Molecular Determinants of Binding to the *Plasmodium* Subtilisin-like Protease 1
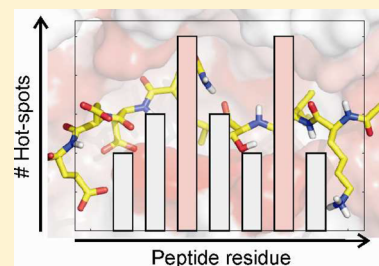
Simone Fulle,*,[†] Chrislaine Withers-Martinez,[‡] Michael J. Blackman,[‡] Garrett M. Morris,[†] and Paul W. Finn[†]

[†]InhibOx Ltd., Oxford Centre for Innovation, New Road, Oxford OX1 1BY, U.K.
[‡]Division of Parasitology, MRC National Institute for Medical Research, Mill Hill, London NW7 1AA, U.K.

Ⓢ *Supporting Information*

ABSTRACT: PfSUB1, a subtilisin-like protease of the human malaria parasite *Plasmodium falciparum*, is known to play important roles during the life cycle of the parasite and has emerged as a promising antimalarial drug target. In order to provide a detailed understanding of the origin of binding determinants of PfSUB1 substrates, we performed molecular dynamics simulations in combination with MM-GBSA free energy calculations using a homology model of PfSUB1 in complex with different substrate peptides. Key interactions, as well as residues that potentially make a major contribution to the binding free energy, are identified at the prime and nonprime side of the scissile bond and comprise peptide residues P4 to P2′. This finding stresses the requirement for peptide substrates to interact with both prime and nonprime side residues of the PfSUB1 binding site. Analyzing the energetic contributions of individual amino acids within the peptide-PfSUB1 complexes indicated that van der Waals interactions and the nonpolar part of solvation energy dictate the binding strength of the peptides and that the most favorable interactions are formed by peptide residues P4 and P1. Hot spot residues identified in PfSUB1 are dispersed over the entire binding site, but clustered areas of hot spots also exist and suggest that either the S4-S2 or the S1-S2′ binding site should be exploited in efforts to design small molecule inhibitors. The results are discussed with respect to which binding determinants are specific to PfSUB1 and, therefore, might allow binding selectivity to be obtained.

INTRODUCTION

Malaria remains one of the most important infectious diseases worldwide, causing 300−500 million clinical cases and over one million deaths each year. The disease results from infection with apicomplexan protozoan parasites of the genus *Plasmodium*. The most severe form of malaria is caused by infection with *Plasmodium falciparum*, but recent alarming reports have indicated that infections with both *Plasmodium vivax* and *Plasmodium knowlesi* can also be fatal.[1,2] Although there are several antimalarial drugs available, the emerging spread of multidrug-resistant parasite strains stresses the need to identify new targets that can be exploited with therapeutic agents.

A promising malarial drug target is a subtilisin-like protease, named PfSUB1 in *P. falciparum*, which is central to the replication of the parasite in human red blood cells (RBC). Each cycle of replication results in the formation of 16 or more daughter merozoites. PfSUB1 is thought to trigger a cascade of proteolytic events that culminate in the release of the merozoites from the infected RBC (a process called egress) and in addition is required for maturation of merozoite surface proteins to enable their invasion of new RBCs.[3,4] The potential of PfSUB1 as a drug target is supported by several findings. First, pharmacological inhibition of PfSUB1 blocks egress and/or reduces the invasion of RBCs by merozoites.[4−6] Second, a peptidyl α-ketoamide based on an authentic PfSUB1 substrate inhibits PfSUB1 as well as its orthologues in *P. vivax* and *P. knowlesi*.[7] This suggests that it should be possible to design

substrate-based compounds that inhibit the enzyme in all three major human pathogens. Third, several small molecules have so far emerged as selective inhibitors, including a natural compound called MRT12113,[5,8] covalent inhibitors based on chloroisocoumarins,[6] maslinic acid,[9] and a quinolylhydrazone.[10] Retesting of maslinic acid using a fluorescence assay, however, has questioned its activity against PfSUB1 (data not shown), and the hydrazone class shows a very flat structure−activity relationship in the 20 μM range.[10] Thus, there is still a need to identify small druglike molecule modulators amenable to chemical modification that inhibit the function of PfSUB1.

Insights into the determinants of binding interactions between an enzyme and its natural substrates can be exploited to facilitate the identification of small-molecule inhibitors for therapeutic intervention. Previous identification of several validated physiological substrates have provided some initial insights into the substrate preference of PfSUB1 (Figure 1) revealing a strong tendency for hydrophobic amino acids at P4, no preference at P3, a restriction to small amino acids at P2, a tendency for amide-containing and acidic amino acids at P1, and a tendency for hydroxyl-containing and acidic residues at the prime side positions P1′-P5′.[4,5,11] Mutational analysis of peptide substrates has implied, furthermore, that the P1 to P4 residues are not sufficient for substrate binding to PfSUB1,[4]
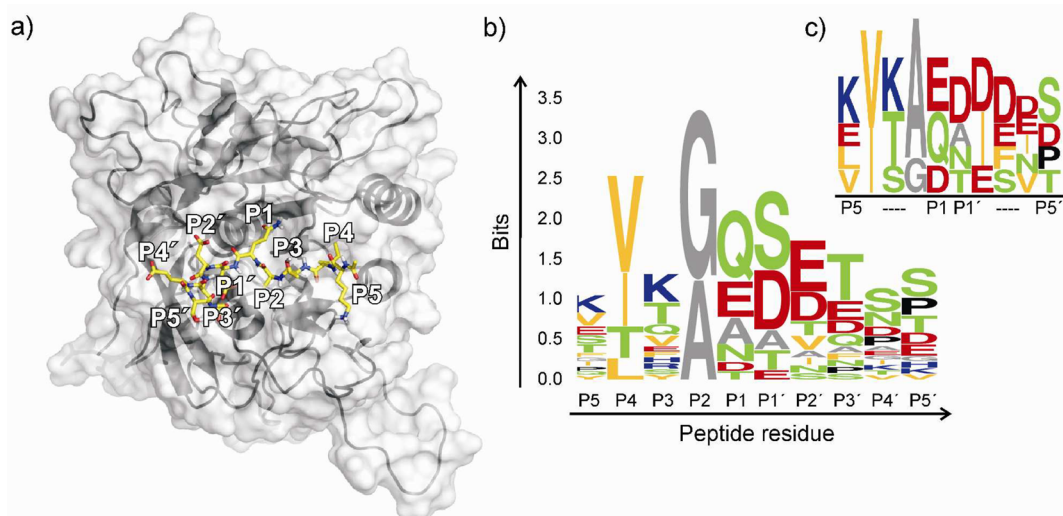
Figure 1. Structural model and substrate specificity of the active site of PfSUB1. (a) Homology model of a PfSUB1-peptide complex structure used as starting structure for the MD simulations.[7,42] The structural model is in line with the substrate specificity shown in (b), revealing distinct hydrophobic S4 and polar S1 pockets, a small uncharged S2 pocket (explaining the restriction to Ala or Gly at P2), and a less distinct S3 pocket. The latter is common in members of the subtilase family of serine proteases[29] resulting in the side-chain of P3 pointing away from the active site toward the solvent. (b) Sequence logo of the known PfSUB1 peptide substrates[4,11] reveals the substrate preference: hydrophobic amino acids at P4 (Ile, Leu, Val, Thr), no preference at P3, small amino acids at P2 (Ala, Gly), amide containing and acidic amino acids at P1 (Asp, An, Glu, Gln), and hydroxyl-containing polar and acidic residues at the prime side positions P1'-P5' (Ser, Thr, Asp, Glu). (c) For comparison, sequence logo of the PfSUB1 peptide substrates used in this study.

suggesting that prime side residues are required for optimal substrate recognition. Knowledge about the substrate preference, however, does not provide direct insights into the binding determinants of the PfSUB1 binding site. Important questions that remain include the following: Which residues of the PfSUB1 binding site are essential for binding affinity and should, therefore, be exploited in compound design efforts? What are the key interactions with known substrates that can be used to guide the identification and optimization of small molecule inhibitors?

In order to provide a better understanding of the binding modes and the origin of the binding determinants of PfSUB1 substrates, we have performed molecular dynamics (MD) simulations in combination with MM-GBSA free energy calculations of a homology model of PfSUB1 in complex with different deca-peptide substrates (Table 1). The peptides were chosen based on validated substrates, including members

of the serine-rich antigen (SERA) family and merozoite surface proteins (MSP). In addition, PfSUB1-peptide complex structures with sequence variations at the prime side (e.g., Ala mutations) were simulated in order to investigate the stabilizing effect of prime side residues. The present results extend an initial MD study performed by us where the formation of hydrogen bonds to the SERA4 peptide substrate was investigated.[7] Here, we have significantly extended that previous analysis by investigating the effect of sequence variations of PfSUB1-peptide complex structures, by analyzing key hydrogen bonds formed between peptide substrates and the PfSUB1 binding cleft, and by investigating energetic contributions of individual residues to binding. The MD results provide critical insights into the binding determinants of the PfSUB1 structure, which are currently being used to guide structure–activity experiments of peptidic compounds and which will facilitate future inhibitory compound design efforts in general.

Table 1. List of Simulated PfSUB1-Peptide Complex Structures[a]

native sequences	processing site identity	notes
LVSAD↓NIDIS	PfSUB1	internal cleavage site
KITAQ↓DDEES	SERA4 site 1	most efficiently cleaved peptide
EIKAE↓TEDDD	SERA5 site 1	variation at P1 and P1'-P5'
KVKAQ↓DDFNP	SERA6 site 1	P1' = D; P3' not polar
VVTGE↓AISVT	MSP1-42	P1' not polar
prime side variations		notes
KITAQ↓ADEES		single Ala mutation
KITAQ↓ADAES		double Ala mutation
KITAQ↓ADDES		double mutation
KITAQ↓AAAAA		quintuple Ala mutation
KITAQ↓D		

[a]Downward-pointing arrow denotes the scissile bond.

■ MATERIALS AND METHODS

Preparation of MD Starting Structures. Starting structures for the MD simulations were derived using a homology model of the catalytic domain of PfSUB1[12] and modeling substrate peptides by superposition of the backbone coordinates of homologue complexes (PDB codes: 1LW6[13] and 1MEE[14]). Side-chain atoms in 1LW6 and 1MEE were kept whenever possible in order to guide the peptide side-chain conformations. In more detail: P5-P2 were modeled by using the corresponding residues in 1LW6 (TIVL) and P1-P5' were modeled by using the corresponding residues in 1MEE (D↓LRYN). Remaining missing atoms in the side-chains were built by mutation of the corresponding amino acid using tleap from the Amber 11 package.[15] The peptides were then minimized using MOE in the complex structures using the MMFF94x force field, keeping the receptor structure fixed, and tethering the heavy atoms with a force constant of 100 kcal/(mol Å²). To

avoid terminal charges on the peptides, the N- and C-terminal residues were capped with acetyl (ACE) and *N*-methylamine (NME) groups, respectively.

Setup of MD Simulations. MD simulations of PfSUB1 bound to ten different peptide substrates (Table 1) were performed with the Amber 11 suite of programs[15] together with the ff99SB modifications[16,17] of the Cornell et al. force field.[18] In all cases, the system was neutralized by adding sodium counter-ions and solvated in a truncated octahedron box of TIP3P water molecules,[19] forming a solvent shell of at least 11 Å between each face of the box and the solute. The systems were minimized by 250 steps of steepest descent minimization followed by 250 steps of conjugate gradient minimization. The particle mesh Ewald (PME) method[20] was used to treat long-range electrostatic interactions, and bond lengths involving hydrogen atoms were constrained using the SHAKE algorithm.[21] The integration time step for all MD simulations was 2 fs, with a direct-space nonbonded cutoff of 9 Å. After minimization, MD in the canonical ensemble (NVT) was carried out for 50 ps, during which time the systems were heated from 100 to 300 K. Harmonic restraints with force constants of 5 kcal/(mol Å^2) were applied to all receptor and peptide atoms in this step. Subsequent isothermal isobaric ensemble (NPT)-MD was performed for 50 ps to adjust the solvent density. Finally, the force constants of the harmonic restraints on the receptor atoms were gradually reduced to zero over 250 ps in the NVT ensemble. An additional 50 ps of unconstrained NVT-MD at 300 K with a time constant of 2.0 ps for heat bath coupling were performed to relax the system without constraints. The production runs of all simulations achieved lengths of 50 ns of which snapshots saved at 20 ps intervals of the last 40 ns were used for analysis of hydrogen bonds and calculation of effective binding free energies.

Analysis of MD Trajectories. The 'ptraj' module of Amber 11 was used for analyzing the root-mean square deviation (RMSD) between structure pairs, the root-mean square fluctuations (RMSF) about the mean position of atoms, and the formation of hydrogen bonds. For RMSD and RMSF calculations, overall translational and rotational motions were removed with respect to all heavy atoms of the core region of the respective PfSUB1-peptide complex structures; i.e. the six surface-exposed loops were not included (excluded residues: 384−392, 396−411, 520−534, 543−551, 561−574, and 639−646). RMSD were accordingly calculated for all heavy atoms of the core region of the respective PfSUB1-peptide complexes. RMSF values were calculated for the backbone atoms of the peptide structures.

Hydrogen bonds were defined by a distance cutoff of 3.2 Å and an angle cutoff of 120°. Hydrogen bonds were only considered if their occupancies attained > 20% (percent of simulation time in which the hydrogen bond is formed). PfSUB1 residue numbering used in this study refers to the *P. falciparum* 3D7 sequence (PlasmoDB ID PF3D7_0507500; previous ID: PFE0370c).

Calculation of Effective Binding Free Energies and Per-Residue Contributions. MM-GBSA calculations[22−24] were carried out following the "single trajectory method", where snapshots of the binding partners were extracted from MD trajectories of PfSUB1-peptide complexes. The single trajectory method neglects energetic contributions due to conformational changes but leads to a drastic reduction in the statistical uncertainty of the free energy components.[22] The basic idea of the MM-GBSA approach is that the free energy of

binding can be calculated by considering only the end points of the thermodynamical cycle of ligand binding (bound and free states).

All counterions and water molecules were stripped from the snapshots and the analysis performed using the MM-PBSA Perl script provided in the Amber 11 suite of programs.[15] The binding free energy, ΔG_{bind}, can be calculated as

$$G_{bind} = G_{complex} - G_{protein} - G_{ligand}$$
$$= \Delta E_{MM} + \Delta G_{GB/PB} + \Delta G_{SA} - T\Delta S \qquad (1)$$

where ΔE_{MM} is the gas-phase interaction energy between the PfSUB1 receptor and the peptide including the electrostatic and van der Waals energies, $\Delta G_{GB/PB}$ and ΔG_{SA} are the electrostatic and nonpolar contributions to desolvation upon peptide binding, respectively, and $T\Delta S$ are the entropy contributions arising from changes in the degrees of freedom of the solute molecules, which were not considered here, i.e., all values reported for the MM-GBSA calculations should thus be considered as "effective energies" (ΔG_{Eff}) rather than free energies. For comparison, MM-PBSA calculations were also carried out.

In order to detect hot spot residues, the effective binding energies were decomposed into contributions of individual residues using the MM-GBSA energy decomposition scheme introduced by Gohlke et al.[25]

MM-GBSA Calculations. For each snapshot, ΔE_{MM} was calculated based on the ff99SB force field[17] without applying any nonbonded cutoff. ΔG_{GB}, the polar contribution to the solvation free energy, was determined by applying the 'OBC' Generalized Born (GB) method (*igb* = 2) and using mbondi2 radii. $\Delta G_{GB/PB}$ calculations are sensitive to the solute dielectric constant. Following findings from Hou et al.[26] for moderately charged binding interfaces, as present in the PfSUB1 binding cleft, the internal dielectric constant was set to 2 and the external dielectric constant to 80. The effect of the chosen internal dielectric constant on the predicted binding free energies was, furthermore, investigated by setting different internal dielectric constants (e_{in} = 1, 2, 3, and 4). The polar contributions were computed at 100 mM ionic strength (*saltcon* = 0.1 M). ΔG_{SA}, the nonpolar contribution to the solvation free energy, was estimated using the ICOSA method (*gbsa* = 2) by a solvent accessible surface area (SASA)-dependent term using a surface tension proportionality constant of γ = 0.0072 kcal/ (mol Å^2) and an offset of 0 kcal/mol. The autocorrelation function of the free energy was calculated,[22] using snapshots saved at 200 fs intervals of the 11 ns, to determine the ideal frequency to extract snapshots to avoid analyzing correlated structures.

MM-PBSA Calculations. ΔE_{MM} was calculated in the same way as in the MM-GBSA calculation (using the ff99SB force field). ΔG_{PB} was determined by solving the linearized Poisson−Boltzmann (PB) equation (*proc* = 2) using Parse radii and a solvent probe radius of 1.4 Å. A dielectric constant of 2 and 80 for the interior and exterior of the solute was applied, respectively. The polar contributions were computed at 100 mM ionic strength (*istring* = 100 mM). ΔG_{SA} was estimated using the NPOTP method by a solvent accessible surface area (SASA)-dependent term using a surface tension proportionality constant of γ = 0.00542 kcal/(mol Å^2) and an offset of 0.92 kcal/mol.
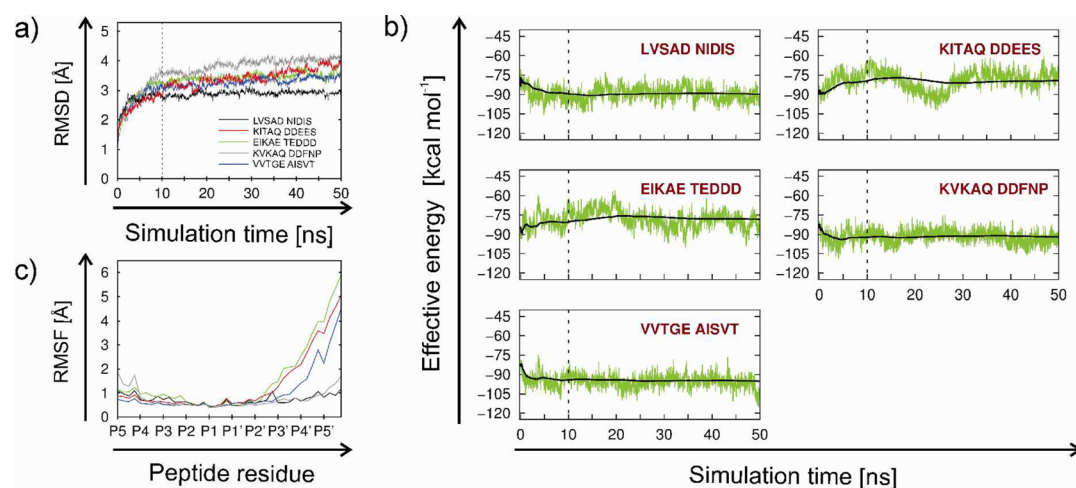
Figure 2. Convergence and stability examination for MD simulations of PfSUB1-peptide complex structures. (a) Time series of RMSD values of PfSUB1-peptide complexes are shown with respect to structures obtained at the end of the equilibration procedure. (b) Effective binding free energies (green) are shown together with accumulated mean values (red) for PfSUB1-peptide complexes. (c) Atomic fluctuations (RMSF) of peptide backbone atoms during MD trajectories of the respective PfSUB1-peptide complex structure. Color-coding of PfSUB1-peptide complexes is as in (a). Vertical lines in (a) and (b) indicate the time after which snapshots were extracted for further analysis. Equivalent plot for PfSUB1-peptide complex structures with sequence variations at the prime side of the KITAQ↓DDEES peptide are shown in Figure S1 in the Supporting Information.

Table 2. Binding Free Energies and Selected Individual Energy Contributions of PfSUB1-Peptide Complex Structures[a]

sequence	processing site identity	ΔG_{Eff}	$\Delta G_{nonpolar}$	ΔG_{GBELE}
LVSAD↓NIDIS	PfSUB1	−89.6 ± 0.1	−94.6 ± 0.1	5.0 ± 0.1
KITAQ↓DDEES	SERA4 site 1	−79.1 ± 0.2	−86.3 ± 0.2	7.2 ± 0.1
EIKAE↓TEDDD	SERA5 site 1	−77.8 ± 0.2	−84.1 ± 0.2	6.3 ± 0.1
KVKAQ↓DDFNP	SERA6 site 1	−91.8 ± 0.1	−102.9 ± 0.1	11.2 ± 0.1
VVTGE↓AISVT	MSP1−42	−95.3 ± 0.1	−104.1 ± 0.1	8.8 ± 0.1

[a]All values are given in kcal/mol (±standard error of the mean); calculated for trajectory range 10–50 ns. ΔG_{Eff} = sum of the gas-phase interaction energy and the electrostatic and nonpolar contributions to desolvation upon peptide binding ($\Delta G_{Eff} = \Delta G_{MM} + \Delta G_{GB} + \Delta G_{SA}$). $\Delta G_{nonpolar}$ = sum of van der Waals contribution from the molecular mechanics force field and the nonpolar contribution to the solvation free energy ($\Delta G_{nonpolar} = \Delta G_{MM-VDW} + \Delta G_{GBSUR}$). ΔG_{GBELE} = sum of electrostatic energy as calculated by the molecular mechanics force field and the electrostatic contribution to the solvation free energy ($\Delta G_{GBELE} = \Delta G_{MM-ELE} + \Delta G_{GB}$).

■ RESULTS

In order to investigate the binding mode and the binding determinants of PfSUB1 substrates, MD simulations of PfSUB1 bound to five known peptide substrates were carried out in combination with free energy calculations. The investigated peptides were LVSAD↓NIDIS (an internal cleavage site of PfSUB1; scissile bond indicated by a downward-pointing arrow), KITAQ↓DDEES (SERA4 site 1; the most efficiently cleaved peptide identified to date), EIKAE↓TEDDD (SERA5 site 1), KVKAQ↓DDFNP (SERA6 site 1), and VVTGE↓AISVT (MSP1-42). Previous results pointed to the unusual requirement of PfSUB1 to interact with both prime and nonprime side residues of the substrate recognition motif.[4,7] In order to further investigate the role of the P1′-P5′ residues, several sequence variations at the prime side of the KITAQ↓DDEES peptide were also investigated (Table 1).

Convergence and Stability Examination. The convergence and stability of the simulations were monitored through the examination of structural and energetic properties, which included the root-mean-square deviation (RMSD) of heavy atoms with respect to structures obtained at the end of the equilibration procedure and the effective free energy during the MD simulations (Figure 2 and Figure S1 in the Supporting Information).

RMSD Examination. PfSUB1 contains six surface-exposed loops that undergo large conformational changes during the MD simulations. Considering the remaining part of the PfSUB1 receptor structures, including the peptide binding site, the RMSD for all complex structures initially rises to 3.8 Å during the first 5−10 ns but remains constant for the remainder of the simulations (Figure 2 a)). The initial structures were manually modeled, which justifies the longer MD equilibration compared to those generally found during other MD simulations.

Free Energy Examination. Fluctuations of effective binding energies for the PfSUB1-peptide complex structures are shown together with cumulative mean values in Figure 2 b). The effective energies are quite variable, but the accumulated mean values become stable for most cases after 10 ns of simulation time (Figure 2 b)). This is consistent with the RMSD analysis in Figure 2 a). Thus, the following analysis of binding energies, binding determinants, and hydrogen bond formation is based on snapshots obtained after 10 ns of the production runs.

To obtain reliable estimates of binding energies, the snapshots used for the binding free energy evaluation must be independent and the averaged free energy values must be converged. Encouragingly, both aspects are fulfilled: First, as the correlation time for the "effective energies" for all systems is about 1−2 ps, the snapshots used for the binding free energy evaluation, which were extracted at time intervals of 20 ps, should be independent.[22] Second, the plots in Figure 2 b)

indicate rather stable time series for effective energies for the last 40 ns implying that the obtained averaged values are converged. This is supported by a low standard error for the calculated mean energies of ~0.2 kcal/mol (Table 2). Thus, there was no need to extract snapshots more frequently from the trajectories.

Overall, the calculated effective energies of complex formation amount to high negative values (e.g., for the KITAQ↓DDEES PfSUB1-peptide complex to −79.1 ± 0.2 kcal/mol (Table 2)) indicating that favorable protein-peptide complexes are formed. The calculated values, however, overestimate the binding free energy which can be partly explained due to two missing contributions: the lack of entropic contributions, which can be expected to be unfavorable in the case of the flexible peptides; and the lack of energetic contributions due to conformational changes, which were not considered here because of the use of the single trajectory approach.

A previous study from Hou et al.[26] showed that MM-PBSA performed better in calculating absolute binding free energies but that MM-GBSA performed better in calculating relative free energies. For comparison, effective energies were also calculated by the MM-PBSA approach. The predicted values resulted in general in lower effective energy values (by ca. −17 kcal/mol) but did not change relative binding free energies compared to the MM-GBSA approach (Table S1 in the Supporting Information). Considering in addition the computational efficiency of the MM-GBSA approach, we decided to calculate the free energy decompositions using the MM-GBSA approach.

Hydrogen Bonds Formed via the Peptide Backbone Include P4-P2′. To seek initial insights into the binding determinants of the PfSUB1-substrates, we investigated hydrogen bonds formed between PfSUB1 and the simulated peptide substrates along the MD trajectories (Figures 3 and 4). The analysis shows that in at least four of the five peptide complexes very strong canonical backbone hydrogen bonds (with an occupancy of at least 60%, indicated in orange and red; Figure 3) are formed between the amino and carbonyl group of P4 and the corresponding backbone groups of Gly467, between the amino group of P2 and the carbonyl group of Lys465, and between the amino group of P1 and the carbonyl group of Ser490 as well as between the amino group of P2′ and Asn603 in the S2′ pocket (Figure 3). Further strong canonical hydrogen bonds are formed, in three out of the five peptide complexes, between the amino and carbonyl group of P3 and Ser492 and between the carbonyl group of P1 and either the amino group of Thr605 or Ser606 (Figure 3). Interestingly, strong backbone-backbone (canonical) hydrogen bonds formed with P3 are present in peptides with a Val (V) at P4 but not with an Ile (I) at that position. The only backbone component of the nonprime segment that does not form hydrogen bonds at all is the carbonyl group of P2, which points toward the solvent during the course of all simulations.

The identification of strong canonical backbone hydrogen bonds formed between P4-P2′ and the PfSUB1 binding site is in agreement with restricted fluctuations of the P5-P2′ residues, whereas the prime side residues P3′-P5′ undergo pronounced conformational changes (Figure 2 c)) in the case of most of the simulated PfSUB1-peptide complexes. Only PfSUB1-peptide complexes LVSAD↓NIDIS and KVKAQ↓DDFNP have restricted fluctuations of the P3′-P5′ residues, which can be explained by strong hydrogen bonds formed by these residues.
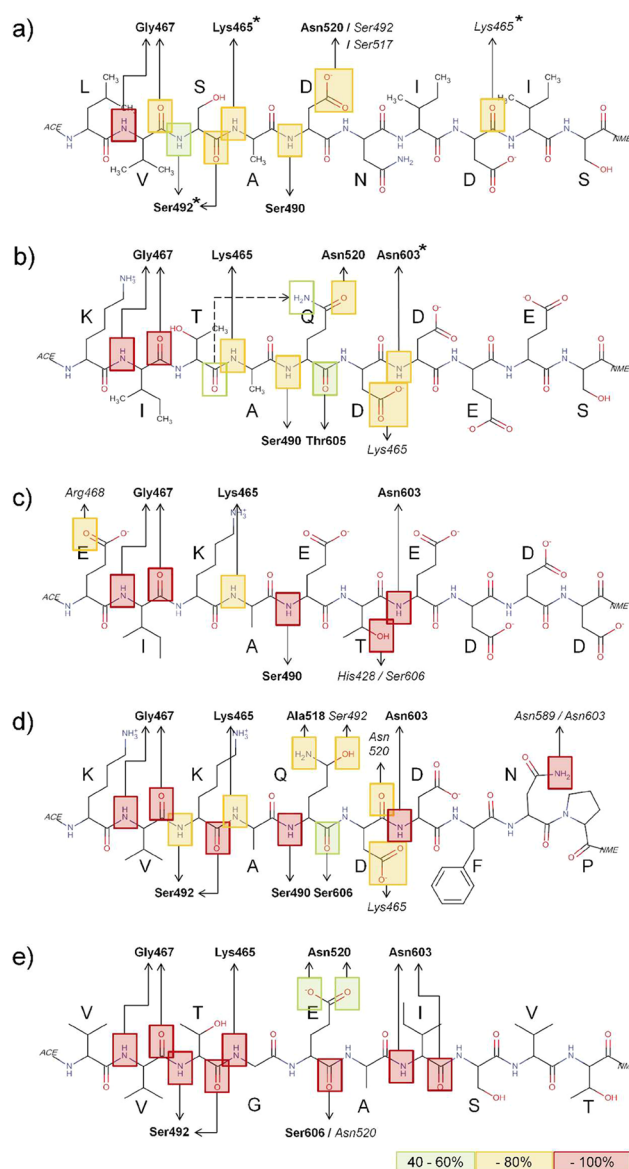

Figure 3. Scheme of hydrogen bonds formed between PfSUB1 and peptide substrates along the MD trajectories. The peptide orientation is shown from N- to C-terminus left-to-right and so is opposite to the orientation shown in (Figure 1). Very strong hydrogen bonds are boxed in red (occupancy of 80−100%) and orange (occupancy of 60−80%); strong hydrogen bonds are boxed in green (occupancy of 40−60%). Residues of the PfSUB1 binding site are labeled in bold or italics depending on whether the hydrogen bond is formed with a backbone group (bold) or a side-chain group (italics) of PfSUB1. In (b) an intramolecular hydrogen bond formed between the amine group of P1 and the carbonyl group of P3 is indicated as dashed line. In the case of residues which form more than one hydrogen bond, the coloring is based on the stronger hydrogen bond. In the case of the ammonium functionality of Lys465, the occupancy values of formed hydrogen bonds are added up and the box accordingly colored. The identification of strong hydrogen bonds based on occupancy values is supported by the analysis of the per-residue decomposition of relative free energies of the PfSUB1 binding site. Considering the backbone and side-chain part of the per-residue contributions separately revealed that for almost all of the identified hydrogen bonds the favorable electrostatic interaction as calculated by the molecular mechanics force field (ΔG_{MM-ELE}) outbalances the unfavorable electrostatic contribution due to desolvation (ΔG_{GB}) (exceptions are marked via *), respectively.
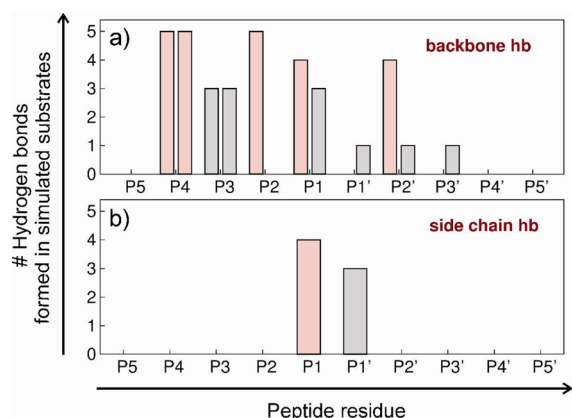
Figure 4. Summary scheme of hydrogen bonds formed between PfSUB1 and peptide substrates along the MD trajectories. The number of substrate peptides which form strong hydrogen bonds (occupancy > 40% as shown in Figure 3) are depicted as bar plots in (a) for hydrogen bonds formed with the peptide backbone and in (b) for hydrogen bonds formed with the peptide side-chains. In (a) hydrogen bonds formed with the backbone carbonyl group (left bar plot) are distinguished from those formed with the backbone amino group (right bar plot). Residues which form hydrogen bonds in at least four investigated substrates are depicted in red.

Hydrogen Bonds Formed via Peptide Side-Chains Include P1 and P1'. In addition, stabilizing hydrogen bond interactions are formed by the peptide side-chains, but these interactions are very dependent on the peptide sequence. The only two peptide side-chains that form stable hydrogen bonds with the PfSUB1 binding site are residue positions P1 and P1' (Figure 4).

Hydrogen bonds formed via the P1 side-chains include most commonly the backbone of Asn520 (Figure 3 a, b, and e) but also occasionally the backbone of Ala518 (Figure 3 d) and the side-chains of Ser492 and Ser517 (Figure 3 a and d). Hence, it is of note that the polar P1 side-chain forms strong side-chain hydrogen bonds only in the case of the LVSAD↓NIDIS and KVKAQ↓DDFNP peptides (Figure 3 a and d). This was initially considered surprising given the presence of three spatially adjacent Ser residues in the polar S1 pocket (Ser517, Ser519, and Ser492). We conclude that peptide binding is preferentially mediated by stabilizing hydrogen bonds between the P1 side chain and the backbone group of Asn520 and, in the case of the KITAQ↓DDEES peptide, another internal hydrogen bond with the carbonyl backbone group of P3. An interesting finding in this context is that the only investigated peptide with an Asp (D) at P1 does form stabilizing hydrogen bonds with the hydroxyl groups of Ser492 and Ser517.

In contrast to the P1 residue, the P1' residues form in general strong polar side-chain interactions (Figure 3 b, c, and d). In the case of the KITAQ↓DDEES (Figure 3 b) and KVKAQ↓DDFNP (Figure 3 d) peptides, these interactions are salt bridges formed between an Asp (D) and Lys465 in the S2 pocket.

Identification of Hot Spots by MM-GBSA Free Energy Decomposition. We next analyzed the energetic contributions of individual amino acids to PfSUB1-peptide complex formation to search for the dominant factors that dictate binding specificity of peptide substrates. For this, calculated binding free energies were decomposed on a per-residue level using the MM-GBSA approach in Amber 11[15] and the single trajectory method.[25] Calculations of this type not only

rationalize molecular recognition processes but also can guide the identification of small inhibitor molecules that mimic the determinants of binding of protein–protein complexes (hot spots).[27] Similarly, we expected that hot spots of PfSUB1-peptide complexes are equally important for small molecule binding to PfSUB1.[28]

Per-Residue Contributions of Peptides to Binding Free Energies Identifies P1 and P4 To Be the Main "Hot Spot" Residues. The decompositions of relative free energies revealed that among all five investigated peptide substrates, the most favorable interactions are formed by peptide side-chain position P4 ($\varnothing\Delta G_{Eff} = -8.2 \pm 0.5$ kcal/mol) and P1 ($\varnothing\Delta G_{Eff} = -6.4 \pm 1.4$ kcal/mol) (Table S2 in the Supporting Information).

In the case of P4, this is consistent with the importance of P4 residues in substrate recognition by subtilases in general[29] and agrees with an experimental mutation study on the LVSAD↓NIDIS peptide, where substitution of P4 Val (V) to Ala (A) had the greatest effect on substrate cleavage efficiency.[4] The large contribution of P4 to the binding free energy may be attributed to the van der Waals contribution and the nonpolar part of solvation free energy (Tables S3–S7 in the Supporting Information), which is not surprising given the highly hydrophobic nature of the S4 pocket.

In the case of P1, our earlier study (Koussis et al.) indicated that some flexibility is tolerated at P1 but that Asp (D), Ser (S), and Ala (A) are cleaved with the best efficiency.[4] Given the tolerance of P1 to accommodate different polar and small hydrophobic side-chains, the finding that P1 has the second greatest effect on the binding affinities is an unexpected outcome of this study. The large contribution of the P1 residue to complex stability does not arise from electrostatic interactions as one might have expected given the finding of strong hydrogen bonds formed by the P1 side-chain (Figure 4). An explanation is that in the case of P1, the favorable electrostatic interaction as calculated by the molecular mechanics force field ($\Delta G_{MM\text{-}ELE}$) is canceled by the unfavorable electrostatic contribution due to desolvation (ΔG_{GB}). In contrast, the van der Waals contribution ($\Delta G_{MM\text{-}VDW}$) and the nonpolar part of solvation free energy (ΔG_{GBSUR}) contributes favorably to binding (Tables S3–S7 in the Supporting Information). The decomposition analysis suggests, furthermore, that having a large side-chain at P1, such as in Gln (Q) and Glu (E), increases the binding affinity toward PfSUB1 as compared to the smaller Asp (D). This is in line with the substrate preference of known substrates (Figure 1).

Considering only those residues whose contributions to the effective energy ΔG_{Eff} are ≤ -4 kcal/mol (Figure 5), it becomes apparent that the distribution of further hot spot residues (as indicated in this study) varies depending on the investigated peptide sequence. P2 and P1' have a key contribution to binding in three of the five investigated peptide substrates, while P5, P3, and P2' have a key contribution to binding in two of the five peptide substrates. Based on the decomposition results, it seems to be beneficial to have an Ala (A) at P2 over a Gly (G). At P1' an uncharged residue (e.g., A or T) seems to be favorable over a charged Glu (D) (Figure 5; see Discussion and Conclusion for further discussion).

Effect of the Internal Dielectric Constant on the Predicted Binding Free Energies. MM-GBSA predictions are quite sensitive to the choice of the internal dielectric constant e_{in}, and therefore this parameter should be carefully chosen
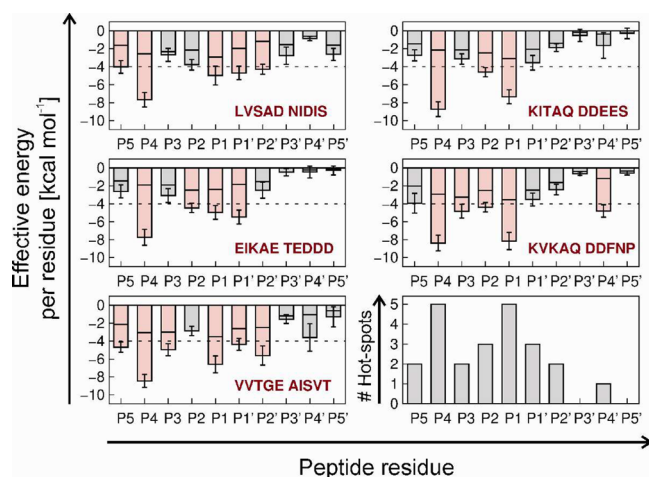
Figure 5. Per-residue contribution to the binding effective energy of PfSUB1-peptide complexes are depicted as bar plots. Per-residue contributions were calculated by the MM-GBSA decomposition method. Residues whose contributions to the effective energy $\Delta G_{Eff} \leq -4$ kcal/mol are depicted in red. The backbone and side-chain contributions to the effective free energy are indicated by partitioning the bar plots. The areas at the top of the bar plots correspond to the backbone and the areas at the bottom correspond to the side-chain contributions, respectively. The per-residue contributions were calculated by applying the MM-GBSA decomposition approach to MD trajectories of PfSUB1 in complex with five substrate peptides (red labels). The distribution of hot spot residues ($\Delta G_{Eff} \leq -4$ kcal/mol) along the investigated substrates is shown in the lower right panel. Repeating the analysis for the first and second half of the trajectory resulted in the same pattern of per-residue contributions.

according to the binding site characteristics.[26] For hydrophobic binding interfaces, a low internal dielectric constant ($e_{in} = 1$) can be recommended, while for highly charged binding interfaces, where a strong change in polarization occurs upon binding, higher values ($e_{in} = 4$) should be used.[26] Following findings from Hou et al.[26] for moderately charged binding interfaces (2−3 charged residues), as present in the PfSUB1 binding cleft (Arg600 (P3′), Lys465 (P1′), Arg468 (S5)), e_{in} was set to 2. However, we have also tested the effect of e_{in} on the predicted binding free energies by setting different e_{in} constants (1, 2, 3, and 4).

These analyses revealed that increasing e_{in} results in lower effective free energies but that the trend of calculated binding free energies remained the same for the investigated PfSUB1-complex structures (Table S1 in the Supporting Information). Similarly, the effect of increasing e_{in} on the per-residue contribution to the effective binding energy of PfSUB1-peptide complexes revealed that the contribution of hydrophobic residues barely changed (the contribution of 94% of the hydrophobic residues changed only with a standard deviation of < 0.5 kcal/mol). However, the contribution of charged residues increased significantly (the contribution of 55% of the charged residues increased with a standard deviation of > 1.0 kcal/mol). The reason for this was that the shielding effect of larger e_{in} values on the favorable electrostatic interaction as calculated by the molecular mechanics force field (ΔG_{MM-ELE}) was lower than the effect on the unfavorable electrostatic contribution due to desolvation (ΔG_{GB}) (Tables S3−S7 in the Supporting Information). In the case of the analyzed PfSUB1-complex structures, this resulted in a relative larger energy contribution of the charged P1′ residues in the case of the $e_{in} = 3$ and $e_{in} = 4$

analysis, which can be explained by strong polar interactions formed with the nearby Lys465 residue and the effect described above. Encouragingly, in all settings, the most favorable interactions are formed by peptide residues P4 and P1 which demonstrates the robustness of the main results.

Sequence Variations at the Prime Side. Motivated by the finding that the prime side residues might be important for optimal substrate recognition,[4,7] several sequence variations at the prime side of the KITAQ↓DDEES peptide were simulated by changing individual residues to Ala (A) as well as removing the P2′-P5′ prime part (Figures S1 and S2 in the Supporting Information). The results obtained seem to contrast with the suggested role of acidic amino acids at the prime side for binding specificity: mutating P1′ to Ala did not change the energy contribution of P1′. In contrast, removing the P2′-P5′ residues reduced the contribution of P1′ (Figure S2). The latter is in agreement with the high fluctuations of P1′ in the KITAQ↓D peptide (Figure S1). Overall, these results point to the importance of the backbone-backbone hydrogen bonds formed by P2′ (Figure 3 and 4) to ensure the stabilizing role of P1′.

Per-Residue Contributions of the PfSUB1 Binding Site to Binding Free Energies Identifies Four Cluster of Hot Spot Residues. The MM-GBSA calculations identify, in all peptide-PfSUB1 complex structures, four subsites whose energy patterns make a major contribution to the binding free energy (Figure 6). These "hot spot" regions include, in at least four of the five investigated PfSUB1-peptide structures, PfSUB1 residues Phe491, Leu469, and/or Phe493 (S4 pocket), Lys465 and Leu466 (lower rim of the PfSUB1 binding site cleft; forming S1′ and the lower part of the S2−S4 pockets), Ser492 and/or Asn520 (S1 pocket), and Asn603 (S2′ pocket). This is in line with the hydrogen bond analysis and energy decompositions of peptide structures above. The hydrogen bond analysis identified Lys465 and Asn603 as forming strong hydrogen bonds with the backbone groups of P2 and P2′, respectively, and Asn520 as forming strong hydrogen bonds with the side-chain group of P1. The energy decompositions of peptide structures identified, furthermore, that peptide residues P4 and P1 dominate binding.

■ DISCUSSION AND CONCLUSIONS

In order to enhance our understanding of the binding modes and the origin of binding determinants of PfSUB1 substrates, molecular dynamics (MD) simulations were performed and analyzed in combination with free energy calculations. The motivation of this study was to address the following questions: Which residues of the PfSUB1 binding site are essential for binding affinity and should, therefore, be exploited in compound design efforts such as docking experiments? What are the key interactions with known substrates that can be used to guide the identification and optimization of small molecule inhibitors?

Which Residues Are Essential for Binding Affinity? The analysis of hydrogen bonds formed along the MD trajectories indicated that, overall, a large number of canonical backbone hydrogen bonds are formed between peptide residues P4-P2′ and the PfSUB1 binding site cleft. This is consistent with the outcome of the free energy decompositions, which revealed that, in general, peptide residues P4 and P2-P1′ have the largest contribution to the effective free energy. This finding stresses the requirement of peptide substrates to interact with both prime and nonprime side residues of the PfSUB1 binding
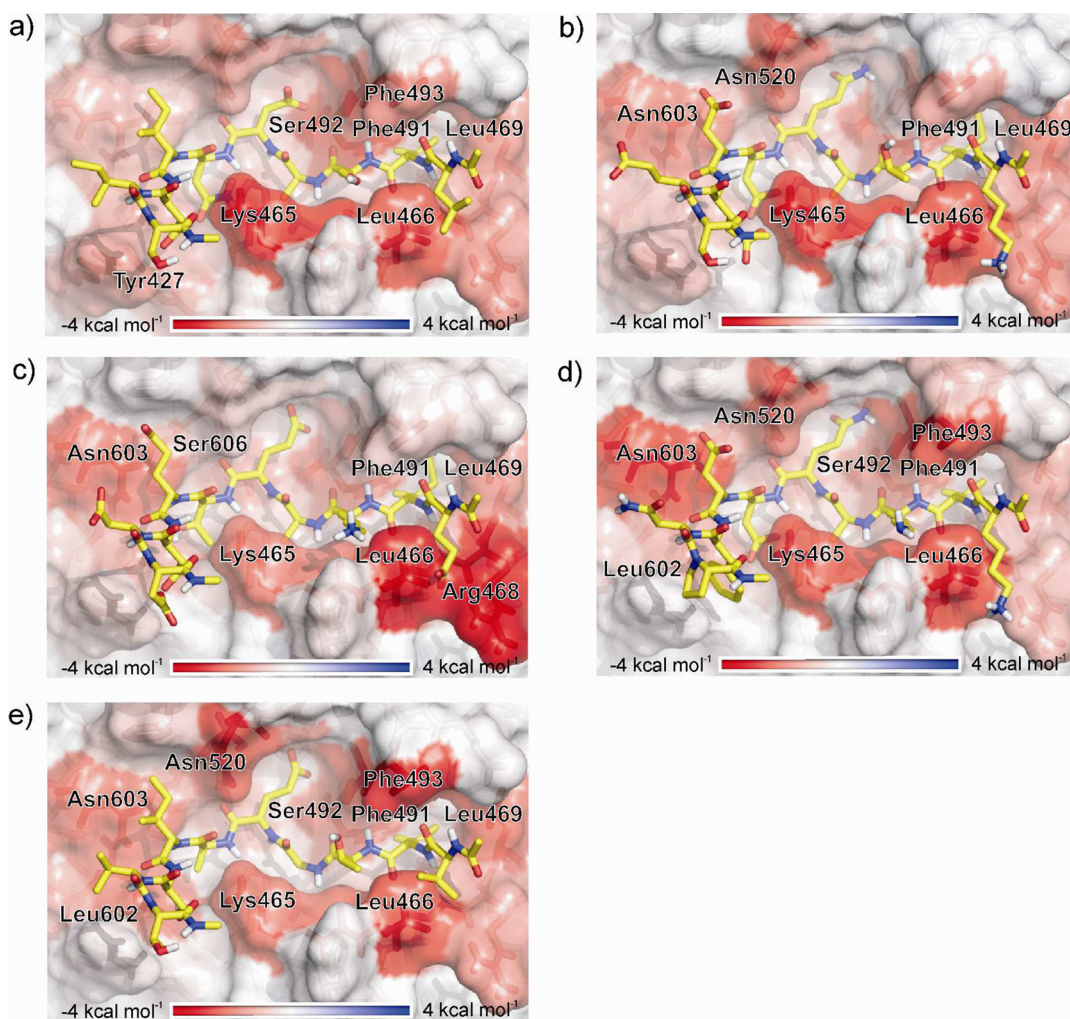
Figure 6. Per-residue contribution to the binding effective energy of the PfSUB1 binding site bound to different peptide substrates. The per-residue contributions were calculated by applying the MM-GBSA decomposition approach to MD trajectories of PfSUB1 in complex with (a) LVSAD↓NIDIS, (b) KITAQ↓DDEES, (c) EIKAE↓TEDDD, (d) KVKAQ↓DDFNP, and (e) VVTGE↓AISVT. The per-residue contributions are mapped onto the starting structures of the simulations using a color code with a linear scale. Residues whose contributions to the effective free energy ΔG_{Eff} ≤ −2 kcal/mol (Table S9 in the Supporting Information) are labeled.

site as indicated before.[4,7] Among the peptide residues, the most favorable interactions are formed by residues P4 and P1. Encouragingly, this result was obtained independent of the choice of the internal dielectric constant. The two peptide residues are, thus, potentially the main hot spot residues in the PfSUB1 substrates. The free energy decomposition revealed, furthermore, that van der Waals interactions and the nonpolar part of solvation free energy dictate the binding strength of the peptides, whereas the binding specificity is determined by electrostatic interactions and the polar part of solvation free energy (Table 2; Table S8 in the Supporting Information). Similar findings with respect to the binding determinants were found for other complex structures.[30]

The present finding of P4 and P1 to be the most important residues for the interaction with PfSUB1 agrees partly with a former study on a PcFK1-PfSUB1 complex model.[31] Based on the number of interacting residues and a free energy decomposition also using the MM-GBSA approach, the residues P2 and P1 were found to be the most important residues for the interaction between PfSUB1 and the YVPAQ↓NPCCR loop region in PcFK1.[31] Several reasons could explain the differences to the current study, including the fact that

different peptide sequences were investigated and that the latter analysis was based on a much shorter simulation trajectory.

Which Residues Are Essential for Binding Specificity and Selectivity? For most members of the S8A subfamily of subtilisins, of which PfSUB1 is a member, enzyme specificity is primarily dictated by interactions of the P1-P4 residue side-chains with the corresponding S1-S4 pockets.[29] The most obvious difference between PfSUB1 and other known members of the S8A subfamily is the small S2 pocket,[7] as it is delimited by a side-chain in PfSUB1 (Lys465) but is a Gly in most other S8A subfamily members. This explains not only the restriction to small residues at the P2 position but also might be exploited in the design of selective peptide inhibitors; larger S2 pockets will lose favorable van der Waals and other interactions formed in the small S2 pocket in PfSUB1.[32] In line with this argument, P2 makes a major contribution to the effective binding energy (ΔG_{Eff} ≤ −4 kcal/mol in three of the five substrates).

The analysis of hydrogen bonds revealed that in the case of the KITAQD↓DEESS and KVKAQ↓DDFNP peptides strong salt bridges are formed between the side-chain keto group of the P1′ residue and the side-chain ε-amino group of Lys465 in the S2 pocket. The impact of the salt bridges formed with

Lys465 in these two peptides is supported by individual energy contributions to the binding free energy: Favorable electrostatic interactions, as calculated by the molecular mechanics force field (ΔG_{MM-ELE}), outweigh the unfavorable electrostatic contribution due to desolvation (ΔG_{GB}) (Figure 3). This was true independent of the choice of the internal dielectric constant. To ensure stabilizing and selective interactions with the Lys465 residue, it seems, thus, to be beneficial to have an Asp (D) at the P1′ position of substrate based inhibitors.

Another region that might be exploited in the design of selective inhibitors is the S1 pocket, which is characterized in PfSUB1 by a cluster of Ser residues, i.e. Ser492, Ser517, and Ser519, which correspond to Gly127, Ala152, and Gly154 in bacterial subtilisin BPN′ structures (PDB code 1SBT). Our analysis revealed that in three of the five investigated peptides no stabilizing hydrogen bonds are formed between the peptide substrates and the side-chains of these residues. This indicates that the S1 pocket might not be a good subpocket to achieve binding selectivity. The only exception seems to be the aspartic acid (D) in the LVSAD↓NIDIS complex structure, which forms stabilizing hydrogen bonds with two of these serine residues.

Which Residues of the PfSUB1 Substrate-Binding Groove Should Be Exploited in Compound Design Efforts? The identification of binding "hot spots" in receptor structures can provide a knowledge-driven design of small inhibitor molecules by guiding the selection of the binding site used for docking studies[27] or by prioritizing points in pharmacophoric queries.[33] The identified hot spots residues of the PfSUB1 binding site not only are dispersed over the entire binding but also form four clusters of hot spot residues: S4 pocket, lower rim of the PfSUB1 binding site cleft, S1 pocket, and S2′ pocket. The distribution of residues, which make the major contribution to the binding free energy, suggests that the subpockets S4-S2 might be the most promising binding sites, e.g., for docking studies. However, this would disregard the possibility to use the side-chain keto group of P1′ to achieve selectivity versus other serine proteases. Thus, targeting the S1-S2′ binding subpockets should also be exploited in compound design efforts of small molecules. Overall, the spread of hot spot residues in the PfSUB1 binding site indicate that, although PfSUB1 has some potential as a drug target, it is also a difficult target.

Implication for Structure−Activity Experiments. Peptides that are most efficiently recognized and cleaved can be linked to inhibitory groups to generate specific inhibitors.[34] In line with this strategy, we published recently the capacity of N-acetylated peptidyl derivates of PfSUB1 substrates to inhibit SUB1 from the human malaria pathogens *P. falciparum*, *P. vivax*, and *P. knowlesi*.[7] In agreement with the free energy calculation presented in this study, a terminal carboxylic acid extension in the peptidyl α-ketoamide KS466 (designed to mimic P1′ interactions with PfSUB1) increased the inhibitory activity by about 2−6-fold.[7] The analysis of hydrogen bonds indicated that the backbone-backbone hydrogen bonds formed by P2′ might be necessary to ensure the stabilizing role of P1′. By combining an α-ketoamide functionality with the peptide, a reversible covalent bond with the hydroxyl side-chain of the catalytic serine of PfSUB1 is formed.[7] In such a case, it can be expected that the P1′ moiety in a hexapeptide (P5-P1′) forms the same stabilizing interactions as observed for the simulated deca-peptide substrates (P5-P5′). The results obtained from the MD simulations suggest, furthermore, that the P5 residue might not be necessary to achieve strong binding. Corresponding

structure activity experiments are currently under way for the α-ketoamide peptide structure.

What Are Key Interactions (Binding Determinants) of Known Substrates Which Can Be Used To Guide the Identification and Optimization of Small Molecule Inhibitors? Previously validated substrates have provided the first insights into the substrate preference of PfSUB1 (Figure 1) revealing, e.g., that P1 has a preference for amide containing and acidic amino acids. The decomposition of effective energy in this study has extended our knowledge of the binding determinants by revealing I) that P1 is a potential hot spot residue and II) that the large contribution of P1 to the binding free energy is driven by van der Waals interactions and the nonpolar part of solvation energy; i.e. in the case of the P1 residue, favorable electrostatic interactions formed by hydrogen bonds are canceled by unfavorable electrostatic contribution to desolvation (independent of the choice of the internal dielectric constant). Thus, having a small hydrophobic moiety in S1 or changing the amide group to a less polar group might be favorable for binding.

Role of Prime Side Residues. The analysis of hydrogen bonds and free energy decomposition identified P1′ as playing a key role for peptide stabilization. However, the tendency of PfSUB1 substrates to contain hydroxyl-containing and acidic residues at the prime side positions P1′-P5′[4,5,11] suggests that further prime side residues might form stabilizing interactions. One reason that, e.g., potential hydrogen bonds formed by P2′-P5′ are not found to be stable along the present MD trajectories might be that, in the native PfSUB1-substrate complex, the investigated (inhibitor) peptides can be expected to be part of a surface loop of a globular protein (such as in homologous complex structures, e.g., PDB code 1MME). The protein−protein complex forces the investigated peptide loop region into a bent shape protruding out of the active site, whereas further stable hydrogen bonds might be formed by the prime side residues P3′-P5′ (e.g., P3′ with Lys465 or Tyr427; suggested based on the starting model structure). In the course of most of the PfSUB1-peptide trajectories the P3′-P5′ residues undergo enhanced fluctuations (Figure 2c), thereby binding only occasionally for a short period of time to different regions of the PfSUB1 structure (either protruding out of the active site or extending along the S′ pockets). This explains partly why strong hydrogen bonds as well as hot spot residues are identified to comprise peptide residues P4-P2′ but not P3′-P5′. The goal of the present study was, however, not to investigate protein−protein complex structures but to get insights into the binding determinants of PfSUB1 that can be transferred to small molecule binding. For this objective, simulating PfSUB1-peptide complex structures was more insightful in our opinion.

Summing up. The present study enhanced our understanding of the binding determinants of peptide substrates to PfSUB1. The results are currently being used to guide structure−activity experiments of the α-ketoamide peptide structure described in ref 7. It is encouraging that, although the present study is based on a homology model of PfSUB1, the results obtained are in line with the experimental data obtained so far. This indicates that, in the absence of a crystal structure, the use of homology models in combination with MD and MM-GBSA free energy calculations can provide critical insights into the origin of binding determinants. We expect that the insights obtained in this study will facilitate compound design efforts against a promising antimalarial drug target.

The present study also underlines that using MM-GBSA calculations in order to estimate absolute binding free energies remains difficult. However, free energy calculations can still be useful in practical applications even with a considerable level of inaccuracy,[35] and a number of successful applications to various systems have been reported.[36,37] A more comprehensive discussion about potential assets and drawbacks of the MM-GB/PBSA approach in comparison with other free energy methods can be found elsewhere.[36,38,39] The success of the MM-GBSA approach is quite sensitive to simulation protocols, such as the sampling strategy to generate snapshots, the simulation length, the choice of the internal dielectric constant, the way the entropic contribution is included, etc.[26,36,40] As no experimental binding affinity values are currently available for PfSUB1 substrates, it was not possible to experimentally justify the used simulation protocol, and, thus, the parameters used in this study were guided by the literature (e.g., the internal dielectric constant e_{in} was set to 2 as recommended for moderately charged binding interfaces as present in PfSUB1).[22,26]

To establish nevertheless the effect of e_{in} on the predicted binding free energies, different e_{in} constants (1, 2, 3, and 4) were tested. In agreement with a previous report,[41] these tests showed that the absolute value of the electrostatic binding free energy varies inversely with the value of the dielectric constant but that the main results are not affected (i.e., the most favorable interactions were formed in all settings by peptide residues P4 and P1). However, by increasing e_{in}, the contribution of charged residues increased significantly and resulted, e.g., into a relative larger contribution of the P1′ residues in the case of the KITAQ↓DDEES and KVKAQ↓ DDFNP peptides (Figure S3 in the Supporting Information). This shows that an adequate choice of the internal dielectric constant for a specific binding pocket might be especially critical if one aims to understand the determinants of binding of charged residues.

ASSOCIATED CONTENT

ⓢ Supporting Information

Tables with binding free energies of PfSUB1-peptide complex structures (Table S1), per-residue contribution for residues of substrate peptides bound to the PfSUB1 binding site (Table S2), individual energy contributions for binding free energy of the simulated substrate peptides (Tables S3–S8), and per-residue contribution for residues of the PfSUB1 binding site bound to different peptide substrates (Table S9) as well as graphical representations of results obtained for PfSUB1-peptide complex structures with sequence variations at the prime side of the KITAQ↓DDEES peptide (Figures S1 and S2) and per-residue contribution of peptide residues calculated by using different internal dielectric constants (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: simone.fulle@inhibox.com

Notes

The authors declare no competing financial interest.

ABBREVIATIONS:

PfSUB1: subtilisin-like serine protease present in *P. falciparum*; RBC: human red blood cells; MD: molecular dynamics; RMSD: root-mean square deviation; RMSF: root-mean square fluctuations; MM-PBSA: Molecular Mechanics Poisson–Boltzmann surface area; MM-GBSA: Molecular Mechanics Generalized Born surface area

REFERENCES

(1) Price, R. N.; Douglas, N. M.; Anstey, N. M. New developments in *Plasmodium vivax* malaria: Severe disease and the rise of chloroquine resistance. *Curr. Opin. Infect. Dis.* **2009**, *22*, 430−435.

(2) William, T.; Menon, J.; Rajahram, G.; Chan, L.; Ma, G.; Donaldson, S.; Khoo, S.; Frederick, C.; Jelip, J.; Anstey, N. M.; Yeo, T. W. Severe *Plasmodium knowlesi* malaria in a tertiary care hospital, Sabah, Malaysia. *Emerging Infect. Dis.* **2011**, *17*, 1248−1255.

(3) Blackman, M. Malarial proteases and host cell egress: An 'emerging' cascade. *Cell. Microbiol.* **2008**, *10*, 1925−1934.

(4) Koussis, K.; Withers-Martinez, C.; Yeoh, S.; Child, M.; Hackett, F.; Knuepfer, E.; Juliano, L.; Woehlbier, U.; Bujard, H.; Blackman, M. A multifunctional serine protease primes the malaria parasite for red blood cell invasion. *EMBO J.* **2009**, *28*, 725−735.

(5) Yeoh, S.; O'Donnell, R.; Koussis, K.; Dluzewski, A.; Ansell, K.; Osborne, S.; Hackett, F.; Withers-Martinez, C.; Mitchell, G.; Bannister, L.; Bryans, J.; Kettleborough, C.; Blackman, M. Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes. *Cell* **2007**, *131*, 1072−1083.

(6) Arastu-Kapur, S.; Ponder, E. L.; Fonovic, U. P.; Yeoh, S.; Yuan, F.; Fonovic, M.; Grainger, M.; Phillips, C. I.; Powers, J. C.; Bogyo, M. Identification of proteases that regulate erythrocyte rupture by the malaria parasite *Plasmodium falciparum*. *Nat. Chem. Biol.* **2008**, *4*, 203−213.

(7) Withers-Martinez, C.; Suarez, C.; Fulle, S.; Kher, S.; Penzo, M.; Ebejer, J. P.; Koussis, K.; Hackett, F.; Jirgensons, A.; Finn, P.; Blackman, M. J. *Plasmodium* subtilisin-like protease 1 (SUB1): Insights into the active-site structure, specificity and function of a pan-malaria drug target. *Int. J. Parasitol.* **2012**, *42*, 597−612.

(8) Blackman, M. J.; Corrie, J. E.; Croney, J. C.; Kelly, G.; Eccleston, J. F.; Jameson, D. M. Structural and biochemical characterization of a fluorogenic rhodamine-labeled malarial protease substrate. *Biochemistry* **2002**, *41*, 12244−12252.

(9) Moneriz, C.; Mestres, J.; Bautista, J. M.; Diez, A.; Puyet, A. Multi-targeted activity of maslinic acid as an antimalarial natural compound. *FEBS J.* **2011**, *278*, 2951−2961.

(10) Gemma, S.; Giovani, S.; Brindisi, M.; Tripaldi, P.; Brogi, S.; Savini, L.; Fiorini, I.; Novellino, E.; Butini, S.; Campiani, G.; Penzo, M.; Blackman, M. J. Quinolylhydrazones as novel inhibitors of *Plasmodium falciparum* serine protease PfSUB1. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 5317−5321.

(11) Silmon de Monerri, N.; Flynn, H.; Campos, M.; Hackett, F.; Koussis, K.; Withers-Martinez, C.; Skehel, J.; Blackman, M. Global identification of multiple substrates for *Plasmodium falciparum* SUB1, an essential malarial processing protease. *Infect. Immun.* **2011**, *79*, 1086−1997.

(12) Withers-Martinez, C.; Saldanha, J. W.; Ely, B.; Hackett, F.; O'Connor, T.; Blackman, M. J. Expression of recombinant *Plasmodium falciparum* subtilisin-like protease-1 in insect cells: characterization, comparison with the parasite protease, and homology modelling. *J. Biol. Chem.* **2002**, *277*, 29698–29709.

(13) Radisky, E. S.; Koshland, D. E., Jr. A clogged gutter mechanism for protease inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 10316–10321.

(14) Dauter, Z.; Betzel, C.; Genov, N.; Pipon, N.; Wilson, K. S. Complex between the subtilisin from a mesophilic bacterium and the leech inhibitor eglin-C. *Acta Crystallogr., B* **1991**, *47*, 707–730.

(15) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(16) Simmerling, C.; Strockbine, B.; Roitberg, A. E. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* **2002**, *124*, 11258–11259.

(17) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725.

(18) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(19) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(20) Cheatham, T. E.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. Molecular dynamics simulations on solvated biomolecular systems: The Particle Mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J. Am. Chem. Soc.* **1995**, *117*, 4193–4194.

(21) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(22) Gohlke, H.; Case, D. A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J. Comput. Chem.* **2004**, *25*, 238–250.

(23) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897.

(24) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.

(25) Gohlke, H.; Kiel, C.; Case, D. A. 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.* **2003**, *330*, 891–913.

(26) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69–82.

(27) Metz, A.; Pfleger, C.; Kopitz, H.; Pfeiffer-Marek, S.; Baringhaus, K. H.; Gohlke, H. Hot spots and transient pockets: Predicting the determinants of small-molecule binding to a protein-protein interface. *J. Chem. Inf. Model.* **2012**, *52*, 120–133.

(28) Zerbe, B. S.; Hall, D. R.; Vajda, S.; Whitty, A.; Kozakov, D. Relationship between hot spot residues and ligand binding hot spots in protein–protein interfaces. *J. Chem. Inf. Model.* **2012**, *52*, 2236–2244.

(29) Siezen, R. J.; Leunissen, J. A. Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Sci.* **1997**, *6*, 501–523.

(30) Delgado-Soler, L.; Pinto, M.; Tanaka-Gil, K.; Rubio-Martinez, J. Molecular determinants of bim(BH3) peptide binding to pro-survival proteins. *J. Chem. Inf. Model.* **2012**, *52*, 2107–2118.

(31) Bastianelli, G.; Bouillon, A.; Nguyen, C.; Crublet, E.; Petres, S.; Gorgette, O.; Le-Nguyen, D.; Barale, J. C.; Nilges, M. Computational reverse-engineering of a spider-venom derived peptide active against *Plasmodium falciparum* SUB1. *PLoS One* **2011**, *6*, e21812.

(32) Huggins, D. J.; Sherman, W.; Tidor, B. Rational approaches to improving selectivity in drug design. *J. Med. Chem.* **2012**, *55*, 1424–1444.

(33) Barillari, C.; Marcou, G.; Rognan, D. Hot-spots-guided receptor-based pharmacophores (HS-Pharm): A knowledge-based approach to identify ligand-anchoring atoms in protein cavities and prioritize structure-based pharmacophores. *J. Chem. Inf. Model.* **2008**, *48*, 1396–1410.

(34) Raju, R. M.; Goldberg, A. L.; Rubin, E. J. Bacterial proteolytic complexes as therapeutic targets. *Nat. Rev. Drug Discovery* **2012**, *11*, 777–789.

(35) Shirts, M. R.; Mobley, D. L.; Brown, S. P. Free-energy calculations in structure-based drug design. In *Drug design: Structure- and Ligand-based approaches*; Merz, K. M., Ringe, D., Reynolds, C. H., Eds.; Cambridge University Press: 2010.

(36) Homeyer, N.; Gohlke, H. Free energy calculations by the Molecular Mechanics Poisson–Boltzmann Surface Area method. *Mol. Inf.* **2012**, *31*, 114–122.

(37) Fulle, S.; Gohlke, H. Molecular recognition of RNA: Challenges for modeling interactions and plasticity. *J. Mol. Recognit.* **2010**, *23*, 220–231.

(38) Steinbrecher, T.; Labahn, A. Towards accurate free energy calculations in ligand protein-binding studies. *Curr. Med. Chem.* **2009**, *17*, 767–785.

(39) Singh, N.; Warshel, A. Absolute binding free energy calculations: On the accuracy of computational scoring of protein-ligand interactions. *Proteins* **2010**, *78*, 1705–1723.

(40) Genheden, S.; Ryde, U. How to obtain statistically converged MM/GBSA results. *J. Comput. Chem.* **2010**, *31*, 837–846.

(41) Basdevant, N.; Weinstein, H.; Ceruso, M. Thermodynamic basis for promiscuity and selectivity in protein-protein interactions: PDZ domains, a case study. *J. Am. Chem. Soc.* **2006**, *128*, 12766–12777.

(42) Withers-Martinez, C.; Saldanha, J. W.; Ely, B.; Hackett, F.; O'Connor, T.; Blackman, M. J. Expression of recombinant *Plasmodium falciparum* subtilisin-like protease-1 in insect cells: Characterization, comparison with the parasite protease, and homology modelling. *J. Biol. Chem.* **2002**, *277*, 29698–29709.