J Mol Model (2013) 19:485–496
DOI 10.1007/s00894-012-1572-3

ORIGINAL PAPER

Molecular and structural insight into plasmodium falciparum RIO2 kinase

Devendra K. Chouhan · Ashoke Sharon · Chandralata Bal

Received: 6 June 2012 / Accepted: 14 August 2012 / Published online: 5 September 2012

Abstract Among approximately 65 kinases of the malarial genome, RIO2 (right open reading frame) kinase belonging to the atypical class of kinase is unique because along with a kinase domain, it has a highly conserved N-terminal winged helix (wHTH) domain. The wHTH domain resembles the wing like domain found in DNA binding proteins and is situated near to the kinase domain. Ligand binding to this domain may reposition the kinase domain leading to inhibition of enzyme function and could be utilized as a novel allosteric site to design inhibitor. In the present study, we have generated a model of RIO2 kinase from *Plasmodium falciparum* utilizing multiple modeling, simulation approach. A novel putative DNA-binding site is identified for the first time in PfRIO2 kinase to understand the DNA binding events involving wHTH domain and flexible loop. Induced fit DNA docking followed by minimization, molecular dynamics simulation, energetic scoring and binding mode studies are used to reveal the structural basis of PfRIO2-ATP-DNA complex. Ser105 as a potential site of phosphorylation is revealed through the structural studies of ATP binding in PfRIO2. Overall the present study discloses the structural facets of unknown PfRIO2 complex and opens an avenue toward exploration of novel drug target.

Keywords *Archaeoglobus fulgidus* RIO2 kinase · Induced fit docking · Molecular docking · Molecular modeling · *Plasmodium falciparum* RIO2 kinase

Abbreviations

DNA	Deoxyribonucleic acid
ATP	Adenosine triphosphate
RIO	Right open reading frame

D. K. Chouhan · A. Sharon · C. Bal (✉)
Department of Applied Chemistry, Birla Institute of Technology, Mesra,
Ranchi, Jharkhand 835215, India
e-mail: cbal@bitmesra.ac.in

Introduction

Malaria is a mosquito-borne infectious disease caused by parasitic protozoa (an eucaryote) of the genus *Plasmodium,* and severe disease is largely caused by *Plasmodium falciparum.* Today drug-resistant malaria is a persistent global health threat following the emergence of chloroquine resistance. Artemisinin combination therapy has been useful, but the recent emergence of resistance [1, 2] provides strong motivation to identify novel drug targets.

Metabolism of host haemoglobin as a nutrient source by the parasite involves production of pro-oxidant heme, peroxide and other free radicals which are detoxified by the glutathione and thioredoxin dependent parasitic antioxidant system as well as by up-regulation of membrane transporters that acquire small molecule antioxidants from the host [3]. Several drug targets have been identified to disturb this redox balance and induce parasitic death due to an apoptotic like process [4–6]. On the other hand, the parasite can upregulate its own antioxidant defence to acquire drug resistance [7–9]. Transcriptome studies of intra-erythrocytic stages of parasite indicate that oxidative stress activates the transcriptional machinery to produce antioxidant enzymes, proteins involved in merozoite invasion, drive different phases of lifecycle in erythrocytes [10–13]. In eukaryotic system, protein kinases are involved in responses to oxidative stress and survival of the organism. The malarial genome contains approximately 65 protein kinases including an atypical family of RIO kinases [14–16]. One of the members of RIO kinase family, RIO2 kinase is unique because it has a kinase domain capable of catalyzing phosphorylation, similar to other kinases [17, 18] as well as a highly conserved N-terminal winged-helix (wHTH) domain [19–21]. No such combination of winged helix domain and kinase catalytic domain has been observed previously. Although the function of winged helix domain is still

unknown, it resembles the wing like domain found in DNA binding proteins (capable of binding DNA and single stranded nucleotides). Recently, the crucial role of RIO2 enzymatic activity for cleavage of 20S pre-rRNA, ribosome biogenesis [22, 23] and the possibility to develop therapeutic agents against this protein have been reported [24–26]. Therefore, we chose to explore PfRIO2 kinase as a plausible novel anti-malarial drug target as an effort to support our ongoing anti-malarial drug discovery program. The present era of drug discovery heavily depends on significant 3D structural information regarding the target protein. As no co-crystal structure or validated model of PfRIO2 is available, a 3D model utilizing the available crystal structure (pdb:1ZAO) of RIO2 kinase from *Archaeoglobus fulgidus* (AfRIO2) was built [27]. Molecular modeling and molecular dynamics (MD) simulations were conducted to study the putative auto-phosphorylation site. A site map analysis was carried out to predict possible binding sites in the overall protein, which suggested four possible sites. Among them a plausible DNA binding groove was identified and a DNA probe (DNA trinucleotide base-pair) was successfully docked by induced fit (IF) mechanism to produce the PfRIO2-DNA-ATP complex structure. To the best of our knowledge this is the first report of application of IF methodology for DNA docking into a protein-ATP complex.

All the known kinase inhibitors (drugs) target the ATP binding site. Therefore, we first studied ATP binding to PfRIO2 to validate the newly built model. The wHTH domain of PfRIO2 is situated very close to the catalytic kinase domain. The binding of any ligand to this domain may reposition the kinase domain leading to inhibition of kinase function, which could be utilized as a novel allosteric site to design kinase inhibitor. Therefore, we analyzed the possibility of DNA binding to utilize this site as a novel kinase inhibitors (drugs) target.

Materials and methods

In-absence of PfRIO2 co-crystal structure, a valid RIO2 template was necessary to build the model of PfRIO2. Until today the only crystal structure reported for RIO2 is from *Archaeoglobus fulgidus* (AfRIO2). Therefore, we intended to utilize AfRIO2 as a template for our study. Further, to understand the sequence similarity pattern of RIO2 in various organism, a sequence alignment of RIO2 from *Plasmodium falciparum* (protozoa), *Archaeoglobus fulgidus* (archaea), *Homo sapiens* (Human), *Saccharomyces cerevisiae* (yeast) and protein kinase alfa (from human) was conducted. All these protein sequences were obtained from universal protein resource UniProtKB database (ID Q8I1N8, O30245, Q9BVS4, P40160 and P17612 respectively). The sequences were aligned using multiple sequence alignment utility in sequence viewer tool of Schrödinger suite 2011 and the sequence alignment is shown in Fig. 1.

Most of the modeling and simulation were carried out with the modeling suite from Schrödinger 2011 [28]. The 3D models of RIO2 Kinase-ATP complex were built and refined using the Prime module [29]. The Protein Preparation Wizard (PPW) and Epik module of Schrödinger suite were utilized iteratively to check for any errors related to bond-order, metal charges (+2 for Mn ion), proton assignment, protonation state or H-bonding to ensure the chemical correctness of model. MD simulations were conducted through the Maestro interface to Desmond [30]. As ATP is the natural substrate, it was used to validate the docking protocol. In one case, ATP from the crystal structure (pdb: 1ZAO) and in the other case manually built and conformationally optimized ATP (using Maestro) was docked into the active site. SiteMap analysis followed by IF docking [31] was used to identify the DNA binding site. No constraints or restraints were applied to the enzyme residues, cofactors or ATP during the simulation studies. The crystallographic waters were removed to avoid conformational discrepancies associated with water sampling during the simulations. Overall, the implicit-water model was used during minimization, conformational search, molecular docking, SiteMap analysis, loop refinement and energetic calculation using Schrodinger 2011. Only the MD simulations studies were conducted using the explicit-water model using Desmond.

AfRIO2-ATP model development and optimization

The available crystal structure of AfRIO2 (pdb: 1ZAO) provided a major modeling platform to construct a model, which can work for structural simulation and detail investigation. The experimental crystal structure has several missing residues in flexible loop region (amino acid 128–142) indicative of its dynamic role for conducting biological catalysis. Therefore, a complete model structure is necessary to understand the structural perspectives and its relation with biological processes. Firstly, the template structure was taken from protein data base and missing residues 128–142 were added using Prime module of PPW followed by loop refinement through Prime using extended sampling [29]. Then the structure was pre-processed using PPW module of Schrodinger 2011 which started with assignment of bond order and addition of hydrogens. Then all heteroatoms were identified and manually corrected for proper ionization states. All the waters were removed and the protein assignment was run to select the most likely position for hydroxyl, thiol protons as well as the most probable protonation states and tautomers of His residues and to check the "χ-flip" assignments for Asn, Gln and His residues. Finally a restrained minimization using OPLS2005 force field was conducted to refine disallowed torsion angles and eliminate

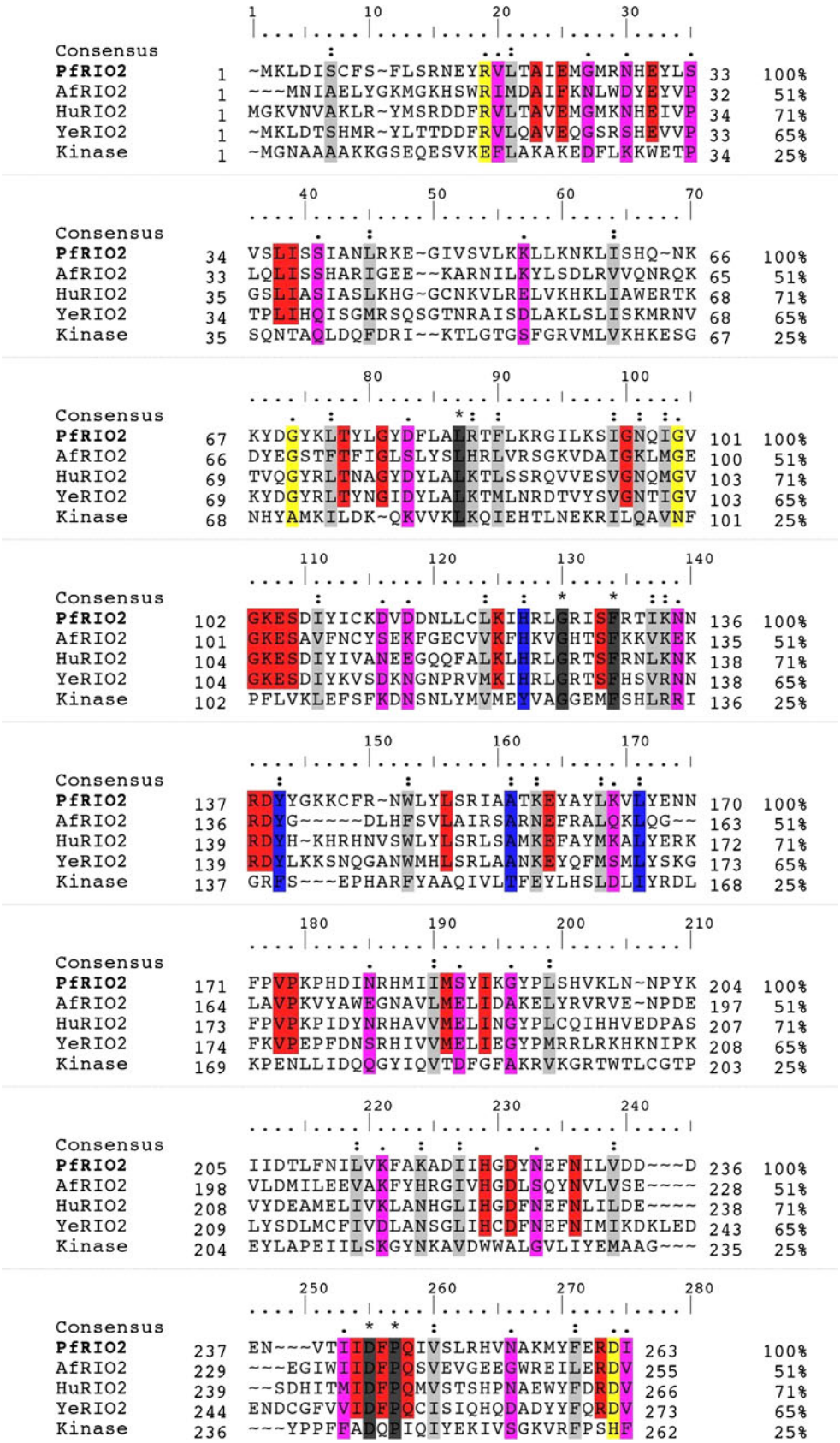

Fig. 1 Multiple sequence alignment between RIO2 protein from *Plasmodium falciparum* (PfRIO2), *Archaeoglobus fulgidus* (AfRIO2), *Homo sapiens* (HuRIO2), *Saccharomyces cerevisiae* (YeRIO2) and kinase alfa (Kinase from human) with UniprotKB ID Q8I1N8, O30245, Q9BVS4, P40160, P17612 respectively. The conserved, strongly similar and weakly similar residues marked as (*, *dark gray*), (:, *light gray*) and (., *magenta*) respectively. Additionally, the conserved residues only in RIO2 (*red*), similar residues in all but conserved only in RIO2 (*blue*) and weakly similar in all but conserved only in RIO2 (*yellow*) has been observed

unfavorable contacts. The heavy atoms were restrained to within a maximum RMSD of 0.3 Å . Further, side chain refinement was performed on residues within 7.5 Å of the large loop (newly added residues 128–142) using default parameters of Prime. Finally, the Polak-Ribiere conjugate gradient (PRCG) minimizations using OPLS2005 force field for 5000 iterations with convergence of 0.05 $kJ\,mol^{-1}$ was conducted through MacroModel module [32]. The water solvation using generalized-Born/surface-area (GB/SA) algorithm with constant dielectric treatment (1.0) for the electrostatic part of the calculation was chosen during energy minimization. The generated model of AfRIO2-ATP complex was subjected to a 5 ns MD simulation to further check overall stability and an average structure of AfRIO2-ATP complex (Fig. 2a) was chosen for further studies.

PfRIO2 model development

PfRIO2 protein sequence (UniProt accession number Q8I1N8, isolate 3D7) with 581 amino acids [33] was used as the query sequence and aligned with the template of AfRIO2. The comparative modeling structure prediction tool in Prime module of Schrodinger suite 2011 was utilized to build a 3D model of PfRIO2. In the final model building step ATP and two Mn^{2+} were included as was present in the crystal structure of AfRIO2. The similar loop refinement and optimization work-flow were conducted as was discussed for AfRIO2 for final PfRIO2 model optimization (Fig. 2b).

ATP binding by molecular docking

The ATP was extracted from the crystal structure (pdb 1ZAO) and manually corrected for bond order using builder module of Maestro. The minimization of PfRIO2 model without ATP was conducted using MacroModel, followed by redocking of the manually corrected ATP using Glide [34]. To validate our docking protocol, ATP was built manually using maestro interface followed by minimization and conformational search through MacroModel to pick the lowest energy conformer and docked into the ATP binding site of PfRIO2 kinase.

MD simulation

The newly added large loop can possibly impart steric clashes in constructed models. Therefore, both AfRIO2 and PfRIO2 structures were refined through 5 ns MD simulation using Desmond. An SPC water box was placed around each model structure to mimic the biological solvent environment. The solvated structures were minimized with the steepest-descent method and was further relaxed with a 24 ps molecular dynamics simulation at 300 K and 1 atm through “quick relaxation” protocol of Desmond. After the equilibration phase, 5 ns production simulations were run without any constraint using the default temperature of 300 K [35] and 1 atm pressure [36]. All simulations used a 10 Å cut-off radius for both VdW’s and electrostatic interactions along with the smooth particle mesh Ewald method [37] to calculate long-range electrostatic interactions. The energy and structural variation as a function of simulation time was used to analyze the convergence of the final result. The last 3 ns trajectories were used to extract the snapshots for binding free energy calculation. An average structure between 3 ns and 5 ns was generated for final MacroModel minimization and structural analysis. Ramachandran plot was generated to check the overall 3D quality of the built model.

SiteMap analysis for DNA binding site identification

The RIO2 model and crystal structure are able to provide significant information regarding ATP binding site, however there is no information available regarding possible binding site of DNA. Thus, we intended to find the active sites that might be useful for DNA binding. Binding site analysis

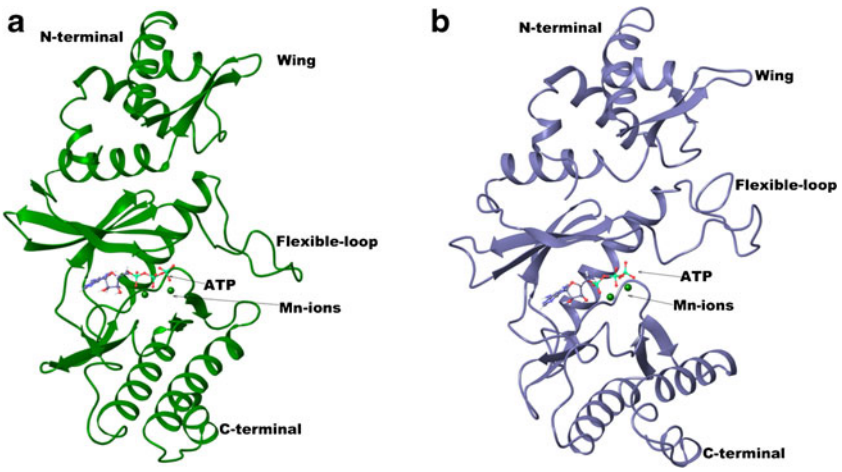

Fig. 2 **a** AfRIO2 model based on crystal structure (pdb: 1ZAO). **b** PfRIO2 model developed using the AfRIO2 model. The major domains such as winged helix, flexible loop, ATP binding site including two Mn2+ ions, C- and N-terminal have been labeled

through SiteMap Program [38, 39] can successfully suggest possible binding sites [40]. Therefore, this method was used and "four site points" were disclosed as possible ligand binding site in PfRIO2 kinase. Online server of MetaDBSite [41] (http://projects.biotec.tu-dresden.de/metadbsite/) was used for prediction of residues on protein responsible for DNA binding.

Induced-fit DNA docking

The PfRIO2 model was not built from a DNA complex crystal structure template. Therefore, docking a long DNA chain may not be a rational step. Thus, a DNA probe comprising trinucleotide base-pairs (GC, TA, CG) was randomly selected for the present docking studies. The DNA probe was manually edited from the extracted crystal structure of DNA (pdb: 1ZLK) using Maestro. The active site was defined by choosing centroid of the selected basic residues of DB region (Gln64, Lys67, Lys142, and His29). Enzymes are highly dynamic and undergo side chain or backbone movements, or both due to IF ligand binding. Thus, DNA binding in PfRIO2 might also proceed through induced fit mechanism. Therefore, IF docking protocol [42, 43] was used to conduct docking of the built DNA probe to the identified site in PfRIO2 followed by side chain refinement through Prime to allow receptor flexibility according to binding mode. The IF docking strategy uses molecular mechanics-generalized Born surface area (MM-GBSA) scoring function with protein flexibility algorithm and overall docking processed through iteratively approach of docking and protein chain refinement, respectively [31]. Several docked poses were generated and the best receptor-DNA docked complex (with lowest G-score) was selected for final minimization in MacroModel module using OPLS2005 force field to relax and optimize the structure.

Energy calculation by Prime/MM-GBSA simulation and eMBrAcE minimization

The Prime/MM-GBSA simulation was performed on the selected poses DNA-PfRIO2 complex (pose 1 to pose 3) obtained from induced fit molecular docking. The poses were minimized using the local optimization feature in Prime, whereas the energies of complex were calculated with the OPLS2005 force field and Generalized-Born/Surface Area continuum solvent model. During the simulation process, the ligand strain energy was also considered. The Prime/MM-GBSA method [44] was used to calculate the binding-free energy (DG) of DNA, using equation $DG = \Delta E_{MM} + \Delta G_{solv} + \Delta G_{SA}$, where ΔE_{MM} is the difference between the minimized energy of the DNA-PfRIO2complex and the sum of the minimized energies of PfRIO2 and DNA. ΔG_{solv} is the difference in the generalized-Born/surface area (GBSA) solvation energy of the DNA-PfRIO2 complex and the sum of the solvation energies of the PfRIO2 and DNA. Similarly, ΔG_{SA} is the difference in surface area energies for the complex and the sum of the surface area energies for the individual molecules. To consider the final docked poses, the pose energies have been further verified by the eMBrAcE module of Schrödinger [32]. This module has been used to determine the binding energy differences through minimization using OPLS2005 force field. All the calculations were performed with GB/SA continuum water solvation model. During eMBrAcE minimization, residues were allowed to move freely. Minimization calculations were performed for 5000 steps or until the energy difference between subsequent conformations was 0.05 $kJmol^{-1}$. These calculations provided energy difference in terms of VdW, electrostatic and total ($\Delta E = E_{complex} - E_{ligand} - E_{protein}$).

Results and discussion

Multiple sequence alignment

The sequences alignment of RIO2 from *Plasmodium falciparum* (protozoa), *Archaeoglobus fulgidus* (archaea), *Homo sapiens* (Human), *Saccharomyces cerevisiae* (yeast) and protein kinase alfa (from human) was performed to understand the sequence similarity pattern of RIO2 in diverse organisms. The sequence alignment is shown in Fig. 1. The overall sequence similarity of RIO2 from *Plasmodium falciparum* with RIO2 from archaea, human, yeast and protein kinase alfa from human was found to be 51 %, 71 %, 65 % and 25 % respectively. The conserved residue verification of PfRIO2 in comparison to AfRIO2 was done. Glu103, Ser104, Lys120 Asp218, Asn223, Asp235 and Gln238 were found to be conserved in AfRIO2 active site. The similar conserved residues were found in PfRIO2 as Glu104, Ser105, Lys121 Asp225, Asn230, Asp243 and Gln246 which was further confirmed through ATP-binding mode in PfRIO2 model (Fig. 2b). The presence of these conserved residues at PfRIO2 ATP binding site suggests that the two RIO2 kinases follow a similar kind of binding mechanism in both organisms.

Model optimization of PfRIO2

The RMSD-plot (Fig. 3a) of 5 ns MD simulation reveals considerable stability in overall protein, however significant movement was observed in the flexible loop region. The study of generated model for secondary structural elements revealed ten helices (H), 11 strands, three beta (β) sheets (A,B, C), six beta hairpins, 27 beta turns, and seven gamma (γ) turns (Fig. 3b). The major residues that interact with ligand (ATP) and Mn ions are shown as red and blue dots respectively.

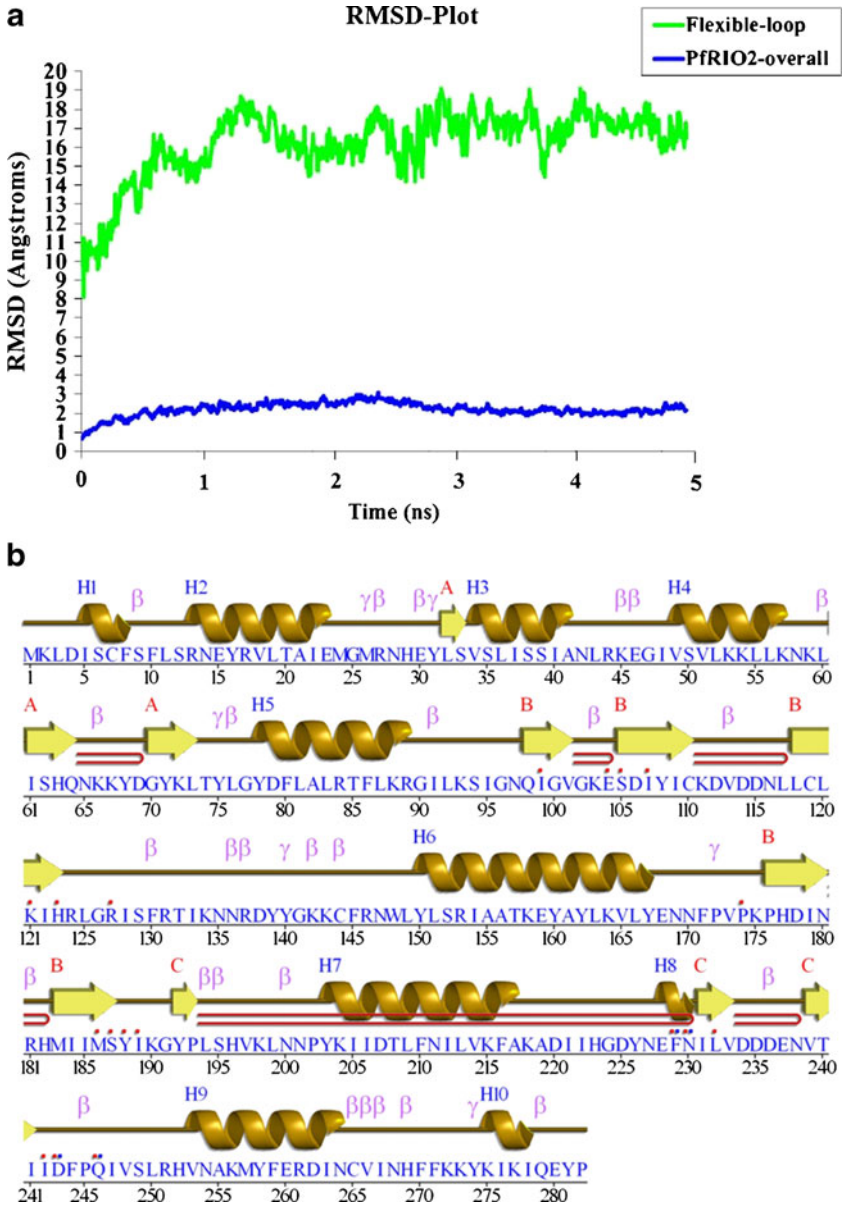

Fig. 3 **a** RMSD vs. simulation time for the 5 ns molecular dynamics simulation of the PfRIO2 model. **b** The secondary structure elements of PfRIO2 showing (helix, H; beta sheet, β; gamma turn, γ) The major residues those interact with ligand (ATP) and Mn+2 ions are shown as *red and blue dots* respectively

Molecular recognition study of ATP binding and serine-phosphorylation in PfRIO2

The superimposition of docked pose (Fig. 4a) of manually built ATP (using maestro interface) on manually corrected ATP (from the crystal structure) shows significant similarity in their binding conformation, which validates the docking protocol to investigate the binding mode analysis.

Based on our docking results, we found that the adenine moiety of ATP is located in the active site via hydrophobic interactions with Pro174, Met186, Tyr188, Ile189, Ile242 and Leu232 (Fig. 4b). ATP catalytic site is comprised of conserved amino acids like Glu104, Ser105, Lys121, Asp225, Asn230, Asp243 Gln246, Met186 and Ile189 (Fig. 4c). The three phosphate groups of ATP attained proper catalytic orientation for phosphate transfer via a bridge like network through interaction with conserved residues and two metal ions (Mn^{2+}). The first metal ion (Mn_1) coordinates with oxygens of α and β phosphate (P_α and P_β) at a distance of 1.9 and 2.0 Å respectively. Also it coordinates with Asp243 (2.0 Å), Asn230 (2.1 Å) and Asp225 (1.9 Å). The conserved residue Lys121 forms a strong electrostatic interaction with oxygen of P_α at 1.4 Å. The Mn_2 orients at the center of phosphate (P_γ), at 1.9 Å distance via the interaction with Gln246 and Asp243 through the coordination at 2.0 Å and 1.9 Å respectively. Further Pγ is stabilized at its location with an electrostatic interaction with Arg127 (1.4 Å) and His123 (1.4 Å). All these interactions stabilize the phosphate backbone and orients Pγ in favorable conformation toward Ser105 at a distance of 4.2 Å. Lys121 was observed in close circumference of

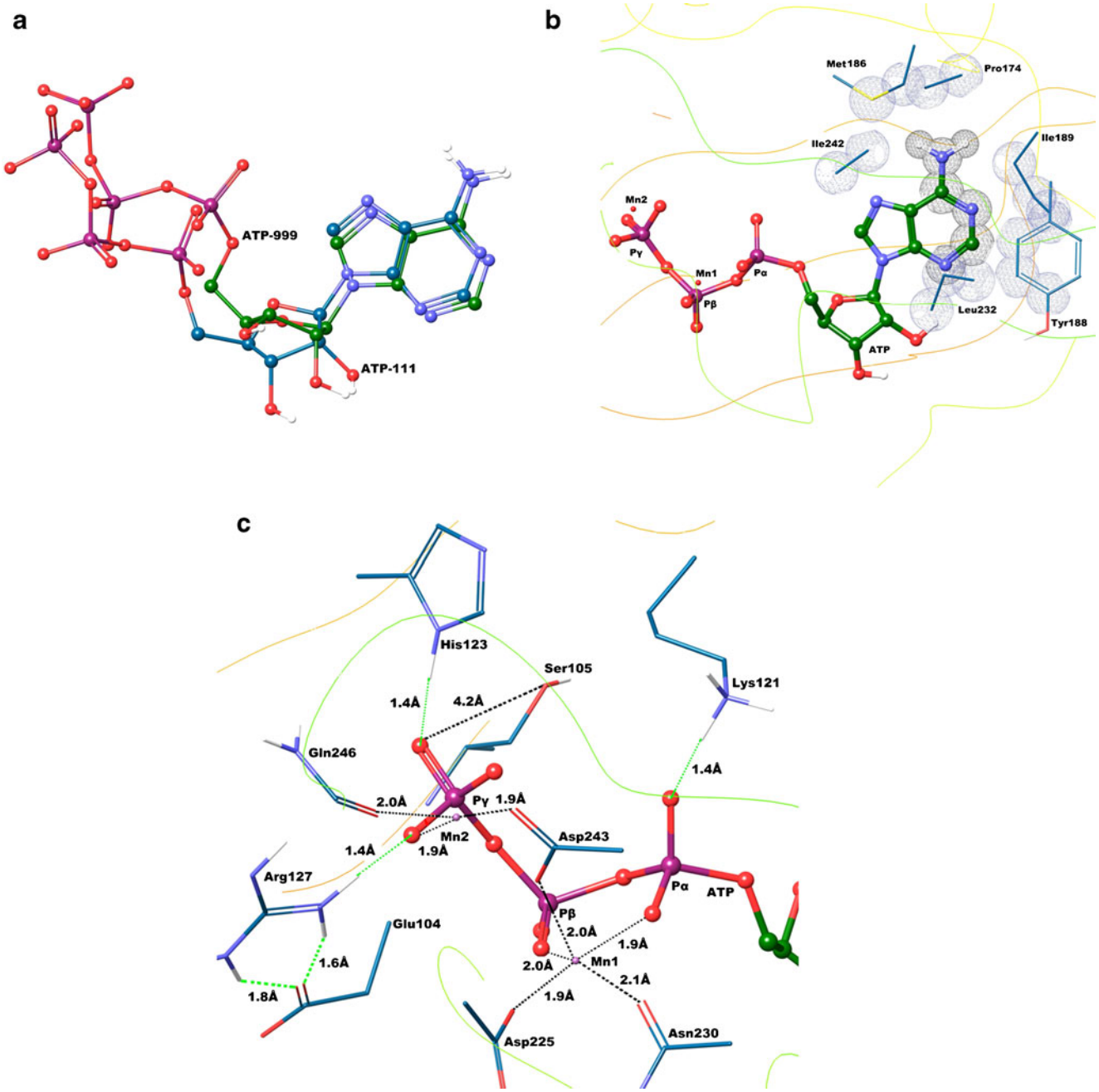

Fig. 4 **a** The superimposition of docked pose of manually built ATP (999: *green*) with ATP (111: *blue*) extracted from crystal structure shows significant similarity in their binding conformation. **b** Hydrophobic interactions of the ATP adenine moiety. Van der Waal surfaces around the purine ring and the surrounding residues are also shown as *mesh* to highlight the interaction profile of the adenine moiety of ATP. **c** Only the triphosphate moiety of ATP is shown for better clarity. A close view of the binding mode of the triphosphate backbone of ATP with the active site residues and two Mn2+ ions (Mn_1 and Mn_2)

Ser105 and Pγ, which enhances the nucleophilicity of hydroxyl group of Ser105 to form bonds with Pγ to conduct auto-phosphorylation. These interaction profiles and structural orientations significantly support Ser105 to be the major conserved residue of PfRIO2 for auto-phosphorylation. However, it opens an avenue to conduct experimental investigation to confirm Ser105 as possible auto-phosphorylation site.

Analysis of ramachandran plot

Ramachandran plot estimates the 3D structural quality of optimized model. It determines the number of residues that are present in the allowed or disallowed regions. Rmachandran plot for AfRIO2 (Fig. 5a) and PfRIO2 (Fig. 5b) shows the presence of 91.1 % and

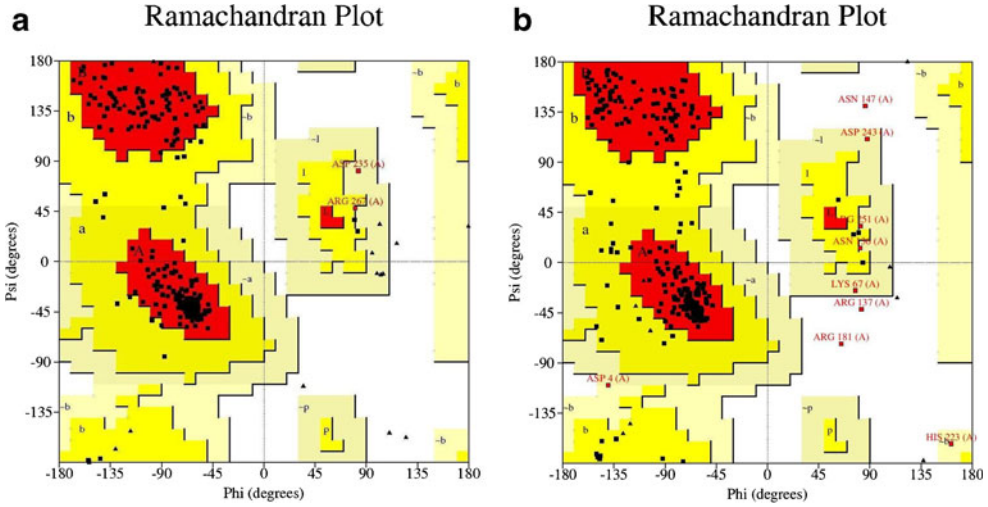

Fig. 5 **a** Ramachandran plot for AfRIO2 model showing 91.1 % residues in most favored region. **b** Ramachandran plot for PfRIO2 model showing 82.4 % residues in most favored region

82.4 % residues in the most favored phi-psi regions respectively.

The lower percentage of residues in most favored region of PfRIO2 ramchandran plot is obvious due to the presence of several missing residues in template sequence. The additionally allowed regions were observed with 8.2 % residues for AfRIO2 and 14.1 % residues for PfRIO2 model. Generously allowed region consist of 0.8 % and 2.3 % residues for AfRIO2 and PfRIO2 model respectively. Disallowed region consist of 0.0 % and 1.1 % residues for AfRIO2 and PfRIO2 model respectively.

DNA binding site identification

The SiteMap analysis revealed "four site points" as plausible ligand binding site (Fig. 6a). The major active sites for ligand binding were observed in wHTH (blue) and flexible loop (green) region. The green location is found closely associated with ATP binding (including auto-phosphorylation site) and wHTH domain (blue). The other two locations (red and violet) are fragmented and significantly smaller in size in comparison to blue and green regions. Thus, a putative DNA binding site has been considered as a combination of blue and green region. The electrostatic charge distribution map (Fig. 6b) was created in PyMol using APBS tool [45]. The positive location around wing-helix & flexible loop region indicated the plausible DNA binding region. Online server of MetaDBSite [41] (http://projects.biotec.tu-dresden.de/metadbsite/) was used for prediction of residues on protein responsible for DNA binding. The analysis reveals that residue from wHTH (Gln64, Lys66, Lys67, Lys88 and Arg89) and flexible loop (Arg137, Tyr139, Lys142, Lys143 and Arg146) have DNA binding properties. Therefore, the results obtained from MetaDBSite strongly support our studies. Further, on the basis of

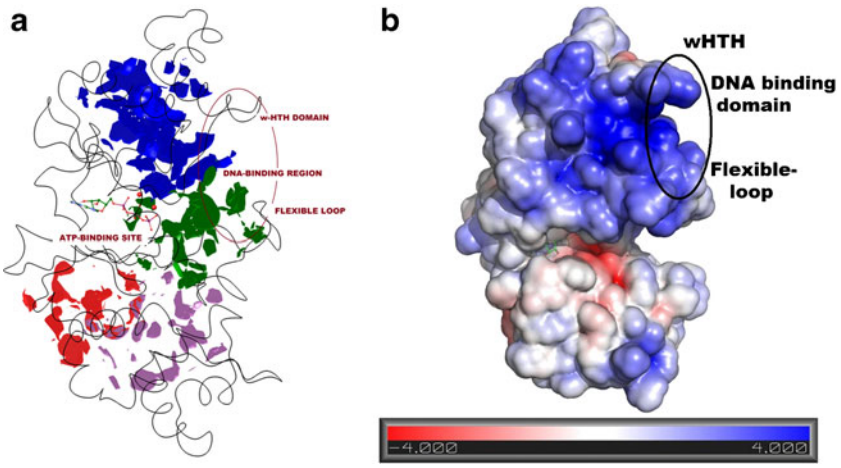

Fig. 6 **a** The site-map analysis of PfRIO2 showing four locations as plausible ligand binding site. The *blue and green* location is situated in wHTH and flexible loop region respectively. The *green* location is found closely associated with ATP binding (including auto-phosphorylation site) and wHTH domain. The other two locations (*red and violet*) are fragmented and are significantly smaller in size in comparison to *blue and green* regions. Thus, a putative DNA binding site has been considered as combination of *blue and green region*. **b** The electrostatic surface shows the major positive region situated near wHTH and flexible loop region, which supports the site-map analysis to identify DNA binding region

reported domain studies of RIO2 kinase it has been suggested that the possible DNA binding site could be close to the wHTH domain region [27]. Thus, the intersection of blue and green regions adjacent to ATP binding site was identified as a putative DNA binding region. These types of domains have been reported in DNA binding proteins having two different modes of binding [19, 20].

The first mode is reported in the crystal structure of transcription factor HNF-3 bound to DNA, where site-specific interactions of DNA major groove plays important role toward its binding to the protein [46]. The second mode reported in the crystal structure of hRFX1 bound to DNA, where interactions between W1, β2, and β3 with the major groove of the DNA are responsible for DNA positioning in protein [47]. In continuation, the other experimental evidence indicated that RIO2 may bind non-specifically to single-stranded oligonucleotides and also suggest the binding potential with oligonucleotides [27]. The sequence analysis of RIO2 shows the presence of charged amino acids in helix 4 of the wHTH domain and basic residues in flexible loop region, which supports its possible DNA binding ability. Also the result of loop refinement using Prime suggests the location of a large loop closed to the wHTH domain. Thus, PfRIO2 kinase may functions by induced fit mechanism involving flexible loop in association with wHTH region to conduct DNA binding. Further, the identified blue and green SiteMap regions are adjacent to ATP binding site, which supports the hypothesis to consider the intersection of these two regions as a putative DNA binding (DB) region. Thus, we believe that intersection associated with wHTH domain and flexible loop will assist in DNA binding, which is indirectly supported by experimental crystal structure, where the electron density of flexible loop is absent due to its high dynamic nature. Further, we can hypothesize that the dynamic nature of flexible loop is responsible for autophosphorylation followed by conformational change to interact with wing-region for completing the DNA binding events.

Induced fit docking

There is no native ligand or DNA reported in any experimental RIO2 kinase structure and this is the first report where we identified a possible DNA binding region in PfRIO2. The earlier report also supports the possible DNA location near helix 3 in AfRIO2 [27]. Further, the built model did not contain DNA and was also not built from a DNA complex crystal structure template, which put a major challenge to study the DNA binding mode. The rigid receptor docking methods may not represent the real binding process, because enzymes are in dynamic motion and undergo side chain or backbone movements, or both due to induced fit ligand binding. Thus, it was necessary to allow flexibility to DNA binding region of protein for DNA docking. Therefore, docking of the DNA to the identified SiteMap regions (intersection of blue and green) was conducted using IF docking.

DNA-PfRIO2 binding energetic

The top three induced fit docking (IFD) poses of DNA docking were investigated in terms of binding mode and binding energies. The energy values are shown in Table 1 for all three poses.

The poses from IF docking were ranked on the basis of their G-score and pose 1 showed significantly lower energy state (−8.7 kcalmol^{-1}) in comparison to pose 2 and 3. The poses were further rescored by Prime and eMBrAcE (multi-ligand bimolecular association with energetics) module of Schrödinger [32]. The rescoring using MMGBSA free energy of binding, which includes DNA-PfRIO2 interaction energy (DG: Table 1) shows relatively better profile of

Table 1 The molecular docking score and binding energy calculation (kcalmol^{-1}) of short DNA into the PfRIO2 kinase by induced-fit method followed by Prime-MMGBSA and eMBrAcE minimization

S. no	DNA-PfRIO2 complex pose	Induce fit docking[a]	MM-GBSA score[b]	eMBrAcE score[c] energy difference		
		G-score	DG	VdW	Electrostatic	Total (ΔE)
1	Pose-1	−8.7	−77.1	−17.4	−1436.2	−52.4
2	Pose-2	−7.8	−75.1	−9.9	−1008.8	−33.1
3	Pose-3	−7.3	−54.5	−13.4	−1420.9	−25.2

[a] Induce fit docking score: G-score (Glide score): The minimized poses are rescored using Schrödinger's proprietary GlideScore scoring function and the binding affinity can be estimated by G-score

[b] MM-GBSA score: DG = E_complex(minimized) − [E_ligand(from minimized complex) + E_receptor]; DG is the ligand/receptor interaction energy of the complex

[c] eMBrAcE (multi-ligand bimolecular association with energetics): ($\Delta E_{Total} = E_{complex} - E_{ligand} - E_{protein}$); VdW: van der Waals interaction. [b] ΔE = Energy difference

−77.1 kcalmol^{-1} for pose 1. The MM-GBSA scoring is based on the difference between the energy of the DNA bound complex and sum of the energy of unbound protein and DNA [$E_{binding} = E_{complex} - (E_{protein} + E_{DNA})$]. The eMBrAcE module significantly supports pose 1 with three major scoring functions; VdW energy, electrostatic energy and overall total energy difference. The VdW energy difference of −17.4 kcalmol^{-1} (for DNA binding) was excellent in comparison to −9.9 and −13.4 kcalmol^{-1} for pose 2 and 3 respectively. The higher total energy difference of −52.4 kcalmol^{-1} for pose 1 supports it as the possible DNA binding structure of PfRIO2. Thus, pose-1 has been finally considered as DNA-PfRIO2 kinase complex (Fig. 7a).

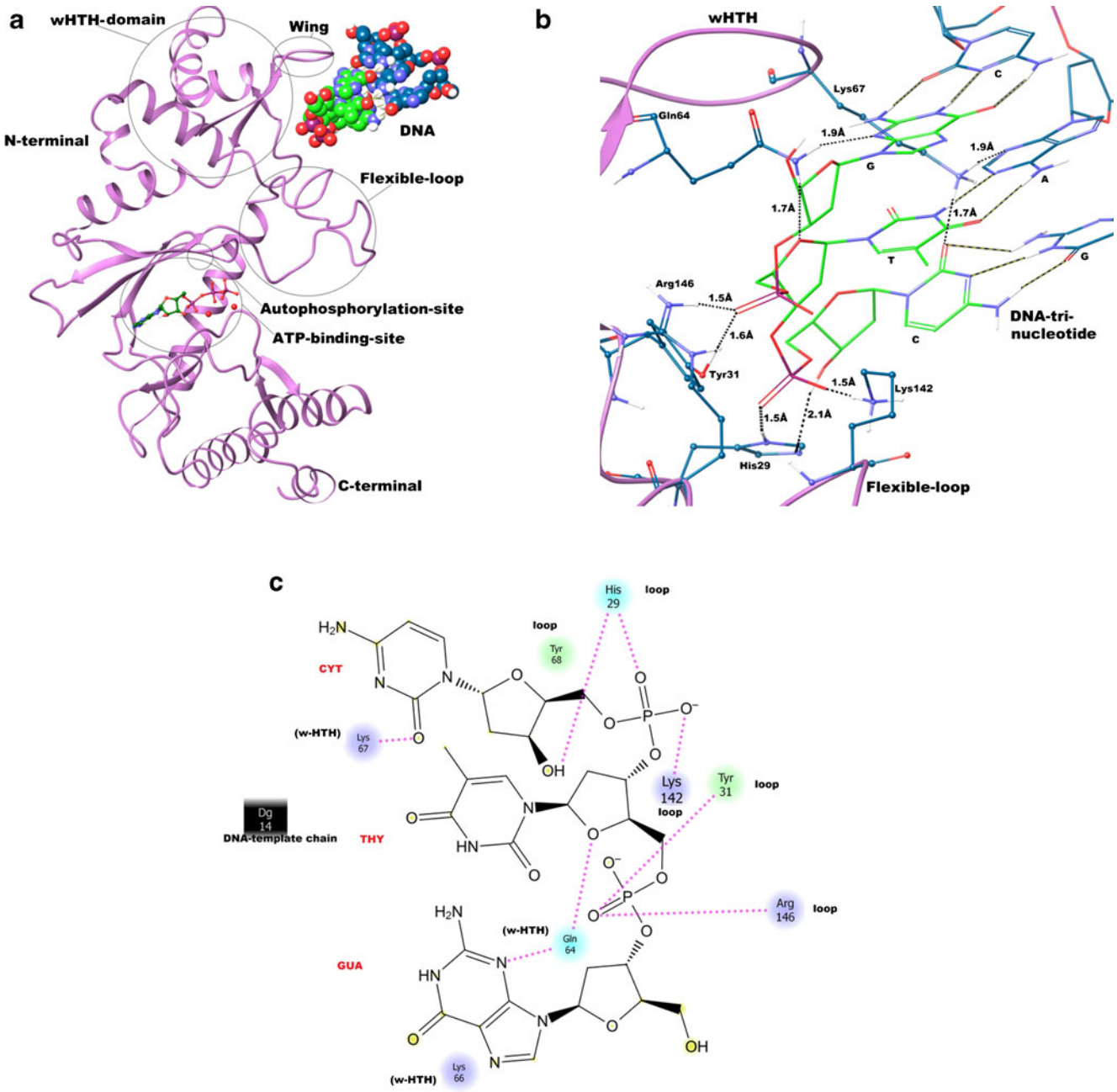

Fig. 7 **a** The best docked pose obtained after induced fit docking of short-DNA (CPK representation of DNA-trinucleotide chain) into the DNA binding groove. The DNA-trinucleotide accommodated within a groove between wHTH domain and flexible loop region. Overall the figure demonstrates DNA binding region, ATP binding and phosphorylation site. **b** The final selected pose of DNA-PfRIO2 showing residues from wHTH domain (Gln64 and Lys67) and residues from flexible loop (His29, Tyr31, Lys142 and Arg146) primarily participating in interaction with DNA chain. **c** A closer view of DNA-binding interaction with wHTH domain and flexible loop residues. The major basic residues interacting through H-bonding (*dotted pink line*) with DNA are shown in *blue shade with labeling*

DNA binding mode analysis

The binding mode obtained from docked pose 1 (Fig. 7b) reveals proper DNA base pairing, H-bonding interactions with wHTH region (Gln64 and Lys67) and loop region (Lys142, His29, Arg146 and Tyr 31), which validates the pose selection to demonstrate the DNA binding in PfRIO2 kinase. His29 and Tyr31 form H-bonds with nucleotide phosphate at 1.5 Å and 1.6 Å respectively. His29 forms an additional H-bond with the hydroxyl group of ribose sugar of cytosine at 2.1 Å. Arg142 and Lys146 from flexible loop strengthen this interaction via H-bonds with phosphate oxygen at 1.5 Å each. The DNA binding interactions are further stabilized by H-bonding interactions of Gln64 with guanine at 1.9 Å and with sugar ring oxygen of thymine at 1.7 Å. Similarly Lys67 from wing part of wHTH domain forms H-bond at 1.7 and 1.9 Å with cytosine and adenine respectively. Thus, overall DNA binding complex is stabilized by all these non-covalent interactions. A closer view of DNA-binding interaction with wHTH domain and flexible loop residues is shown in Fig. 7c.

Conclusions

In summary, our in-depth structural analysis enlightens the salient biochemical processes involving 3D structural implication toward PfRIO2 recognition phenomenon involving ATP-binding, auto-phosphorylation, and DNA interactions. The present structural exploration suggests a novel putative DNA-binding site in PfRIO2 kinase to understand the DNA binding events involving wHTH domain and flexible loop. The formation of a stable PfRIO2-ATP-DNA complex enhances hope for exploiting this novel allosteric site of PfRIO2 as a potential anti-malarial drug target. Thus the structural facets of unknown PfRIO2 complex were investigated by multiple modeling-simulation approaches, which opens an avenue toward exploration of new drug target.

Acknowledgments This research received funding from Department of Biotechnology (DBT), New Delhi, India through grant no BT/41/NE/TBP/2010. Authors' thanks to DBT (BT/PR14237/MED/29/196/2010) for providing Schrödinger Software. DKC is thankful to DBT for the Junior Research Fellowship. Authors are thankful to Dr. Robert J. Woods, Complex Carbohydrate Research Centre, University of Georgia, Athens, USA for critical reading and technical comments, who is also an external Ph. D. supervisor of DKC. We thank Dr. Vishal Trivedi, IIT Guwahati, for helpful discussions on biochemical function of RIO2 Kinase.

References

1. Noedl H, Se Y, Schaecher K, Smith BL, Socheat D, Fukuda MM (2008) Evidence of artemisinin-resistant malaria in western Cambodia. N Engl J Med 359:2619–2620
2. Noedl H, Socheat D, Satimai W (2009) Artemisinin-resistant malaria in Asia. N Engl J Med 361:540–541
3. Kumar S, Bandyopadhyay U (2005) Free heme toxicity and its detoxification systems in human. Toxicol Lett 157:175–188
4. Becker K, Tilley L, Vennerstrom JL, Roberts D, Rogerson S, Ginsburg H (2004) Oxidative stress in malaria parasite-infected erythrocytes: host–parasite interactions. Int J Parasitol 34:163–189
5. Gromer S, Urig S, Becker K (2004) The thioredoxin system—from science to clinic. Med Res Rev 24:40–89
6. Müller S (2004) Redox and antioxidant systems of the malaria parasite Plasmodium falciparum. Mol Microbiol 53:1291–1305
7. Ramya T, Surolia N, Surolia A (2002) Survival strategies of the malarial parasite Plasmodium falciparum. Curr Sci 83:818–825
8. Nogueira F, Diez A, Radfar A, Pérez-Benavente S, Rosario VE, Puyet A, Bautista JM (2010) Early transcriptional response to chloroquine of the Plasmodium falciparum antioxidant defence in sensitive and resistant clones. Acta Trop 114:109–115
9. Vega-Rodríguez J, Franke-Fayard B, Dinglasan RR, Janse CJ, Pastrana-Mena R, Waters AP, Coppens I, Rodríguez-Orengo JF, Jacobs-Lorena M, Serrano AE (2009) The glutathione biosynthetic pathway of Plasmodium is essential for mosquito transmission. PLoS Pathog 5:e1000302
10. Ginsburg H, Stein W (2004) The new permeability pathways induced by the malaria parasite in the membrane of the infected erythrocyte: comparison of results using different experimental techniques. J Membr Biol 197:113–134
11. Bruchhaus I, Roeder T, Rennenberg A, Heussler VT (2007) Protozoan parasites: programmed cell death as a mechanism of parasitism. Trends Parasitol 23:376–383
12. Vaidya AB, Mather MW (2009) Mitochondrial evolution and functions in malaria parasites. Annu Rev Microbiol 63:249–267
13. Dhangadamajhi G, Kar SK, Ranjit M (2010) The survival strategies of malaria parasite in the red blood cell and host cell polymorphisms. Malar Res Treat 2010:973094–973103
14. Doerig C, Billker O, Haystead T, Sharma P, Tobin AB, Waters NC (2008) Protein kinases of malaria parasites: an update. Trends Parasitol 24:570–577
15. Bozdech Z, Zhu J, Joachimiak MP, Cohen FE, Pulliam B, DeRisi JL (2003) Expression profiling of the schizont and trophozoite stages of Plasmodium falciparum with a long-oligonucleotide microarray. Genome Biol 4:R9
16. Ward P, Equinet L, Packer J, Doerig C (2004) Protein kinases of the human malaria parasite Plasmodium falciparum: the kinome of a divergent eukaryote. BMC Genomics 5:79
17. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. Science 298:1912–1934
18. Angermayr M, Bandlow W (2002) RIO1, an extraordinary novel protein kinase. FEBS Lett 524:31–36
19. Gajiwala KS, Burley SK (2000) Winged helix proteins. Curr Opin Struct Biol 10:110–116
20. Wolberger C, Campbell R (2000) New perch for the winged helix. Nat Struct Mol Biol 7:261–262
21. Dong G, Chakshusmathi G, Wolin SL, Reinisch KM (2004) Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. EMBO J 23:1000–1007
22. Geerlings TH, Faber AW, Bister MD, Vos JC, Raue HA (2003) Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20S Pre-rRNA in Saccharomyces cerevisiae. J Biol Chem 278:22537–22545
23. LaRonde-LeBlanc N, Wlodawer A (2005) The RIO kinases: an atypical protein kinase family required for ribosome biogenesis and cell cycle progression. Biochim Biophys Acta 1754:14–24
24. Campbell BE, Boag PR, Hofmann A, Cantacessi C, Wang CK, Taylor P, Hu M, Sindhu ZU, Loukas A, Sternberg PW, Gasser RB (2011) Atypical (RIO) protein kinases from Haemonchus contortus

—promise as new targets for nematocidal drugs. Biotechnol Adv 29:338–350
25. Campbell BE, Hofmann A, McCluskey A, Gasser RB (2011) Serine/threonine phosphatases in socioeconomically important parasitic nematodes—prospects as novel drug targets? Biotechnol Adv 29:28–39
26. Canduri F, Perez PC, Caceres RA, de Azevedo WF Jr (2007) Protein kinases as targets for antiparasitic chemotherapy drugs. Curr Drug Targets 8:389–398
27. LaRonde-LeBlanc N, Wlodawer A (2004) Crystal structure of A. fulgidus Rio2 defines a new family of serine protein kinases. Structure 12:1585–1594
28. Schrödinger LLC (2011) Schrödinger LLC, New York, NY
29. Prime. V 3.0 (2011) Schrödinger, LLC, New York, NY, V 3.0, Schrödinger, LLC, New York, NY
30. Desmond-V-3.0 (2010) Molecular Dynamics System, V 3.0, D. E. Shaw Research, Schrödinger, LLC, New York, NY
31. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R (2006) Novel procedure for modeling ligand/receptor induced fit effects. J Med Chem 49:534–553
32. MacroModel-V-9.9 (2011) Schrödinger LLC, New York, NY
33. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 419:498–511
34. Glide-V-5.7. (2011) V 2.5, Schrödinger, LLC, New York, NY
35. Hoover WG (1985) Canonical dynamics: equilibrium phase-space distributions. Phys Rev A 31:1695–1697
36. Tobias D, Martyna G, Klein M (1994) Constant pressure molecular dynamics algorithms. J Chem Phys 101:4177
37. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8593
38. SiteMap-V-2.5 (2011) V 2.5, Schrödinger, LLC, New York, NY
39. Halgren TA (2009) Identifying and characterizing binding sites and assessing druggability. J Chem Inf Model 49:377–389
40. Nayal M, Honig B (2006) On the nature of cavities on protein surfaces: application to the identification of drug-binding sites. Proteins 63:892–906
41. Si J, Zhang Z, Lin B, Schroeder M, Huang B (2011) MetaDBSite: a meta approach to improve protein DNA-binding sites prediction. BMC Syst Biol 5(Suppl 1):S7
42. Sherman W, Beard HS, Farid R (2006) Use of an induced fit receptor structure in virtual screening. Chem Biol Drug Des 67:83–84
43. Kim YA, Sharon A, Chu CK, Rais RH, Al Safarjalani ON, Naguib FNM, el Kouni MH (2007) Synthesis, biological evaluation and molecular modeling studies of N6-benzyladenosine analogues as potential anti-toxoplasma agents. Biochem Pharmacol 73:1558–1572
44. Lyne PD, Lamb ML, Saeh JC (2006) Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. J Med Chem 49:4805–4808
45. DeLano W (2002) The PyMOL molecular graphics system. DeLano Scientific, Palo Alto
46. Clark KL, Halay ED, Lai E, Burley SK (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5
47. Gajiwala KS, Chen H, Cornille F, Roques BP, Reith W, Mach B, Burley SK (2000) Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. Nature 403:916–921