 kcal/(mol Å²) to the solute followed by a stage with 1 kcal/(mol Å²) energy constant. Both minimization stages consisted of 10 steepest-descent and 490 conjugate-gradient steps. During the equilibration/heating and production dynamics all covalent bonds to hydrogen atoms were constrained with the SHAKE [41] algorithm, and we used a time step of 2 fs. We used periodic boundary conditions with a distance cutoff of 8.0 Å for the direct part of the non-bonded interaction and PME [42] (Particle Mesh Ewald) to account for long-range electrostatic interactions. The minimized system was then thermalized and equilibrated by heating from 0 to 300°K over 20 ps under constant-volume conditions followed by 10 ps at constant-pressure. The production MD phase was launched from the final configuration after equilibration with a relaxation time of 2.0 ps for heat bath coupling and a pressure relaxation time of 2.0 ps.

Docking

The pool of conformers used for the ensemble docking was obtained from multiple (5×2 ns) MD trajectories of PvSUB1 (receptor) and EETI-II (ligand). We extracted snapshots every 10 ps of MD for a total of 1000 snapshots of each protein and clustered these structures with a single-linkage algorithm implemented in GROMACS [43] (g-cluster tool), where the RMSD was calculated only for the binding interface residues. We selected for cross-docking the centroids of each of the 50 clusters identified. The ensemble docking was performed with the rigid-body docking software HEX [44] (version 4.5) where the search was restricted to the binding pocket by positioning EETI-II structures around the interface and limiting the search to 30° for twist range, receptor and ligand range. We used a shape + electrostatic correlation type while the other parameters of HEX were left as default. The best 5000 docking solutions according to the HEX score were selected and re-ranked by a mixed score based on the geometry of the interaction between a canonical loop inhibitor and a subtilisin. This score (geometric score) was composed for 40% of the HEX score (based on surface complementarity) and for 60% of an empirical score defined by the conserved distances between atoms in the inhibitory loop of a canonical inhibitor and a subtilisin. A mixed score permits to take into consideration the shape complementary and the conserved structural feature of the interaction. The best 5 solutions were selected for refinement.

Docking Refinement

The refinement is based on a protocol that uses restrained molecular dynamics simulations. The refinement procedure consisted of a total of 360,000 steps based on specific distance restraints at the ligand/receptor interface and used the NMR refinement tools in AMBER9. The chosen conserved distance restraints were 267SER@OG-P1@C (lower bounds = 2.5 Å, lower-intermediate = 3.0 Å, intermediate- upper = 4.0 Å, upper bounds = 5.5 Å), 154SER@HN-P3@O, 154SER@O-P3@HN, 129GLY@HN-P4@O, 129GLY@O-P4@HN, 127LYS@O-P2@HN (lower bounds = 1.5 Å, lower-intermediate = 2.0 Å, intermediate- upper = 3.0 Å, upper bounds = 3.5 Å). In all cases, an energy constant of 20 kcal/mol Å² was employed. The refinement consisted of three phases. In the first phase (120,000 steps), the

Table 4. SUB1 natural substrates.

<i>Plasmodium</i> protein containing a SUB1 processing site	$P_4P_3P_2P_1 \downarrow P'_1P'_2$	References
Pro-region first maturation site of PfSUB1	VSAD ↓ NI	[48]
Pro-region second maturation site of PfSUB1	VSAD ↓ NI	[48]
PfSERA1 site 1	IKAE ↓ AE	[6]
PfSERA2 site 1	TKGE ↓ DD	[6]
PfSERA3 site 1	VKAA ↓ SV	[6]
PfSERA4 site 1	ITAQ ↓ DD	[6]
PfSERA5 site 1	IKAE ↓ TE	[6] [49]
PfSERA6 site 1	VKAQ ↓ DD	[6]
PfSERA7 site 1	FKGE ↓ DE	[6]
PfSERA9 site 1	VKGS ↓ TE	[6]
PfSERA1 site 2	IYSQ ↓ ED	[6]
PfSERA2 site 2	IWGQ ↓ ET	[6]
PfSERA3 site 2	LYGQ ↓ EE	[6]
PfSERA4 site 2	VYGQ ↓ DT	[6]
PfSERA5 site 2	IFGQ ↓ DT	[6] [49]
PfSERA6 site 2	VHGQ ↓ SN	[6]
PfSERA7 site 2	ISAQ ↓ DE	[6]
PfSERA9 site 2	VHGQ ↓ SG	[6]
PvSERA1 site 1	TKGE ↓ DE	
PvSERA2 site 1	MKAQ ↓ DE	
PvSERA3 site 1	AKGE ↓ DE	
PvSERA4 site 1	RKAQ ↓ QQ	
PvSERA5 site 1 TKGE ↓ DE		
PfSERA5 site 3	VRGD ↓ TE	[6]
PfMSP1 clone 3D7 junction MSP1₈₃ -MSP1₃₀	LVAA ↓ SE	[50]
PfMSP1 clone 3D7 junction MSP1₃₀ -MSP1₃₈	ITGT ↓ SS	[50]
PfMSP1 clone 3D7 junction MSP1₃₈ -MSP1₄₂	VTGE ↓ AI	[50]
PfMSP1 clone FCB1 junction MSP1₃₀ -MSP1₃₈	VSAN ↓ DD	[51]
PfMSP1 clone FCB1 junction MSP1₃₈ -MSP1₄₂	VTGE ↓ AV	[51]
PcMSP1 junction MSP1₈₃ -MSP1₃₀	ATGE ↓ SE	[52]
PcMSP1 junction MSP1₃₀ -MSP1₃₈	VSAE ↓ SE	[52]
PcMSP1 junction MSP1₃₈ -MSP1₄₂	ANAA ↓ ST	[52]
PvMSP1 clones Sal1 and (Belem) junction MSP1 ₈₃ -MSP1 ₃₀	LRGA(S) ↓ SA	
PvMSP1 clones Sal1 and Belem junction MSP1 ₃₀ -MSP1 ₃₈	VGGN ↓ SE	
PvMSP1 clones Sal1 and Belem junction MSP1 ₃₈ -MSP1 ₄₂ TTGE ↓ AE		
PfMSP6 clone 3D7	VQAN ↓ SE	[31]
PfMSP7 clone 3D7 site 1	VKAQ ↓ SE	[31]
PfMSP7 clone 3D7 site 2	TQGQ ↓ EV	[31]
PfRAP-1 clone 3D7	IVGA ↓ DE	[32]
PfMSRP2 clone 3D7	LKGE ↓ SE	[32]
Pro-region first maturation site of PvSUB1	VGAD ↓ NI	[15]
Pro-region second maturation site of PvSUB1	SHAA ↓ SS	[15]
Pro-region third maturation site of PvSUB1	HLAG ↓ SK	[15]
Pro-region first maturation site of PbSUB1	VGAD ↓ SI	[15]
Pro-region first maturation site of PySUB1	VGAD ↓ SI	[15]

The table shows the sequences of the cleavage sites recognized by SUB1 in *Plasmodium falciparum* and *Plasmodium vivax*. Cleavage site sequences in bold characters have been experimentally determined, while the ones in normal characters are deducted from sequence alignments. The arrow indicates the site of cleavage between the P1 and the P1'. Cleavage site sequences have a fairly high similarity in particular at the P1, P2 and P4 positions.

doi:10.1371/journal.pone.0109269.t004

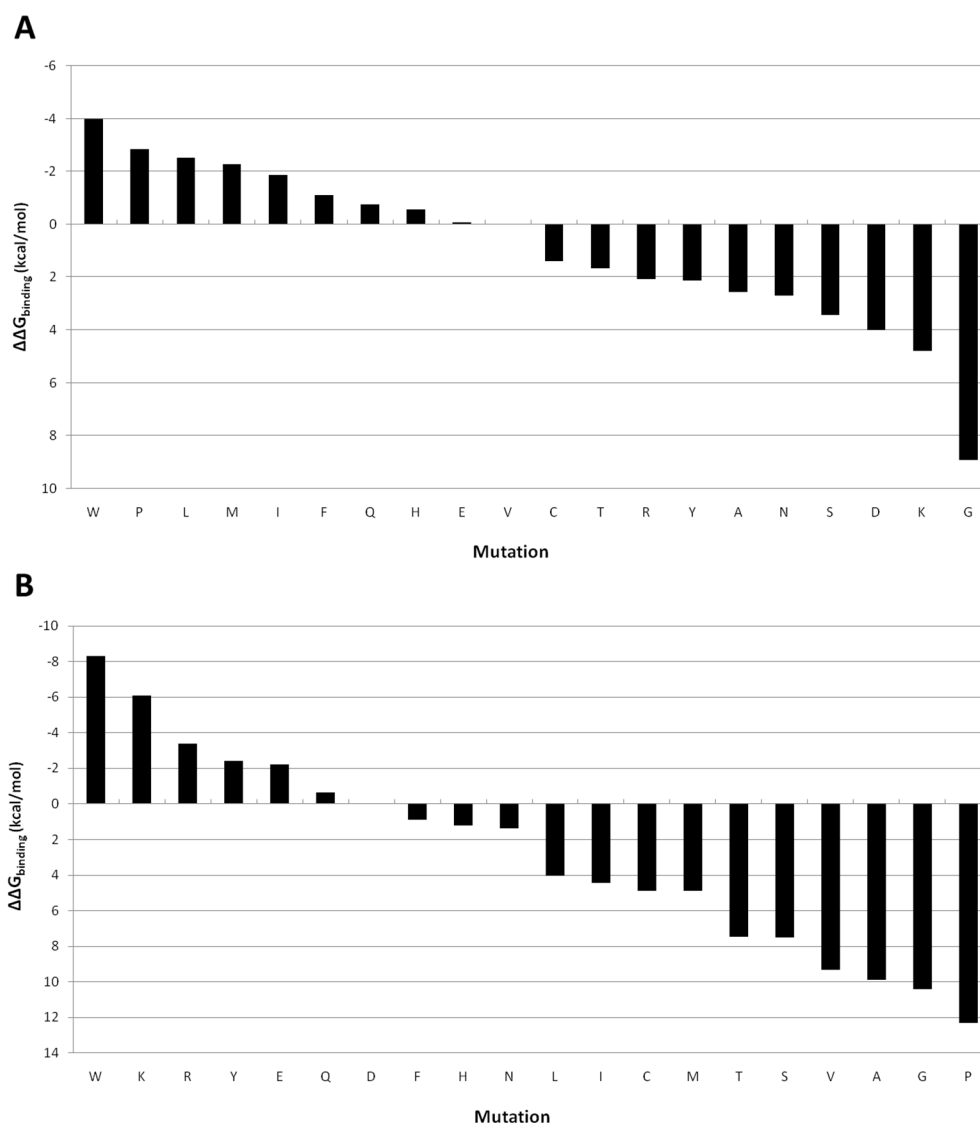


Figure 9. Scoring mutations on P4 and P1. **A:** mutants in position P4. The mutational profile of P4 shows that hydrophobic and bulky residues are preferred for this position. **B:** mutants in position P1. Position P1 instead prefers aromatic residues with polar groups (Tyr, Trp), glutamate and positively charged residues (Lys, Arg).
doi:10.1371/journal.pone.0109269.g009

receptor (PvSUB1) was kept rigid by applying a Cartesian restraint with an energy constant of 10 kcal/(mol Å²) and the distance restraints were switched on/off every 15,000 steps. In the second phase a Cartesian restraint with an energy constant of 10 kcal/(mol Å²) was applied only to heavy main-chain atoms keeping side-chains fully flexible, and in the third phase the energy constant of this Cartesian restraint was reduced to 0.01 kcal/(mol Å²). During these phases, the ligand (EETI-II or hexapeptide) was kept completely flexible. An additional 1 ns of regular MD simulation was performed to allow the system to relax into its final configuration. In the validation, this resulted in a high accuracy complex structure, with no distortions at the interface.

Scoring

The MM/GBSA protocol [45] in AMBER9 was used to calculate the relative free energy of each mutant. The default GBSA model used in the calculations was that of Tsui and Case [46] with an external dielectric of 80 and internal dielectric of 1.0.

For calculating the nonpolar contribution, the surface tension coefficient was set to 0.0072 and the surface offset to 0.0. The solvent accessible surface area was calculated with the ICOSA method. We calculated the relative free energy of binding from snapshots extracted each 10 ps from 10×100 ps trajectories.

Protein Production

EETI-II and mutants. All peptides (desalted, 35%–60% pure as assessed by HPLC) were obtained from GenScript Corporation, Piscataway, NJ, USA. In a typical procedure, peptide (50 mg) was dissolved in 75 mL of KH₂PO₄ buffer (0.2 M, pH 8.2) and allowed to air-oxidize at room temperature under gentle stirring. Monitoring was achieved with Ellman's test [47] and analytical HPLC (column ACE C18, 5 μm×4.6 mm, eluent A: 0.1% TFA/H₂O, eluent B: 60% CH₃CN/H₂O/0.1% TFA) with a 30 min linear gradient of 25% to 55% B at 1 mL flow rate (monitoring at 210 nm). When Ellman's tests are negative and the HPLC monitoring shows no more trace of starting materials (after

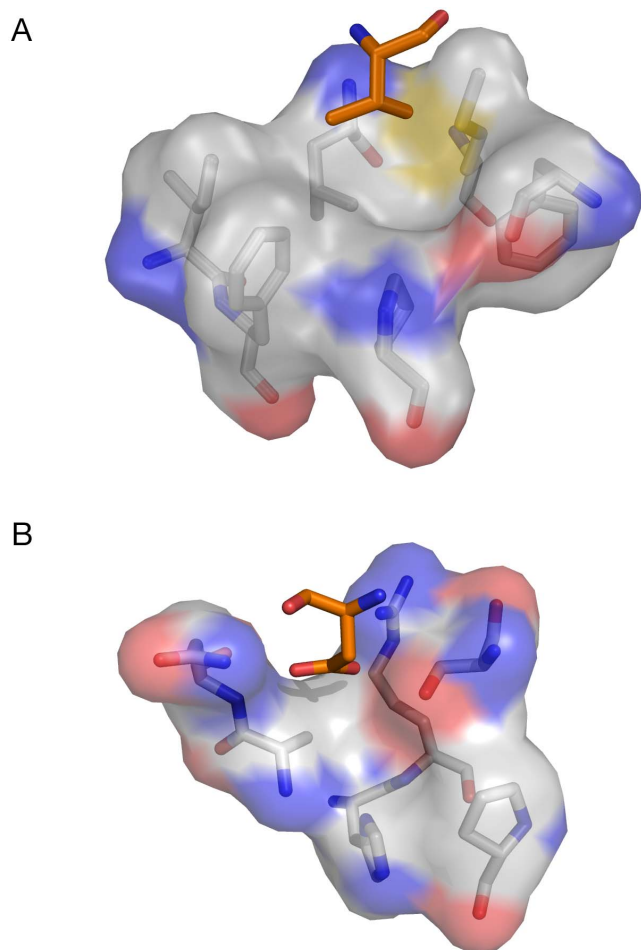


Figure 10. Residues forming the S1 and S4 pockets. The residue P4 (A) and P1 (B) of EETI-II are shown with an orange stick representation.

doi:10.1371/journal.pone.0109269.g010

3 to 5 days), the reaction mixture was centrifuged and the supernatant loaded onto a preparative HPLC column (Merck Lichrospher C₁₈, 10 μ m, 250 $\times$ 25 mm). Elution was achieved with a 90 min linear gradient of 25% to 55% B at 10 mL flow rate (monitoring at 220 nm). The fractions containing the oxidized peptide were combined and lyophilized to yield 15 to 30% of the desired peptide. Successful oxidation was confirmed by mass spectrometry (MALDI-ToF, Bruker Biflex III).

Production and purification of the PvSUB1 recombinant enzyme. The production and purification of the PvSUB1 (Genbank accession number FJ536585) recombinant enzyme was performed essentially as previously described [14,15]. Briefly for large-scale protein production, *Spodoptera frugiperda* Sf9 insect cells (1L at 3 $\times$ 10⁶ cells/mL, Invitrogen) were infected for 72 h with recombinant PvSUB1-recombinant baculovirus at a Multiplicity Of Infection (MOI) of 10 in Insect XPRESS medium (Lonza) supplemented with 50 μ g/mL gentamycin and 0.5 μ g/

mL tunicamycin (Sigma-Aldrich). Culture supernatant containing the secreted PvSUB1 recombinant protein was harvested, centrifuged 30 min at 2150 g to remove cells and cellular debris and concentrated/diafiltered against D-PBS 0.5 M NaCl, 5 mM Imidazole (loading buffer). The proteins were purified on an AKTA purifier system (GE Healthcare). The sample was loaded onto a 3 mL TALON Metal affinity resin (Clontech Laboratories) equilibrated in loading buffer. After extensive washes with loading buffer, the bound protein was eluted with a linear gradient of 5–200 mM imidazole in D-PBS, 0.5 M NaCl. Fractions containing PvSUB1 were pooled and concentrated by using Amicon Ultra 15 (10000 MWCO) and size-fractionated onto a HiLoad 16/60 Superdex 75 column equilibrated with 20 mM Tris pH 7.5, 100 mM NaCl. Fractions were monitored by absorbance (280 nm) and analyzed by Coomassie blue staining of SDS-PAGE gels and enzyme activity assay. The fractions containing the recombinant enzyme activity were pooled and the protein concentration was determined using the BCA Protein Assay following manufacturers recommendations (Bio Basic). Purified PvSUB1 recombinant protein was stored at -20°C following the addition of 30% v/v of pure Glycerol.

Enzymatic Test

For the kinetic assays we used the purified recombinant PvSUB1 enzyme and its specific peptide substrate whose sequence is deduced from PvSUB1 auto-maturation site: KLVGADDVSLA, with cleavage occurs between the two aspartates for PvSUB1. The KLVGADDVSLA sequence was coupled to the fluorophore/quencher dyes Dabsyl/Edans (Exc/Em 360/500 nm) at each edge. The enzymatic assays were performed in 20 mM Tris pH 7.5 and 25 mM CaCl₂ at 37°C as previously described [15]. For the determination of the K_i, the compounds, previously resuspended in ultra-pure distilled water at 10 mM, were tested at ten different concentrations ranging from 1 mM to 2 μ M following sequential 1:2 dilutions. The final mixture was distributed in duplicate into a 384-well black microtiter plate (Thermo Scientific) and the fluorescence was monitored every 3 minutes for 90 min at 37°C in a Labsystems Fluoroskan Ascent spectro-fluorometer. The slope of the linear part of the kinetic was determined in an Excel (Microsoft) spreadsheet. Every steps of the enzymatic assay were done on ice to make sure that the protein was not active before the measure of the fluorescence. The K_i and IC₅₀ values were determined (N = 3) using GraphPad Prism software.

Acknowledgments

We are grateful to Odile Puijalon for constant support and critical reading of the manuscript. This work is dedicated to the memory of our colleague Dung Le-Nguyen.

Author Contributions

Conceived and designed the experiments: GB DLN MN JCB. Performed the experiments: GB AB CN DLN. Analyzed the data: GB AB CN DLN MN JCB. Contributed reagents/materials/analysis tools: GB AB CN DLN MN JCB. Contributed to the writing of the manuscript: GB DLN MN JCB.

References

- Greenwood BM, Targett GA (2011) Malaria vaccines and the new malaria agenda. Clin Microbiol Infect 17: 1600–1607.
- Agnandji ST, Lell B, Soulanoudjingar SS, Fernandes JF, Abossolo BP, et al. (2011) First results of phase 3 trial of RTS,S/AS01 malaria vaccine in African children. N Engl J Med 365: 1863–1875.
- Noedl H, Socheat D, Satimai W (2009) Artemisinin-resistant malaria in Asia. N Engl J Med 361: 540–541.
- Alonso PL, Brown G, Arevalo-Herrera M, Binka F, Chitnis C, et al. (2011) A research agenda to underpin malaria eradication. PLoS Med 8: e1000406.
- Mishra SK, Newton CR (2009) Diagnosis and management of the neurological complications of *falciparum* malaria. Nat Rev Neurol 5: 189–198.

6. Yeoh S, O'Donnell RA, Koussis K, Dluzewski AR, Ansell KH, et al. (2007) Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes. *Cell* 131: 1072–1083.
7. Tawk L, Lacroix C, Gueirard P, Kent R, Gorgette O, et al. (2013) A Key Role for *Plasmodium* Subtilisin-like SUB1 Protease in Egress of Malaria Parasites from Host Hepatocytes. *J Biol Chem* 288: 33336–33346.
8. Keizer DW, Miles LA, Li F, Nair M, Anders RF, et al. (2003) Structures of phage-display peptides that bind to the malarial surface protein, apical membrane antigen 1, and block erythrocyte invasion. *Biochemistry* 42: 9915–9923.
9. Zhu S, Hudson TH, Kyle DE, Lin AJ (2002) Synthesis and in vitro studies of novel pyrimidinyl peptidomimetics as potential antimalarial therapeutic agents. *J Med Chem* 45: 3491–3496.
10. Dagan A, Efron L, Gaidukov L, Mor A, Ginsburg H (2002) In vitro antiplasmodium effects of dermaseptin S4 derivatives. *Antimicrob Agents Chemother* 46: 1059–1066.
11. Smith GP, Petrenko VA (1997) Phage Display. *Chem Rev* 97: 391–410.
12. He M, Taussig MJ (2002) Ribosome display: cell-free protein display technology. *Brief Funct Genomic Proteomic* 1: 204–212.
13. Roberts RW, Szostak JW (1997) RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc Natl Acad Sci U S A* 94: 12297–12302.
14. Bastianelli G, Bouillon A, Nguyen C, Crublet E, Petres S, et al. (2011) Computational reverse-engineering of a spider-venom derived peptide active against *Plasmodium falciparum* SUB1. *PLoS One* 6: e21812.
15. Bouillon A, Giganti D, Benedet C, Gorgette O, Petres S, et al. (2013) In Silico Screening on the Three-dimensional Model of the *Plasmodium vivax* SUB1 Protease Leads to the Validation of a Novel Anti-parasite Compound. *J Biol Chem* 288: 18561–18573.
16. Heitz A, Chiche L, Le-Nguyen D, Castro B (1989) 1 H 2D NMR and distance geometry study of the folding of *Ecballium elaterium* trypsin inhibitor, a member of the squash inhibitors family. *Biochemistry* 28: 2392–2398.
17. Craik DJ, Daly NL, Waine C (2001) The cystine knot motif in toxins and implications for drug design. *Toxicon* 39: 43–60.
18. Craik DJ, Clark RJ, Daly NL (2007) Potential therapeutic applications of the cyclotides and related cystine knot mini-proteins. *Expert Opin Investig Drugs* 16: 595–604.
19. Kolmar H (2009) Biological diversity and therapeutic potential of natural and engineered cystine knot miniproteins. *Curr Opin Pharmacol* 9: 608–614.
20. Le-Nguyen D, Mattras H, Coletti-Previero MA, Castro B (1989) Design and chemical synthesis of a 32 residues chimeric microprotein inhibiting both trypsin and carboxypeptidase A. *Biochem Biophys Res Commun* 162: 1425–1430.
21. Christmann A, Walter K, Wentzel A, Kratzner R, Kolmar H (1999) The cystine knot of a squash-type protease inhibitor as a structural scaffold for *Escherichia coli* cell surface display of conformationally constrained peptides. *Protein Eng* 12: 797–806.
22. Hilpert K, Wessner H, Schneider-Mergener J, Welfle K, Misselwitz R, et al. (2003) Design and characterization of a hybrid miniprotein that specifically inhibits porcine pancreatic elastase. *J Biol Chem* 278: 24986–24993.
23. Reiss S, Sieber M, Oberle V, Wentzel A, Spangenberg P, et al. (2006) Inhibition of platelet aggregation by grafting RGD and KGD sequences on the structural scaffold of small disulfide-rich proteins. *Platelets* 17: 153–157.
24. Souriau C, Chiche L, Irving R, Hudson P (2005) New binding specificities derived from Min-23, a small cystine-stabilized peptidic scaffold. *Biochemistry* 44: 7143–7155.
25. Wentzel A, Christmann A, Kratzner R, Kolmar H (1999) Sequence requirements of the GPNG beta-turn of the *Ecballium elaterium* trypsin inhibitor II explored by combinatorial library screening. *J Biol Chem* 274: 21037–21043.
26. Lippow SM, Wittrup KD, Tidor B (2007) Computational design of antibody-affinity improvement beyond in vivo maturation. *Nat Biotechnol* 25: 1171–1176.
27. Jha RK, Leaver-Fay A, Yin S, Wu Y, Butterfoss GL, et al. (2010) Computational design of a PAK1 binding protein. *J Mol Biol* 400: 257–270.
28. Chung SY, Subbiah S (1996) A structural explanation for the twilight zone of protein sequence homology. *Structure* 4: 1123–1127.
29. Otlewski J, Jelen F, Zakrzewska M, Oleksy A (2005) The many faces of protease-protein inhibitor interaction. *EMBO J* 24: 1303–1310.
30. Genheden S, Ryde U (2010) How to obtain statistically converged MM/GBSA results. *J Comput Chem* 31: 837–846.
31. Komiyama T, VanderLugt B, Fugere M, Day R, Kaufman RJ, et al. (2003) Optimization of protease-inhibitor interactions by randomizing adventitious contacts. *Proc Natl Acad Sci U S A* 100: 8205–8210.
32. Koussis K, Withers-Martinez C, Yeoh S, Child M, Hackett F, et al. (2009) A multifunctional serine protease primes the malaria parasite for red blood cell invasion. *Embo J* 28: 725–735.
33. Silmon de Monerri NC, Flynn HR, Campos MG, Hackett F, Koussis K, et al. (2011) Global identification of multiple substrates for *Plasmodium falciparum* SUB1, an essential malarial processing protease. *Infect Immun* 79: 1086–1097.
34. Fulle S, Withers-Martinez C, Blackman MJ, Morris GM, Finn PW (2013) Molecular Determinants of Binding to the *Plasmodium* Subtilisin-like Protease 1. *J Chem Inf Model*.
35. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98: 10037–10041.
36. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28: 235–242.
37. Feig M, Karanicolas J, Brooks CL, 3rd (2004) MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J Mol Graph Model* 22: 377–395.
38. Case DA, Darden TA, Cheatham TEI, Simmerling CL, Wang J, et al. (2006) AMBER9. University of California, San Francisco.
39. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, et al. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65: 712–725.
40. Jorgensen WL, Chandrasekhar J, Madura JD (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79: 926.
41. Ryckaert J, Ciccotti G, Berendsen H (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23: 327–341.
42. Darden T, York D, Pederson L (1993) Particle mesh ewald-an nlog(n) method for ewald sums in large systems. *J Chem Phys* 98: 10089–10092.
43. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, et al. (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 26: 1701–1718.
44. Ritchie DW, Kemp GJ (2000) Protein docking using spherical polar Fourier correlations. *Proteins* 39: 178–194.
45. Gohlke H, Kiel C, Case DA (2003) Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J Mol Biol* 330: 891–913.
46. Tsui V, Case DA (2000) Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* 56: 275–291.
47. Ellman GL (1959) Tissue sulfhydryl groups. *Arch Biochem Biophys* 82: 70–77.
48. Sajid M, Withers-Martinez C, Blackman MJ (2000) Maturation and specificity of *Plasmodium falciparum* subtilisin-like protease-1, a malaria merozoite subtilisin-like serine protease. *J Biol Chem* 275: 631–641.
49. Debrabant A, Maes P, Delplace P, Dubremetz JF, Tartar A, et al. (1992) Intramolecular mapping of *Plasmodium falciparum* P126 proteolytic fragments by N-terminal amino acid sequencing. *Molecular and Biochemical Parasitology* 53: 89–96.
50. Stafford WH, Blackman MJ, Harris A, Shai S, Grainger M, et al. (1994) N-terminal amino acid sequence of the *Plasmodium falciparum* merozoite surface protein-1 polypeptides. *Mol Biochem Parasitol* 66: 157–160.
51. Heidrich HG, Miettinen-Bauman A, Eckerskorn C, Lottspeich F (1989) The N-terminal amino acid sequences of the *Plasmodium falciparum* (FCBI) merozoite surface antigen of 42 and 36 kilodaltons, both derived from the 185–195 kilodalton precursor. *Molecular and Biochemical Parasitology* 34: 147–154.
52. O'Dea KP, McKean PG, Harris A, Brown KN (1995) Processing of the *Plasmodium chabaudi chabaudi* AS merozoite surface protein 1 in vivo and in vitro. *Molecular and Biochemical Parasitology* 72: 111–119.