Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry 16 (2008) 5254–5265
Bioorganic & Medicinal Chemistry

Structural insights into the *Plasmodium falciparum* histone deacetylase 1 (*Pf*HDAC-1): A novel target for the development of antimalarial therapy

Prasenjit Mukherjee,[a] Anupam Pradhan,[b] Falgun Shah,[a] Babu L. Tekwani[b,c] and Mitchell A. Avery[a,b,d,*]

[a]*Department of Medicinal Chemistry, School of Pharmacy, University of Mississippi, University, MS 38677, USA*
[b]*National Center for Natural Products Research, University of Mississippi, University, MS 38677, USA*
[c]*Department of Pharmacology, University of Mississippi, University, MS 38677, USA*
[d]*Department of Chemistry & Biochemistry, University of Mississippi, University, MS 38677, USA*

Received 14 January 2008; revised 27 February 2008; accepted 3 March 2008
Available online 6 March 2008

Abstract—The histone deacetylase (HDAC) enzyme from *Plasmodium falciparum* has been identified as a novel target for the development of antimalarial therapy. A ligand-refined homology model of *Pf*HDAC-1 was generated from the crystal structures of human HDAC8 and HDLP using a restraint guided optimization procedure involving the OPLS/GBSA potential setup. The model was extensively validated using protein structure checking tools. A predictive docking study was carried out using a set of known human HDAC inhibitors, which were shown to have in vitro antimalarial activity against the chloroquine sensitive D6 and resistant W2 strains of *P. falciparum*. Pose validation and score-based active/inactive separation studies provided independent validation of the geometric accuracy and the predictive ability of the generated model. Comparative analysis was carried out with the human HDACs to identify differences in the binding site topology and interacting residues, which might be utilized to develop selective *Pf*HDAC-1 inhibitors.

1. Introduction

Malaria,[1] a major parasitic disease prevalent throughout the tropical and the subtropical regions of the world is caused by four species belonging to the genus plasmodium including *Plasmodium falciparum* and *Plasmodium vivax*. Approximately 300 million new malaria infections are reported annually and result in nearly one million deaths with the majority of them occurring in children. Emergence of multi drug resistant (MDR) strains of the parasite has rendered the current therapies ineffective and is a growing cause of concern. With the lack of success in the area of vaccine development, identification and drug discovery against novel parasitic targets seems to be the most rational approach towards curbing the menace of MDR malaria.

The histones are nuclear core proteins, which function in the regulation of transcription and cell cycle progression by the acetylation and deacetylation of the ε-amino group of specific lysine residues. The processes of acetylation and deacetylation is controlled by two family of enzymes; histone acetyl transferases (HATs) and the histone deacetylases (HDACs), respectively.[2] A number of HDACs are present in humans and can be classified into three categories[3]; class I (HDAC 1,2,3,8), class II (HDAC 4,5,6,7,9,10) and class III (sirtuins). The HDACs from classes I and II demonstrate zinc dependent mechanisms. The sirtuins are different from the two other classes of HDACs, since they are NAD^+ dependent and demonstrate ADP-ribosyltransferase activity in addition to deacetylase activity. Three histones[4–6] from *P. falciparum*, H2A, H2B and H3 have been identified from cDNA clones and are shown to have close homology with the histones from other organisms. Two histone deacetylase (HDAC) homologues from *P. falciparum*, *Pf*HDAC-1[7] and *Pf*Sir2[8] have been cloned and sequenced so far. *Pf*HDAC-1

Keywords: *Plasmodium falciparum*; Histone deacetylase; Homology modeling; Docking; Molecular dynamics.
* Corresponding author. Tel.: +1 662 915 5879; fax: +1 662 915 5638; e-mail: mavery@olemiss.edu

0968-0896/$ - see front matter © 2008 Published by Elsevier Ltd.
doi:10.1016/j.bmc.2008.03.005

shows sequence homology to human, yeast and other eukaryotic HDACs and is expressed predominantly in the asexual blood stages of the parasite although transcripts have been identified from the sexual stages of the parasite as well. Efforts are currently underway to express and purify a functional form of the *Pf*HDAC-1 enzyme. *Pf*Sir2 is also present in the blood stages of the parasite's lifecycle and it shows sequence homology specifically to the archaeal and eubacterial sirtuins as opposed to other sirtuins sequenced so far. The enzyme's deacetylation activity is dependent upon its ADP-ribosylation and the enzyme is not inhibited by any of the known sirtuin inhibitors.

There are distinct differences in the pharmacophoric requirements and inhibitor design strategies against the zinc dependent and the NAD^+ dependent HDACs. The presence of a zinc binding motif (ZBM) is commonly associated with inhibitors designed against the HDACs from classes I and II. Inhibitors of human HDACs were primarily developed as therapies for cancer. Apicidin,[9,10] a cyclic tetra-peptide natural product based human HDAC inhibitor isolated from *Fusarium* spp. was the first compound to demonstrate in vitro effects against several apicomplexan parasites including the *Plasmodium* species. It was suggested that Apicidin inhibited the plasmodial HDAC leading to the inhibition of cell proliferation. However, Apicidin showed antiproliferative effects in HeLa (mammalian) cells at concentrations similar to what is required for exhibiting antimalarial activity, thereby ruling out its potential development as an antimalarial agent based on toxicity issues. A set of synthetic derivatives[11] of Apicidin derived by modifying the compounds indole moiety with N-alkylated quinolones showed potent antimalarial activity with selectivity over HeLa (mammalian) cells in terms of their anti proliferative effect. These compounds also demonstrated selective binding to *Eimeria tenella* HDAC over avian HDAC in cell nuclear extract-based assays. Another set of human HDAC inhibitors (Fig. 1) containing a hydroxamate ZBM including Trichostatin A (TSA),[12] Salicyl hydroxamic acid (SHA), Suberoic bis hydroxamic acid (SBHA) and Azelaic bis hydroxamic acid (ABHA) were also shown to exhibit antimalarial effects. SBHA[13] treatment was able to cure malarial infection in the murine model of the disease caused by *P. berghei*. ABHA also demonstrated a high selectivity in terms of its antiproliferative effect in cancer cells[14] over normal cells and opened up the possibilities for the development of plasmodial HDAC inhibitors without any major toxicity issues. These scientific observations provide strong evidence that the inhibition of the plasmodial HDACs could be targeted for the development of novel and selective antimalarial therapy.

Ligand bound structure of the target protein provides a suitable starting point for structure-based drug design. In the case of *Pf*HDAC-1, a crystal structure of the enzyme is yet to be elucidated and in its absence, a ligand-refined homology model[15–18] of *Pf*HDAC-1 complexed with a hydroxamate-based inhibitor TSA was generated. The model was suitably refined using a restraint guided optimization protocol applying the optimized potential for liquid simulations (OPLS) force field and generalized born surface atom (GBSA) implicit solvent treatment. The refined model was checked for structural integrity using various protein structure checking tools. The predictive ability of the model was evaluated through a docking study of a set of hydroxamate based human HDAC inhibitors, which demonstrate antimalarial activity against multiple strains of the *P. falciparum* parasite. Comparative analysis with human HDACs was carried out to identify differences in the binding site topology, which might be useful in the design of selective *Pf*HDAC-1 inhibitors. The refined model and the validated docking protocol could be utilized in structure-based drug design studies against *Pf*HDAC-1.

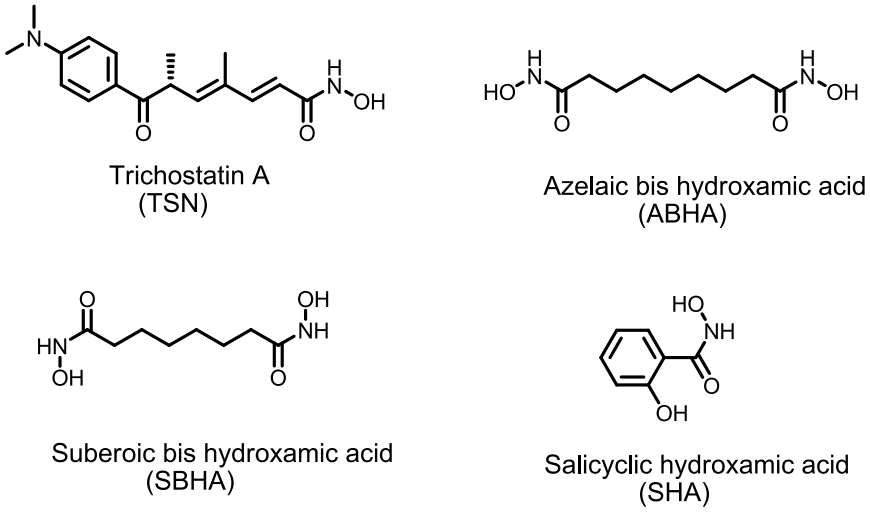

Figure 1. Chemical structures of some known human HDAC inhibitors which show antimalarial activity against the chloroquine sensitive (D6, Sierra Leone) and resistant (W2, Indo China) strains of *P. falciparum*.

2. Discussion

2.1. Model development

A basic local alignment search tool (BLAST)[19] search of the protein data bank (PDB) using the target sequence of *Pf*HDAC-1 showed that human HDAC8[20] and the HDAC homologue from the hyperthermophilic bacterium *Aquifex aeolicus* (HDLP)[21] were the closest structural homologues with ligand bound crystal structures. While developing a ligand-refined model it would be ideal to use structures of the two reference proteins with the same co-crystallized ligand. In this case, common ligand bound structures exist for Suberonylanilide bis hydroxamic acid (SAHA) (1T69-HDAC8, 1C3S-HDLP) and TSA (1T64-HDAC8, 1C3R-HDLP). The TSA bound structures of HDAC8 and HDLP were finally chosen based on two reasons. First, these structures had the highest resolution (1T64-1.90 Å, 1C3R-2.00 Å) and the least number of missing backbone and/or side chain coordinates amongst the available structures. Secondly, we had cell-based antimalarial activity data for TSA against the chloroquine sensitive (D6, Sierra Leone) and resistant (W2, Indo China) strains of *P. falciparum*, making it the ligand of choice.

The analysis of the sequence alignment (Supplementary information, Fig. S2) showed that the target sequence residues from the N-terminus to the end of the L1 (Fig. 2a) loop region had more homology to the HDLP sequence as compared to the human HDAC8 sequence. The overall homology of the target sequence was, however, greater for human HDAC8. It was therefore decided to use the HDLP structure for building the region extending from the N-terminus to the L1 loop and use the human HDAC8 structure for modeling the rest of the protein. The residues from the L1 loop (Fig. 2d) form a part of the binding pocket and affect the shape of the binding pocket as well as the binding conformation of the ligands. In the case of the ligand-refined *Pf*HDAC-1 model, the binding conformation of TSA's alkyl linker and terminal dimethyl anilino moiety which interacts with the L1 loop residues would match more closely with the HDLP-TSA structure. On the contrary, this region is much more flexible in the HDAC8–TSA structure and allows for the binding of two TSA molecules. Therefore, we utilized the coordinates of TSA and Zn^{2+} from the HDLP-TSA structure during the model building process.

Since molecular mechanics treats metals as bodies with a van der Waals radius and a point charge, it may not always be adequate in reproducing the coordination geometry around the metal center. A correct representation of the metal coordination was critical in this case since the ligand coordinates at two points with the metal center and the ligand's pose within the enzyme's binding site is critically dependent upon them. A restraint guided molecular mechanics refinement using the OPLS2001 force field and GBSA implicit solvent treatment with extended cutoffs was therefore carried out to refine the protein structure and preserve the metal coordination. The mean distances for the distance restraints used in the refinement were incorporated from what is seen in the reference crystal structure.

The refined model was subjected to a number of protein structural checks to gauge its quality. Matchmaker analysis[22] showed that the model had a negative temperature factor of −0.18 kT suggesting good geometric accuracy. Prostat analysis[23] was carried out and showed that there were no significant bond length deviations and only 0.5% of residues had a bond angle violation of >10°. The Ramachandran plot (Supplementary information, Fig. S1) for the model showed that only one non glycine residue was located in the disallowed region. The Cα heavy atom RMSD of the modeled structure was 1.21Å with human HDAC8 (1T64, PDB code) and 2.23Å with HDLP (1C3R, PDB code). All of the protein structure checks pointed towards a suitably refined model fit for use in docking studies.

2.2. Structural analysis

The ligand-refined model of *Pf*HDAC-1 (Fig. 2) comprises of a single domain, belonging to the open α/β class of folds, which bears close resemblance to the crystal structures of HDAC homologues from the other species. The structure comprises of a centrally located 8 strands β-sheet enclosed by 14 α helical regions (Fig. 2a–c). The secondary structures are connected by prominent loop regions and seven of these loops L1–L7 (Fig. 2a and d) form the walls of the enzyme's binding site. While the loops L1 and L2 form the upper region of the binding site the loops L3 to L7 forms the walls of the lower region, which also houses the catalytic machinery of the enzyme. The enzyme sequence comprises of a long C-terminal region, which extends beyond the catalytic domain. This region was not modeled for the lack of a reference structural homologue. This extended region is present in the human HDACs 1–3 and is absent in human HDAC8. It is known to be functionally important and helps in the recruitment of protein complexes, which facilitates in the localization and activity of the enzyme.[24] Additional evidence from homologous enzymes suggests that residues in this region undergo posttranslational modifications such as phosphorylation[25] and sumoylation,[26] which affect the enzyme's activity through allosteric modulation and regulation of its complex forming ability. A polyclonal antibody[7] raised against a peptide derived from the C-terminus residues 428–445 of *Pf*HDAC-1 helped in the detection of the enzyme in a fraction of the Triton-100 extraction of the malaria parasite, which was rich in cell nuclei.[7] This suggests that the enzyme is primarily localized in the nuclei and the C-terminal region is likely to play an important role in this localization process.

The second unique structural feature involves the length and composition of the N-terminal L1 loop region of *Pf*HDAC-1, which lines the upper region of the binding site and extends onto the protein surface. This loop region is seen to attain a wide variety of conformations

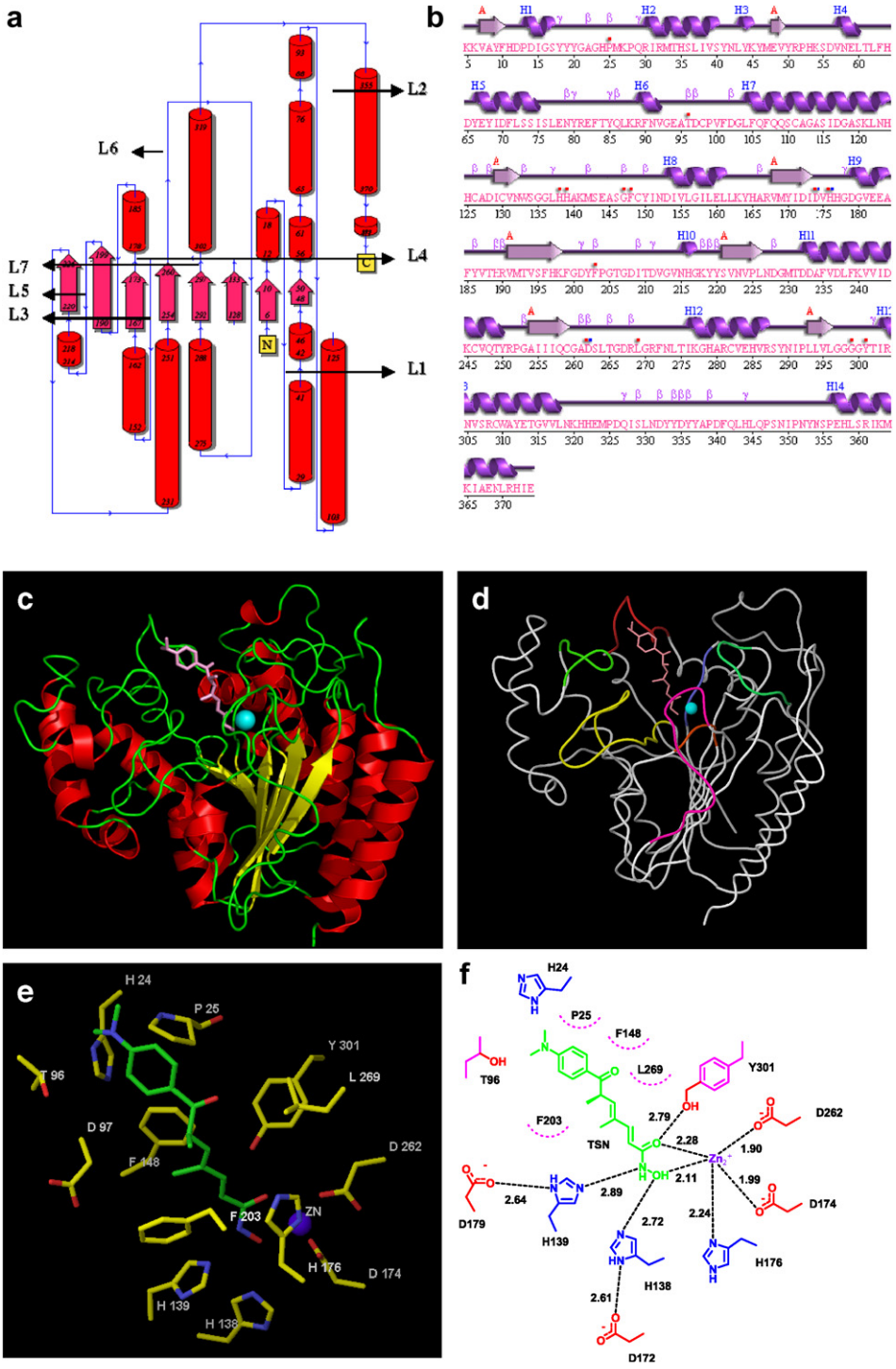

Figure 2. (a) Protein topology diagram of the ligand-refined *Pf*HDAC-1 model showing the α helical regions (red cylinder), β strand (magenta arrow) as well as the loop regions (blue line). The seven key loops (L1–L7) which form a part of the binding site are also labeled. (b) A ProMotif secondary structure wiring diagram for the same model showing the location of the secondary structure elements. The residues demarcated by red dots are in close vicinity of the ligand and those demarcated with blue dots are in close proximity of the metal ion. (c) A cartoon representation of the ligand-refined *Pf*HDAC-1 model showing the spatial arrangement of the α helices (red), β strands (yellow) and the loop regions (green). The binding site is clearly demarcated by the presence of the TSA ligand (pink tube representation) and the metal ion (cyan CPK representation). (d) A Cα tube representation of the same model showing the location of the TSA ligand (pink tube representation), metal ion (cyan CPK representation) as well as the seven loop regions L1 (red), L2 (bright green), L3 (yellow), L4 (brown), L5 (magenta), L6 (light green) and L7 (violet), which line the binding site. (e) A tube representation of the binding site of the ligand-refined *Pf*HDAC-1 model showing the location of TSA (green) and the metal ion (cyan CPK) as well as some key binding site residues (yellow). (f) A schematic diagram of the binding site showing the critical interactions undergone by TSA (green) and the metal ion (purple) as well as the distances for these interactions. The residues undergoing hydrophobic interactions are shown in magenta dotted semicircles while those undergoing polar interactions are shown in red (electro negative) and blue (electro positive). All the distances are between the heavy atoms, and are shown in angstroms.

in the available crystal structures of human HDAC8. In *Pf*HDAC-1, this loop region is located between helices H1 and H2 (Fig. 2b, Supplementary information, Fig. S2) and is longer by 2 residues than that of human HDAC8 and matching with that of HDLP. Sequence analysis showed that the length of the L1 loop region in HDLP matches with those of human HDACs 1–3. A shorter loop region allows the human HDAC8 to attain a much wider cavity as compared to *Pf*HDAC-1, wherein the cavity is smaller and is structurally similar to that of HDLP. A number of small differences were also observed in the other loop regions. The loop between the helices H5 and H6 (Fig. 2b, Supplementary information, Fig. S2) of *Pf*HDAC-1 appears to be longer by three residues as compared to both HDLP and human HDAC8, while the L2 loop region (Fig. 2a) was predicted to be one residue shorter than the corresponding loop region of HDLP. The loop region after the helix H7 appears to be two residues longer, compared to HDLP. On the contrary, the loop following helix H8 was predicted to be shorter by one residue, compared to HDLP. The L5 loop region (Fig. 2a) of *Pf*HDAC-1 located just before helix H10 (Fig. 2b, Supplementary information, Fig. S2) appears to be shorter by one and two residues compared to human HDAC8 and HDLP, respectively. Additionally, a C-terminal loop region of *Pf*HDAC-1 located between helices H13 and H14 (Fig. 2b, Supplementary information, Fig. S2) appears to be longer by two residues as compared to that of human HDAC8.

2.3. Binding site

The binding site of *Pf*HDAC-1 (Table 1) (Fig. 2e and f) comprises of a narrow channel, which leads to a deep cavity lined by the amino acid residues critical to the enzyme's catalytic machinery. During normal substrate processing, this narrow channel is occupied by the four methylene groups of the acetylated lysine. In the ligand-refined *Pf*HDAC-1 model, the aliphatic chain and the dimethyl anilino group of TSA are seen to occupy this region. The terminal hydroxamate group enters into the deep cavity and forms several important interactions with the catalytic residues. The upper region of the binding site is primarily hydrophobic and is lined by the residues His24, Pro25, Thr96, Phe148, Phe203, Leu269 and Tyr301. These residues (Table 1) are mostly conserved except Thr96, which in the case of human HDAC8 and HDLP is a tyrosine; Asp97, which in case of HDLP is a glutamate; and Leu269 which in the case of human HDAC8 is a methionine. These residues are seen to interact with the aliphatic chain and the terminal dimethyl anilino group of TSA.

Table 1. The table shows the key binding site residues of *Pf*HDAC-1 and the nature of their interaction with the ligand and/or the metal ion present in the binding site

	HDAC8	HDLP	*Pf*HDAC-1	Nature of interaction
1	—	P22	P25	Hydrophobic interaction-L1loop
2	Y100	Y91	**T96**	Polar interaction
3	D101	E92	**D97**	
4	H142	H131	H138	Catalytic residue
5	H143	H132	H139	Catalytic residue
6	G151	G140	G147	
7	F152	F141	F148	Hydrophobic interaction
8	D176	D166	D172	Catalytic residue
9	D178	D168	D174	Coordinates Zn^{2+}
10	H180	H170	H176	Coordinates Zn^{2+}
11	D183	D173	D179	Catalytic residue
12	F208	F198	F203	Hydrophobic interaction
13	D267	D258	D262	Coordinates Zn^{2+}
14	M274	L265	**L269**	Hydrophobic interaction
15	Y306	Y297	Y301	Polar/hydrophobic interaction

The corresponding residues from the reference homologues human HDAC8 and HDLP are also listed for comparison. The residues of *Pf*HDAC-1 which are not conserved with either of the reference homologues are shown in bold.

The HDAC enzymes are Zn^{2+} dependent and a single metal ion is present in the catalytic site, which helps in substrate processing. The metal ion is coordinated at five points with a pentacoordinate geometry. It forms three coordinations with the side chain O(δ) of Asp174, Asp262 and the N(δ) of His176. It is also coordinated to the carbonyl and the hydroxyl oxygen of the ligand's hydroxamate group. It is postulated that in the case of the natural substrate, these two coordination sites are occupied by the carbonyl oxygen of the acetylated lysine and a water molecule. The hydroxyl group of the Tyr301 residue forms a hydrogen bond donor interaction with the carbonyl oxygen of the ligand's hydroxamate group. The N(ε) of His138 and His139 form hydrogen bonding interactions with the hydroxamate group of the ligand. The N(δ) of His138 and His139 form hydrogen bonds with the residues Asp172 and Asp262, and appear to create a charge relay system similar to the Asp-His charge relay system present in serine proteases. This interaction is predicted to enhance the basicity of the N(ε) thereby allowing it to initiate proton abstraction. The Zn^{2+} ion plays a major role in the catalytic process of the enzyme by polarizing the carbonyl group of the acetylated lysine, reducing the overall entropy by coordinating to both the substrate and water molecule and enhancing the nucleophilicity of the water by lowering the pK_a of its proton. The proton from the water may be abstracted by His138, initiating a nucleophilic attack of the free hydroxyl on the carbonyl carbon of the substrate. The tetrahedral intermediate is likely to be stabilized by the coordination with the metal ion as well as the hydrogen bond to Tyr301. The collapse of the tetrahedral intermediate would lead to the liberation of free lysine and acetic acid as products. Mutation analysis[27,28] carried out on the homologous human HDAC1 enzyme from human as well as the Rpd3 HDAC from yeast show that the *Pf*HDAC-1 residues corresponding to His138, His139, Asp172, Asp174 and His176, lead to a loss in enzymatic activity upon mutation. This data corroborates with the ligand-refined *Pf*HDAC-1 model wherein the first three residues form a part of the charge relay system, while Asp174 and His176 form the coordinations with the metal ion which is critical for the catalytic process.

2.4. Molecular dynamics simulation

A one nanosecond molecular dynamics simulation was carried out using the ligand-refined model of

*Pf*HDAC-1 to identify the stability of the enzyme's binding site as well as the induced-fit effects that might be present particularly in the upper region of the binding site where the dimethyl anilino group of the ligand binds. The potential energy plot (Fig. 3a) of the production phase of the simulation does not show any major energy drift and is suggestive of an energetically stable protein–ligand complex. The heavy atom RMSD fluctuations of the residue side chains from the upper region of the binding site (Fig. 3b) vary within a range of 0.48 Å and point towards a stable binding site.

In the normal substrate processing, the upper region of the binding site is occupied by the main chain of the substrate. This region is solvent exposed and is prone to induced-fit effects accommodating for different substrates with varying amino acids adjacent to the acetylated lysine residue being processed by the enzyme. In our ligand-refined model the dimethyl anilino group of TSA lies in this region and is solvent exposed while the hydroxamate terminal is coordinated to the metal center. The upper solvent exposed region accounts for the RMSD fluctuations (Fig. 3b) seen in the ligand. This movement also seems to be correlated with an interaction between the nitrogen atom of the dimethyl anilino group and the side chain hydroxyl group of Thr96 (Fig. 3d). The simulation snapshots between 0.2–0.3 ns and 0.8–0.9 ns (Fig. 3b and d) correspond to lower ligand heavy atom RMSD and a shorter T96[O]–TSA[N] distance. On the other hand the simulation snapshots between 0.3–0.4 ns and 0.9–1.0 ns (Fig. 3b and d) correspond to an increase in the ligand heavy atom RMSD and the T96[O]–TSA[N] distance. The T96[O]–TSA[N] distance monitor is $\leqslant$3.5 Å for approximately 28.5% of the overall simulation time and is suggestive of a possible polar interaction between the ligand and the protein. Although the anilinic nitrogen atom at the *para* position is not ideally placed, the possibility of the formation of an interaction with the Thr96 provides us with an interesting clue for structure-based modifications towards

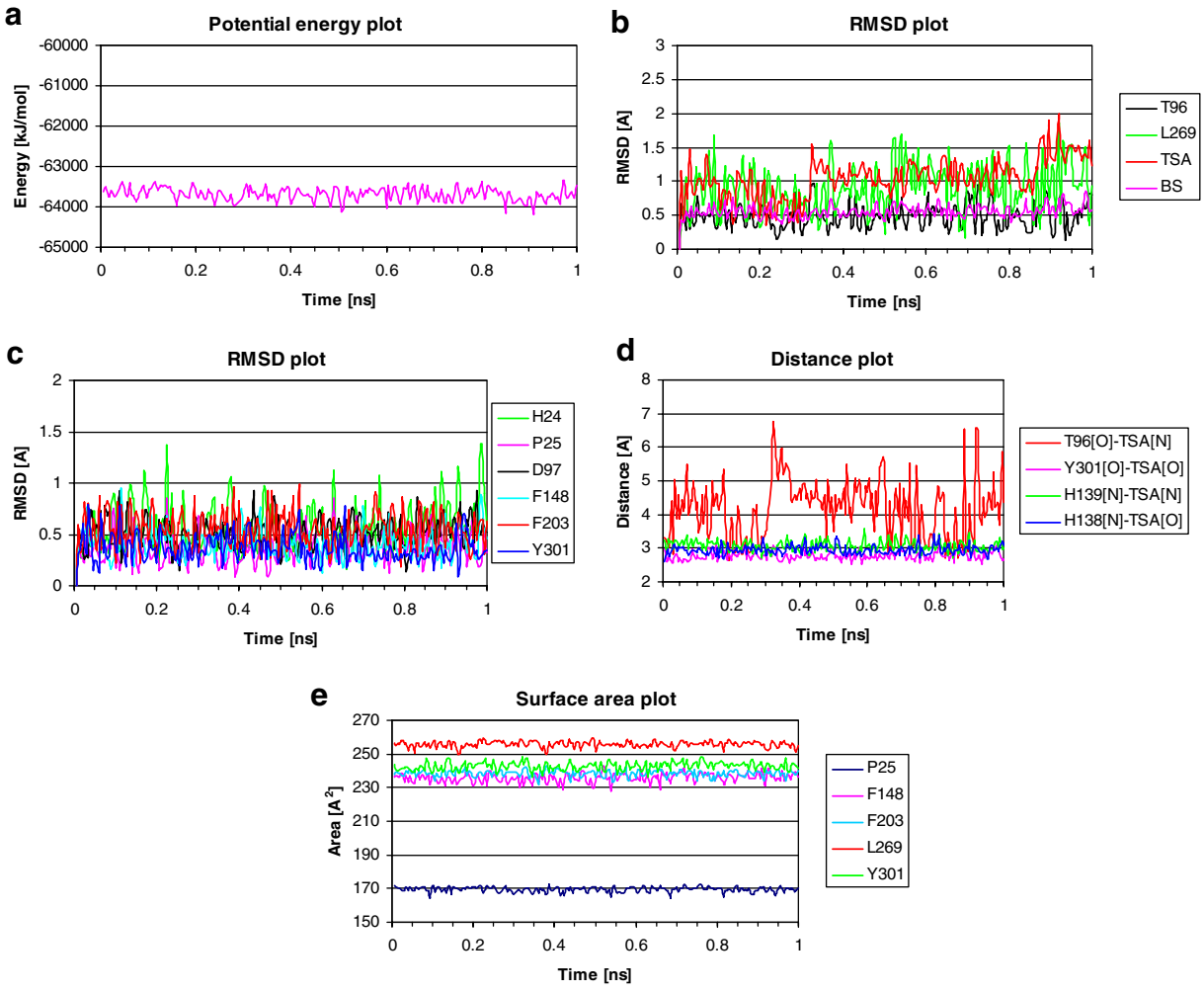

Figure 3. (a) Potential energy plot for the production phase of the dynamics simulation. (b) Heavy atom RMSD of the residue side chains from the upper region of the *Pf*HDAC-1 binding site (BS), TSA and two residue side chains which display maximum fluctuations. The residue side chains considered for BS are H24, P25, T96, D97, F148, F203, L269 and Y301. (c) Heavy atom RMSD of residue side chains from the upper region of the *Pf*HDAC-1 binding site. (d) Monitors for fluctuation of distances between polar atoms of the ligand and relevant heteroatoms of the binding site residues. (e) Van der Waals surface area plots of selected residue side chains from the upper region of the *Pf*HDAC-1 binding site.

enhancing the binding affinity of the putative inhibitors. The other three distance monitors (Fig. 3d) pertaining to hydrogen bonding interactions between the hydroxamate moiety of TSA and the binding site residues His138, His139 and Tyr301 are fairly stable during the course of the simulation. Amongst the amino acid side chains lining the upper region (Fig. 3b and c) of the *Pf*HDAC-1 binding site, Leu269 shows the maximum fluctuation. Leu269 is one of the most exposed residues in the binding pocket (Fig. 3e) and undergoes side chain movements (Fig. 3b), of which some (0.0–0.1 ns, 0.3-0.4 ns and 0.85–0.95 ns) seem to be correlated with the movement of the ligand. Amongst the other residues lining the upper region of the *Pf*HDAC-1 binding site, His24 (Fig. 3c) shows a higher degree of RMSD fluctuation. The residue is located in the upper periphery of the binding pocket and is prone to induced-fit effects accommodating for the enzyme substrate. The residues (Fig. 3b and c) other than Leu269 and His24 do not show any major movement during the course of the simulation. The maximum fluctuation range and standard deviation for side chain heavy atom RMSD for any of these residues are $\leqslant$0.86 Å and $\leqslant$0.17 Å, respectively.

2.5. Docking study

A docking study was conducted to evaluate the predictive ability of the homology model and its suitability for use in the structure-based drug design studies. A set of known human HDAC inhibitors with a hydroxamate ZBM were used for this study. These inhibitors were found to elicit antimalarial activity (Table 2) in an in vitro assay against the chloroquine sensitive D6 and resistant W2 strains of *P. falciparum*. Presuming that *Pf*HDAC-1 inhibition is the likely cause for their antimalarial activity, a scoring metric should be able to evaluate and rank the binding poses of these ligands in *Pf*HDAC-1 in order of their biological activity. The analysis is of semiquantitative nature and a logical correlation between the scores and the biological activity distribution would provide an indirect validation of the predictive ability of the model.

In order to generate a validated docking protocol for this score-based active/inactive separation study, it was necessary to conduct a docking pose validation first. The ligand-refined pose of TSA from the model was utilized for conducting this validation. Ten poses of TSA were generated using the Glide SP (standard precision) mode and re-ranked using Glide score, Emodel and Glide energy score. The top ranked solutions (Fig. 4a and b) for all the three scoring functions had heavy atom RMSD's of <1.0 Å with the modeled pose of TSA suggesting that either of the scoring functions could be utilized for the docking study. The attainment of a pose with low RMSD was also suggestive of the geometric accuracy of the model's binding site and the ligand pose since the modeled pose could be replicated independently, using a docking program.

In the score-based active/inactive separation study the four ligands including TSA were docked using the same docking protocol as used in pose validation and the top ranked poses of each ligand based on Glide score, Emodel and Glide energy score were extracted. Amongst the three scoring functions (Table 2), the best trends were observed for Glide energy followed by Emodel and Glide score. Glide energy ranked the four molecules in the same order as their biological activity with the TSA pose getting the highest score. The Emodel score ranked TSA in the first place but SHA was scored higher than ABHA and SBHA. The rank order for Glide score was the worst wherein the weakest active SHA was ranked highest and the most active TSA was ranked second. A break up of the Glide energy term showed that the van der Waals term showed a better correlation with the biological activity as compared to the coulombic term. Comparison of the ligand structures show that the activity of these molecules correlate well with the size of their hydrophobic head groups and corroborates with the docking results, which point towards the importance of the van der Waals interactions.

All of the ligands used in the docking study have a hydroxamate group as a ZBM and they coordinate with the metal ion through the carbonyl and hydroxyl oxygens of the hydroxamate group. A hydrogen bonding interaction between the Tyr301 hydroxyl and the carbonyl oxygen of the hydroxamate group is observed in all of the poses. Hydrophobic interactions play a major role in the stability of the ligand–protein complex. TSA (Fig. 4a) is the most potent of all the compounds with a two digit nanomolar activity and is seen to form extensive hydrophobic contacts with His24, Pro25, Phe148, Phe203, Leu269 and Tyr301. The compounds ABHA and SBHA are weaker than TSA and both have activities in the same log order range. In the case of ABHA (Fig. 4c), the alkyl chain is longer compared to TSA and the terminal dimethyl anilino moiety is replaced by a hydroxamate group. The alkyl chain occupies the

Table 2. The table shows the results of the in vitro biological evaluation of the four human HDAC inhibitors against the chloroquine sensitive (D6, Sierra Leone) and resistant (W2, Indo China) strains of *P. falciparum*

	IC_{50} (μM) D6	IC_{50} (μM) W2	pIC_{50} (D6)	pIC_{50} (W2)	Glide score	Emodel	Glide energy	evdw	ecoul
TSA	0.0156	0.01	7.81	8.01	−5.08	−61.07	−46.35	−28.90	−17.45
SBHA	2.23	1.53	5.65	5.82	−0.55	−41.68	−42.59	−16.40	−26.82
ABHA	5.5	3.9	5.26	5.41	−0.11	−38.84	−40.80	−17.94	−22.86
SHA	35.7	51.3	4.45	4.29	−5.47	−47.36	−33.31	−10.93	−22.38
CQ[a]	0.0124	0.256							
Art[a]	0.009	0.001							

The docking scores for the best ranked poses of these molecules in *Pf*HDAC-1 are also listed for comparison.
[a] CQ, chloroquine; Art, artemisinin.

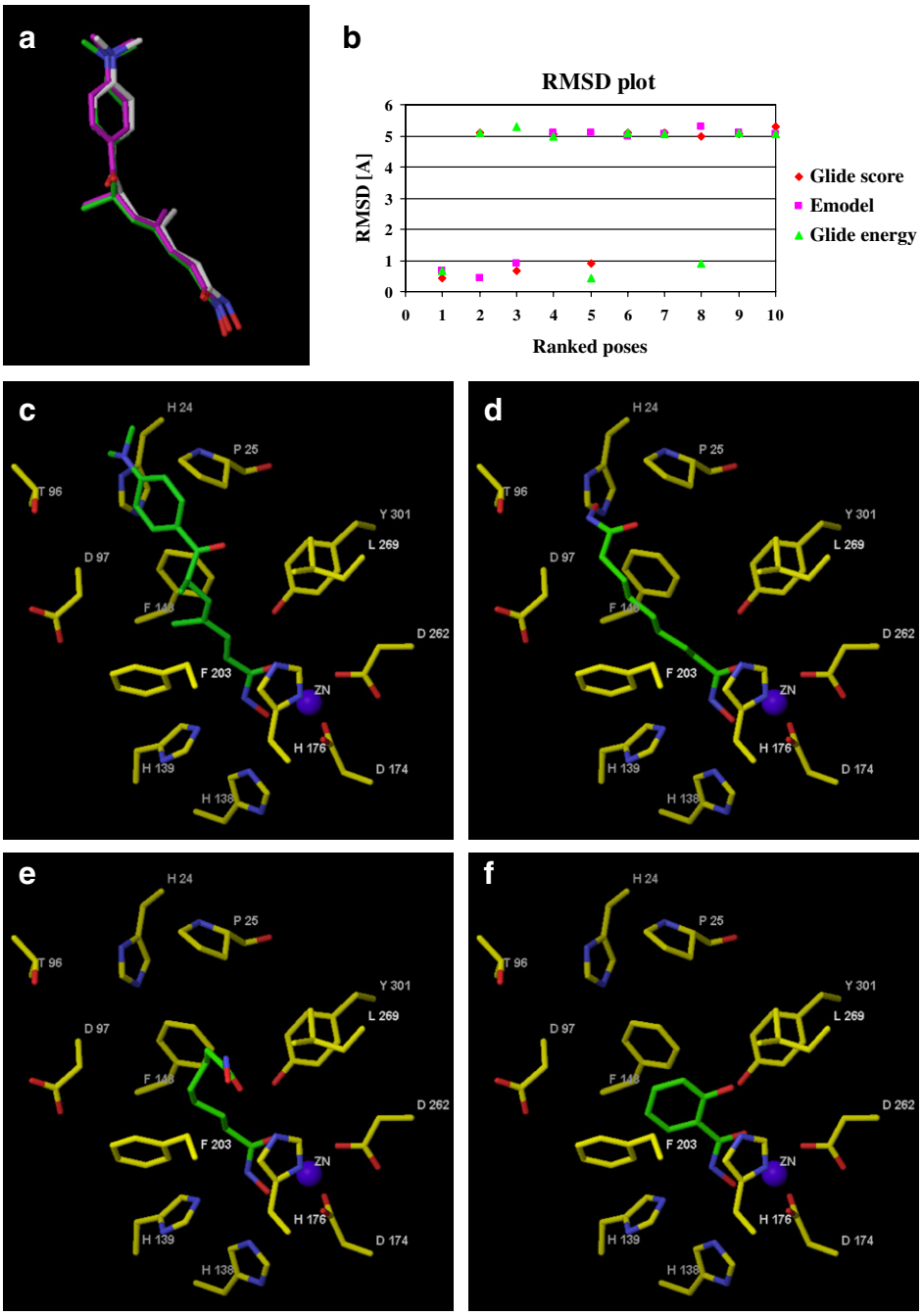

Figure 4. (a) An overlay of the modeled pose of TSA (green) and the top ranked pose based on Glide score (magenta), Emodel/Glide energy score (white). (b) A plot showing the heavy atom RMSD of the ten poses with the modeled pose of TSA ranked on the basis of Glide score (red), Emodel (magenta) and Glide energy score (green). The top ranked poses for (c) TSA, (d) ABHA (e) SBHA and (f) SHA based on Glide energy score.

length of the narrow channel but the replacement of the hydrophobic group with a polar moiety leads to a loss in van der Waals interaction. The hydroxyl oxygen of the free hydroxamate group is at 4.37Å from the N(δ) of His24, while the nitrogen of the hydroxamate group is 4.37Å from the hydroxyl oxygen of Thr96. Minor induced-fit effects in this region might lead to the formation of stronger interactions of polar nature. In the case of SBHA (Fig. 4e), the alkyl linker is one carbon shorter than ABHA and it does not occupy the whole length of the narrow channel. The overall vector of the alkyl linker of SBHA is similar to that of ABHA till the fourth carbon from the hydroxamate group. The rest of the molecule attains a different vector and comes out of the narrow groove, forming hydrogen bonding interactions with the N(ε) of His176 and backbone N of Phe203 through its free hydroxamte group. SHA (Fig. 4f) is the weakest of all the inhibitors and is likely to be so because of the reduced size of the hydrophobic group. The phenyl ring is not able to occupy the whole length of the narrow channel and is less adept at competitively replacing the natural substrate from the binding groove. The hydroxyl group *ortho* to the

hydroxamate group is in close proximity to the hydroxyl group of Tyr301 and can form an alternate hydrogen bonding interaction instead of the carbonyl oxygen of the hydroxamate group. Overall, the biological activity of these molecules correlates well with the energetic stability and the non bonded interaction profile of the docked poses of these molecules in *Pf*HDAC-1.

2.6. Comparison with human HDAC

A structural comparison was carried out with the human HDAC8 (1T64, PDB code) to determine similarities in their overall protein topology as well as pertinent structural features of their respective binding sites. The two structures superimposed with a heavy atom RMSD of 1.21 Å. Comparison of the amino acid sequences as well as the structures of the two proteins helped us in delineating some unique structural differences between the binding sites of *Pf*HDAC-1 and human HDAC8. The residues forming the catalytic centers of the two enzymes are totally conserved (Table 1) and the differences are observed primarily at the top end of the binding site. Sequence analysis showed that *Pf*HDAC-1 showed more identity to the HDLP sequence than the human HDAC8 sequence in the region extending from the N-terminus to the end of the L1 loop. This loop region is shorter by two residues in human HDAC8 as compared to both *Pf*HDAC-1 and HDLP. This loop forms a part of the binding cavity and in the case of human HDAC8[20] attains a wide variety of conformational states. It makes this region of human HDAC8's binding site more malleable and prone to induced-fit effects as compared to HDLP and *Pf*HDAC-1. In the case of the HDAC8–CRA-A complex (1VKG, PDB code), the loop moves away from the binding site to accommodate the larger ligand. On the contrary, in the case of the HDAC8–TSA complex the loop falls back towards the binding site and more order is incorporated into the L2 loop through π–π stacking interactions between the ligand's dimethyl anilino group and Tyr100 (HDAC8). This in turn exposes a secondary binding site, which is occupied by another molecule of TSA.

In the case of HDLP, the loop is larger and does not show any significant movement as seen from the three ligand bound crystal structures of HDLP[21] The larger loop also blocks the second cavity and prevents ligand entry into the secondary site. The *Pf*HDAC-1 structure in this region matches closely with that of HDLP and is likely to be less malleable as compared to what is seen in the case of human HDAC8. Sequence analysis[21] suggests that the binding sites of human HDACs 1–3 would also have a similar conformational state as *Pf*HDAC-1 since they all have the L1 loops of the same length as HDLP. One may further postulate that the nature of the substrates processed by human HDAC8 might be more varied and different from those processed by HDLP, human HDACs 1–3 and the *Pf*HDAC-1.

Comparison of residues (Fig. 5) around the upper region of the binding site suggests two unique differences. One of the differences is seen in terms of the positively charged residue Arg268, located at the upper right hand side of the binding site. In the case of human HDAC8 and HDLP this residue is replaced by a proline and tyrosine, respectively. However, sequence comparison[20,21] with the human HDACs 1–3, 6 and other plasmodial HDAC sequences (unpublished results) suggests that this arginine residue is conserved in all of them except human HDAC6, which has a proline at this position. This residue in more than 4 Å away from any of the ligand atoms in the current model but may be accessible for larger ligands. The other residue of interest is Thr96 of *Pf*HDAC-1, which is replaced by a tyrosine residue in the case of both human HDAC8 and HDLP. Sequence analysis showed that this threonine residue is conserved in a number of plasmodial HDAC sequences (unpublished results) whereas it is replaced by a glutamate in the case of human HDACs 1 and 2, an aspartate in the case of human HDAC3 and a tyrosine in case of human HDAC6. This residue is located at the top end of the binding site and in the case of HDAC8–TSA complex shows a π–π stacking interaction with the ligand's dimethyl anilino group. Possibilities also exist for a polar interaction between the tertiary nitrogen and the hydroxyl group of the tyrosine residue. In the ligand

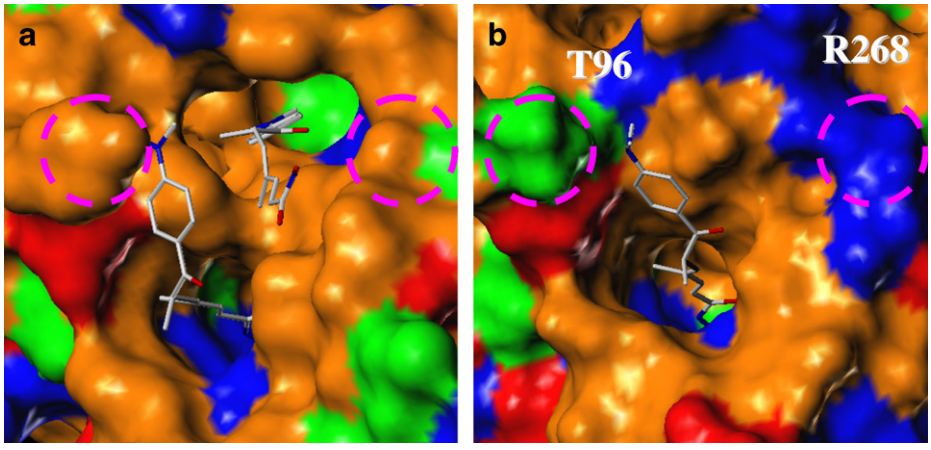

Figure 5. The surface representation of the binding site of (a) HDAC8–TSA complex (1T64, PDB code) and the (b) Ligand-refined *Pf*HDAC-1 model. The surface residues are color coded based on their nature: electronegative (red), electropositive (blue), polar (green) and hydrophobic (orange). The magenta circles demarcate the location of two residues, which are not conserved between human HDAC8 and *Pf*HDAC-1.

bound HDLP-TSA structure, this residue is seen to undergo a change in its rotameric state from its *apo* form to accommodate the ligand's dimethyl anilino group. By comparing the sequences and the available crystal structures of HDACs one might speculate that the presence of an electronegative group in this area, vectored into the solvent could provide binding selectivity towards the plasmodial HDACs. Presence of such a group would enhance the polar interaction with the threonine residue of *Pf*HDAC-1 but impede the interaction with the tyrosine residue of human HDACs 6 and 8, and also provide a mismatch for the acidic residues present at this position in the case of the human HDACs 1–3.

3. Conclusions

With the evolution of drug resistant strains of the malaria parasite, there is an urgent need for the development of highly selective and efficacious antimalarial therapies designed against novel plasmodial targets. The *Pf*HDAC-1 enzyme from *P. falciparum* represents one such target of interest. A ligand-refined model of *Pf*HDAC-1 was generated using the crystal structures of closely related homologues human HDAC8 and HDLP using an OPLS/GBSA potential setup. The structural accuracy of the model was extensively validated using protein structure check tools. A molecular dynamics simulation was carried out to evaluate the stability of the enzyme's binding site and the induced-fit effects present in the upper region of the binding site. Extensive docking studies were carried out using a set of known human HDAC inhibitors, which show antimalarial activity at the in vitro against the chloroquine sensitive (D6, Sierra Leone) and resistant (W2, Indo China) strains of *P. falciparum*. A semiquantitative correlation was observed between the biological activity of these compounds and the docking scores for these compounds suggesting that the inhibition of the *Pf*HDAC-1 might be one of the possible reasons for their antimalarial activity. Comparison of the binding site of the *Pf*HDAC-1 with that of HDAC8 from human suggested certain subtle differences in terms of the topology of the binding cavity as well as the amino acid residues lining the binding pocket. These characteristics may be exploited for the development of selective inhibitors of *Pf*HDAC-1, as potential new antimalarials. The model in conjunction with the validated docking protocol could be used for an extensive structure-based virtual screening study to identify putative inhibitors of *Pf*HDAC-1 with the novel chemical scaffolds. Efforts are also currently underway to express a functional form of the *Pf*HDAC-1 enzyme for utilization in biochemical assay to identify selective *Pf*HDAC-1 inhibitors.

4. Materials and methods

4.1. Computational resources

All the computational studies were performed on a 2.8 GHz Pentium 4 Linux desktop running Fedora 5.0. Sequence alignment, homology modeling, loop and side chain refinement were carried out using Prime 1.6 (Schrodinger. LLC, Portland, OR). Molecular mechanics optimization was carried out in Macromodel 9.5 (Schrodinger. LLC, Portland, OR) and Glide 4.5 (Schrodinger. LLC, Portland, OR) was used for all docking calculations. Protein structure checks were conducted using the Matchmaker module of Sybyl 6.9 (Tripos Inc, St. Louis, MO) and the ADIT validation server (http://deposit.pdb.org/validate/). Figures were generated using Sybyl 6.9, Pymol 1.0 (http://pymol.sourceforge.net/) and PDBsum (http://www.ebi.ac.uk/pdbsum/).

4.2. Sequence alignment and model building

The protein sequence of *Pf*HDAC-1 (primary Accession No. Q9XYC7) was downloaded from the ExPASy proteomics server[29] A protein–protein BLAST[19] search of the PDB was conducted using the target to identify structural homologues. The closest structural homologues with ligand-bound crystal structures were HDAC8 (41% sequence identity) from human and HDLP (32% sequence identity) a HDAC homologue from the hyperthermophilic bacterium *A. aeolicus*. The crystal structures 1C3R (PDB code) from HDLP and 1T64 (PDB code) from human HDAC8 were utilized as reference structures for building the model.

Within Prime, the target sequence was imported in the fasta format. The reference structures were downloaded from the PDB, imported into maestro and then stripped of the water molecules as well as the unwanted monomeric unit and ligand/metal ion associated with it. The structures were saved in the PDB format for use in Prime. Sequences of the reference homologues were extracted from the respective structure files and aligned with the target sequence using the default alignment procedure of Prime applying the BLOSUM62 similarity matrix. The initial alignment was further modified manually to generate the final alignment (Supplementary information, Fig. S2). The 1C3R (PDB code) structure was used as the template for modeling the residues 5–26 of *Pf*HDAC-1 while the residues 27–374 were built from the 1T64 (PDB code) template. Assignment of the reference template for each target sequence residue was carried out using the Set Template Regions functionality. The ligand and the metal ion were incorporated in the model generation process and the coordinates of the ligand TSA and the Zn^{2+} ions were copied from the 1C3R (PDB code) structure. Model building was automatically terminated where no reference sequence existed for building the template. Side chain rotamers of conserved residues were retained and all side chains were optimized. Side chains which were not derived from the template were minimized as well.

4.3. Model refinement

Loop refinement was carried out for residues 76–82 using the extended low sampling setup. The extended sampling procedure is applicable to loops, which extend

from 6 to 11 residues in length. The sampling method setup sets an accessibility limit on the number of processors that could be utilized to distribute the job, which in the case of 'low' is two processors. Side chain optimizations were performed for the residues Phe203 and Arg268. Atom type, bond order, protonation and tautomeric assignments were manually reset. The prepared structure was submitted for restraint guided molecular mechanics refinement using the OPLS/GBSA potential setup, which involved the OPLS2001 force field and the GBSA implicit solvent treatment with an extended non bonded cutoff (Van der Waals = 8 Å, electrostatic = 20 Å). Five flat bottom distance restraints with a potential well half width of ±0.2 Å were applied to the Zn^{2+} ion and the corresponding coordination sites. The Zn^{2+} atom type from the OPLS force field uses a partial charge of +2. The Van der Waals parameters for the atom type involved a vdw radius (r) of 0.72 Å, collision diameter (σ) of 1.952 Å and a well depth (ε) of 0.2338 kcal/mol. The mean distances for the distance constraints were obtained from the 1C3R (PDB code) structure. The heavy atoms of the ligand, metal, residue side chains and backbones were initially constraints using a force constant of 100 kJ/mol-Å^2. The same force constant was applied to the distance restraints. The system was minimized in multiple steps using 5000 steps of Polak Rabiere conjugate gradient minimization and a convergence criterion of 0.05 kJ/mol. In the first four steps, the force constant on the ligand, metal, and residue side chains were reduced from 100 to 0 (100, 50, 25, 10, and 0). Residue backbones and the distance restraints were similarly released in four steps each. The restraint free model was then minimized to a RMS convergence criterion of 0.01 kJ/mol using the same potential setup and minimization protocol.

4.4. Molecular dynamics simulation

Molecular dynamics simulation was carried out on the ligand-refined *Pf*HDAC-1 model using the same OPLS/GBSA potential setup used in the model refinement stage. Five flat bottom distance restraints with a potential well half width of ±0.2 Å and a weak force constant of 25 kJ/mol-Å^2 were applied to the Zn^{2+} ion and the corresponding coordination sites. The backbone heavy atoms of the protein were constrained with a very weak force constant of 10 kJ/mol-Å^2. SHAKE was applied to all bonds with hydrogens. Molecular dynamics was carried out at 300 K with time step of 2 fs. The system was equilibrated for a period of 100 ps following which a production run of 1ns was carried out. A trajectory for post-dynamics analysis was saved by recording snapshots at intervals of 5 ps.

4.5. Docking calculations

The ligands for the docking study were built in maestro and minimized using the same protocol used for the last stage of the model refinement. The geometric center for the scoring grid was defined using the centroid of the TSA pose from the ligand-refined *Pf*HDAC-1 model. The Zn^{2+} ion was identified for setting up a metal constraint. Docking studies were carried out in the Glide SP mode. The metal constraint was kept on during docking and any ligand atom with an acceptor feature was allowed to fulfill the metal coordination geometry. Ten poses were collected for every ligand and re-ranked using the scoring terms available in Glide.

4.6. In vitro antimalarial assay

Antimalarial activity was determined in vitro on the chloroquine sensitive (D6, Sierra Leone) and resistant (W2, Indo China) strains of *P. falciparum*. The 96-well microplate assay was based on the evaluation of the effect of compounds on the growth of asynchronous cultures of *P. falciparum*, determined by the assay of parasite lactate dehydrogenase (pLDH) activity.[30,31] The appropriate dilutions of the compounds were prepared in DMSO or RPMI-1640 medium and added to the cultures of *P. falciparum* (2% hematocrit, 2% parasitemia) setup in clear flat bottomed 96-well plates. The plates were placed into the humidified chamber and flushed with a gas mixture of 90% N_2, 5% CO_2 and 5% O_2. The cultures were incubated at 37 °C for 48 h. Growth of the parasite in each well was determined by pLDH assay using Malstat® reagent.[30] The medium and RBC controls was also setup in each plates. The standard antimalarial agents chloroquine and artemisinin were used as the positive controls.

Acknowledgment

We thank Dr. Dan Schafer of Schrodinger. LLC, Portland, OR for his helpful suggestions.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2008.03.005.

References and notes

1. Choi, S.; Mukherjee, P.; Avery, M. *Curr. Med. Chem.* **2008**, *15*, 161–171.
2. Weinmann, H.; Ottow, E. *Annu. Rep. Med. Chem.* **2004**, *39*, 185–196.
3. de Ruijter, A. J. M.; van Gennip, A. H.; Caron, H. N.; Kemp, S.; van Kuilenburg, A. B. P. *Biochem. J.* **2003**, *370*, 737–749.
4. Longhurst, H. J.; Holder, A. A. *Mol. Biochem. Parasitol.* **1995**, *69*, 111–113.
5. Creedon, K. A.; Kaslow, D. C.; Rathod, P. K.; Wellems, T. E. *Mol. Biochem. Parasitol.* **1992**, *54*, 113–115.
6. Bennett, B. J.; Thompson, J.; Coppel, R. L. *Mol. Biochem. Parasitol.* **1995**, *70*, 231–233.
7. Joshi, M. B.; Lin, D. T.; Chiang, P. H.; Goldman, N. D.; Fujioka, H.; Aikawa, M.; Syin, C. *Mol. Biochem. Parasitol.* **1999**, *99*, 11–19.

8. Merrick, C. J.; Duraisingh, M. T. *Eukaryot. Cell* **2007**, *6*, 2081–2091.
9. Singh, S. B.; Zink, D. L.; Polishook, J. D.; Dombrowski, A. W.; Darkin-Rattray, S. J.; Schmatz, D. M.; Goetz, M. A. *Tetrahedron Lett.* **1996**, *37*, 8077–8080.
10. Darkin-Rattray, S. J.; Gurnett, A. M.; Myers, R. W.; Dulski, P. M.; Crumley, T. M.; Allocco, J. J.; Cannova, C.; Meinke, P. T.; Colletti, S. L.; Bednarek, M. A.; Singh, S. B.; Goetz, M. A.; Dombrowski, A. W.; Polishook, J. D.; Schmatz, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13143–13147.
11. Meinke, P. T.; Colletti, S. L.; Doss, G.; Myers, R. W.; Gurnett, A. M.; Dulski, P. M.; Darkin-Rattray, S. J.; Allocco, J. J.; Galuska, S.; Schmatz, D. M.; Wyvratt, M. J.; Fisher, M. H. *J. Med. Chem.* **2000**, *43*, 4919–4922.
12. Mai, A.; Cerbara, I.; Valente, S.; Massa, S.; Walker, L. A.; Tekwani, B. L. *Antimicrob. Agents Chemother.* **2004**, *48*, 1435–1436.
13. Andrews, K. T.; Walduck, A.; Kelso, M. J.; Fairlie, D. P.; Saul, A.; Parsons, P. G. *Int. J. Parasitol.* **2000**, *30*, 761–768.
14. Parsons, P. G.; Hansen, C.; Fairlie, D. P.; West, M. L.; Danoy, P. A.; Sturm, R. A.; Dunn, I. S.; Pedley, J.; Ablett, E. M. *Biochem. Pharmacol.* **1997**, *53*, 1719–1724.
15. Sabnis, Y. A.; Desai, P. V.; Rosenthal, P. J.; Avery, M. A. *Protein Sci.* **2003**, *12*, 501–509.
16. Sabnis, Y.; Rosenthal, P. J.; Desai, P.; Avery, M. A. *J. Biomol. Struct. Dyn.* **2002**, *19*, 765–774.
17. Patny, A.; Desai, P. V.; Avery, M. A. *Proteins* **2006**, *65*, 824–842.
18. Desai, P. V.; Avery, M. A. *J. Biomol. Struct. Dyn.* **2004**, *21*, 781–790.
19. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403–410.
20. Somoza, J. R.; Skene, R. J.; Katz, B. A.; Mol, C.; Ho, J. D.; Jennings, A. J.; Luong, C.; Arvai, A.; Buggy, J. J.; Chi, E.; Tang, J.; Sang, B. C.; Verner, E.; Wynands, R.; Leahy, E. M.; Dougan, D. R.; Snell, G.; Navre, M.; Knuth, M. W.; Swanson, R. V.; McRee, D. E.; Tari, L. W. *Structure* **2004**, *12*, 1325–1334.
21. Finnin, M. S.; Donigian, J. R.; Cohen, A.; Richon, V. M.; Rifkind, R. A.; Marks, P. A.; Breslow, R.; Pavletich, N. P. *Nature* **1999**, *401*, 188–193.
22. Godzik, A.; Kolinski, A.; Skolnick, J. *J. Mol. Biol.* **1992**, *227*, 227–238.
23. Laskowski, R. A.; Moss, D. S.; Thornton, J. M. *J. Mol. Biol.* **1993**, *231*, 1049–1067.
24. Ayer, D. E. *Trends Cell Biol.* **1999**, *9*, 193–198.
25. Tsai, S. C.; Seto, E. *J. Biol. Chem.* **2002**, *277*, 31826–31833.
26. David, G.; Neptune, M. A.; DePinho, R. A. *J. Biol. Chem.* **2002**, *277*, 23658–23663.
27. Hassig, C. A.; Tong, J. K.; Fleischer, T. C.; Owa, T.; Grable, P. G.; Ayer, D. E.; Schreiber, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 3519–3524.
28. Kadosh, D.; Struhl, K. *Genes Dev.* **1998**, *12*, 797–805.
29. Bairoch, A.; Apweiler, R. *Nucleic Acids Res.* **1997**, *25*, 31–36.
30. Makler, M. T.; Ries, J. M.; Williams, J. A.; Bancroft, J. E.; Piper, R. C.; Gibbins, B. L.; Hinrichs, D. J. *Am. J. Trop. Med. Hyg.* **1993**, *48*, 739–741.
31. Makler, M. T.; Hinrichs, D. J. *Am. J. Trop. Med. Hyg.* **1993**, *48*, 205–210.