Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 22 (2004) 241–247
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Characteristics of the *Plasmodium falciparum* PK5 ATP-binding site: implications for the design of novel antimalarial agents

Susan M. Keenan, William J. Welsh*

Department of Pharmacology, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, NJ 08854, USA

Received 11 September 2003; received in revised form 11 September 2003; accepted 12 September 2003

Abstract

Increasing worldwide resistance of *Plasmodium falciparum* (*P. falciparum*) to traditional chemotherapy strategies such as chloroquine and mefloquine demonstrates the urgent need for the discovery of novel chemotherapeutic agents in the fight against malaria. The recent discovery of *P. falciparum* Protein Kinase 5 (P*f*PK5) invites the possibility of selectively targeting the life cycle of *P. falciparum* in order to prevent cerebral malaria. P*f*PK5 bears a high degree of sequence identity (>58%) to a structurally conserved family of mammalian kinases known as the cyclin-dependent kinases (CDKs). The CDKs are the key regulatory elements governing the ordered progression of the mammalian cell cycle. With numerous X-ray crystal structures of CDK2 to provide a structural template, here we present a three-dimensional structural model of P*f*PK5 constructed using computer-based homology modeling techniques. Our model was used to compare the ATP binding site of P*f*PK5 with that of the mammalian kinase CDK2. Furthermore, kinase–ligand interactions of P*f*PK5 with known inhibitors were investigated and compared to available crystal structures of CDK2 with inhibitors bound. The focus of the study is to identify similarities and differences between the ATP binding sites of the two kinases that can be exploited for future rational drug design.

Keywords: Malaria; *P. falciparum* PK 5 (P*f*PK5); CDK; Drug design; Inhibitors

1. Introduction

There are four parasitic species that are the causative agents of human malaria; namely *Plasmodium falciparum*, *P. vivax*, *P. ovale*, and *P. malariae* [1]. While *P. vivax* and *P. falciparum* are the most widely distributed species, *P. falciparum* alone is responsible for the most severe infections and for over 95% of 1–2 million worldwide deaths resulting annually from malarial infection [1]. As a global threat, malaria is making a dramatic comeback and, over the past 20 years, efforts to control the malaria parasites have met with decreasing success as these parasites are developing resistance to the drugs traditionally used to control this deadly disease [2].

While the stages of parasitic infection are well documented, until recently, little has been known concerning the regulation of the parasitic life cycle. However, the identification of a family of kinases from the *P. falciparum* with a high degree of sequence conservation to the mammalian CDKs has allowed researchers to begin investigation of the mechanisms that control passage through the parasitic life cycle. *P. falciparum* Protein Kinase 5, P*f*PK5, was isolated in 1994 using oligonucleotides based on the sequences of CDK2 and CDK1. Sequence analysis demonstrated ∼58% identity and >65% sequence conservation with the CDKs and, once expressed in *E. Coli*, P*f*PK5 kinase activity could be observed against casein and histone H1 [3], and later the carboxyl-terminal domain (CTD) of the large subunit of RNA polymerase II [4]. Interestingly, expression of P*f*PK5 is observed constantly throughout the intra-erythrocytic stage of the parasitic life cycle indicating a possible role for P*f*PK5 in the regulation of this stage of asexual reproduction [3,5,6]. More importantly, PfPK5 activity peaks about 36 h post infection during the schizont stage [7].

In addition to the observed sequence conservation, evidence suggests that P*f*PK5 is activated in a manner similar to the activation of the mammalian CDKs. First, research into the mechanism of P*f*PK5 activation demonstrates the necessity of Thr158 phosphorylation of P*f*PK5 for activation [7]. Similar phosphorylation events have been shown to be necessary for the activation of the mammalian CDKs. For example, activation of CDK2 requires phosphorylation of Thr160 [8]. Second, the in vitro activation of P*f*PK5 is enhanced as a result of co-incubation with the human cyclins [9]. Full activation of the mammalian CDKs requires the association of the catalytic kinases with a regulatory cyclin subunit [10]. Evaluation of the activational effects of cyclin H and p25,

* Corresponding author. Tel.: +1-732-235-3234; fax: +1-732-235-3475.
E-mail address: welshwj@umdnj.edu (W.J. Welsh).

1093-3263/$ – see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2003.09.002

two mammalian cyclins, on P*f*PK5 activity has led to the isolation of Pfcyc-1, a protein able to activate P*f*PK5 in vitro [9] Taken together, these data strongly suggest a role for P*f*PK5 in the life cycle of *P. falciparum* which is similar to that observed for CDK2 in the mammalian cell cycle.

The similarity between P*f*PK5 and the mammalian CDKs in terms of their mechanism of kinase activation and inhibition has led researchers to investigate the effects of known small molecule CDK inhibitors on the growth of *P. falciparum*. Harmse et al. [11] analyzed the effects of a series of purine analogues on *P. falciparum* growth and CDK1 activity. While some compounds demonstrated similar IC_{50} values against *P. falciparum* and CDK1, surprisingly, the inhibitory effects of other compounds differed markedly. Furthermore, pre-incubation of *P. falciparum* with flavopiradol and olomoucine, two compounds known to inhibit the mammalian CDKs, resulted in the inhibition of *P. falciparum* DNA synthesis [12]. More recently, Le Roach et al. determined the IC_{50} values of P*f*PK5 kinase inhibition by hymenialdisine (HYM), indirubin-3′-monoxime (I3M) and purvalanol A (PURA). The concentrations required to inhibit P*f*PK5 were higher (1, 10 and 8 μM, respectively, for HYM, I3M and PURA) than the IC_{50} values determined for CDK2 inhibition (70, 440 and 70 nM) [9].

In the present study, we have developed a three-dimensional model of P*f*PK5 by applying computational homology modeling techniques and utilizing the high-resolution crystal structure of the T160 phosphorylated CDK2 associated with cyclin A as the template (1JST). We then docked HYM, I3M and PURA to P*f*PK5, and compared the consensus binding modes obtained to the interaction orientations observed for the same inhibitors (or analogs) complexed with CDK2 in their respective X-ray crystal structures. Furthermore, we compared structural features within the ATP-binding sites of P*f*PK5 and CDK2 with the goal of providing the foundation for the development of pharmacophore models. By defining the structural and spatial requisites for molecular recognition and binding of small molecules to P*f*PK5, these pharmacophores will guide the rational de novo design of a new generation of therapeutic agents for the treatment of cerebral malaria.

2. Materials and methods

2.1. Sequence alignment

The sequence of the P*f*PK5 was obtained from GenBank (Accession No. CAA4393) [13]. The CDK2 (1JST) was obtained from RCSB-PDB and the sequence was extracted using Insight II (Accelrys Inc., San Diego, CA). The sequence alignment between P*f*PK5 and CDK2 was deduced using the identity matrix, and the resulting alignment was further refined manually. The gaps resulting from the alignment are localized to loop regions and, thus, were easily accommodated. Both the template and model was analyzed using the *What If* Server (http://www.cmbi.kun.nl/gv/servers/WIWWWI/) [14].

2.2. Homology modeling

All molecular modeling operations were carried out using InsightII and related modules (Accelrys Inc., San Diego, CA). The 3D coordinates of P*f*PK5 were assigned, and undefined loops and gaps in the aligned sequences were accommodated by generation of intervening sequences while conserving the integrity of backbone orientations in structurally defined, flanking residues as determined by the alignment. Inclusion of the randomly modeled regions required splice-site relaxation in order to allow the protein backbone to adopt appropriate psi/phi angles. This operation was completed by relaxing the backbone atoms of those residues involved in splice sites and subjecting the molecule to a localized energy minimization procedure. ATP was extracted from the CDK2 structure (1JST) and superimposed into the putative ATP-binding site of P*f*PK5. The side chain residues of the *Pf*PK5 model were checked for acceptable rotomer positioning using the internal rotomer library, and steric clashes were resolved by cursory energy minimization of the entire molecule (250 iterations, steepest descent). Finally, the model was subjected to a full energy minimization through a series of independent steps in which the ATP and the macromolecule hydrogen atoms, side chains, backbone atoms, and ultimately, the entire molecule were successively allowed freedom of movement. This procedure implemented the CVFF force field with a distance dependent dielectric function ($\varepsilon = \varepsilon_o r$, with $\varepsilon_o = 4$) until the convergence criterion of 0.04184 kJ/mol (0.01 kcal/mol) change in energy between successive iterations was achieved. The dielectric constant $\varepsilon_o = 4$ was chosen as it corresponds to that of water in a fast-moving dynamic field. Later, the consensus orientations of HYM, I3M and PURA obtained from each docking study (described below) were superimposed into the ATP binding site of P*f*PK5. Again steric clashes were resolved by cursory energy minimization of the entire molecule (250 iterations, steepest descent), then each ligand-kinase model was subjected to a full energy minimization as described above.

2.3. GOLD docking

ATP was extracted from the binding site of the energy minimized P*f*PK5-ATP structure, and the resulting model was utilized for the docking analysis. Genetic optimization for ligand docking (GOLD) is a ligand-docking application that utilizes a genetic algorithm (GA) to explore ligand conformation and satisfy ligand-binding requirements [15]. Unlike deterministic algorithms, the stochastic nature of GOLD ensures that the search space is well explored and local energy extremes are less influential. Probabilities for crossover, mutation, and migration were set to 95, 95, and 10, respectively. The maximum distance between hydrogen bond donors and fitting points was set to 5 Å, and the nonbonded

cutoff for the van der Waals energies was adjusted to the manufacturer recommended settings ($10k_{ij}$, well depth of the van der Waals energy of the atom pair *i*, *j*). The ligand orientation accepted was the top scoring consensus orientation of 25 independent GA runs, each with a maximum number of 1000 GA operations performed on a single population of 100 individuals. The selection pressure was 1.1.

3. Results and discussion

As the effects of traditional chemotherapy agents decrease as a result of growing parasitic resistance, the global threat of malaria remains a cause for concern. *P. falciparum* resistance to mainstay chemotherapeutic agents such as fansidar and chloroquine has developed worldwide, creating dire social and health implications for developing countries [16]. Drug resistance of this scale requires scientists to identify new mechanisms by which to inhibit *P. falciparum* and thus prevent the spread of malaria. P*f*PK5 is identified as a protein involved in the regulation of the *P. falciparum* life cycle and, as such, the targeting of P*f*PK5 activity provides a means of inhibiting the asexual reproduction of *P. falciparum* and reveals a novel target for the design of antimalarial agents.

As P*f*PK5 appears to be activated and inhibited in a manner similar to that of CDK2, we chose CDK2 as the preferred template for our P*f*PK5 homology model. The P*f*PK5 and CDK2 sequences were aligned based on identity, as the two proteins in question show significant sequence conservation (identity equals ∼68%, similarity equals ∼72%). The alignment scheme results in the generation of two single-residue gaps in the P*f*PK5 sequence which correspond to Thr28 and Leu96 in CDK2 and were easily accommodated in the P*f*PK5 model (open boxes, Fig. 1). Examination of the CDK2 and P*f*PK5 aligned sequences revealed good conservation of regions of high local sequence identity and functionally important motifs, specifically: (1) the PSTAIRE (single amino acid representation) motif, which is important for the binding of cyclin and characteristic of all cyclin dependent kinases [17]; (2) the activational domain, which spans from the DEG motif (D145) to the APE motif (E172) in CDK2 providing a deep cleft for ATP binding [18,19]; and (3) residues 80–84 in CDK2, which have been shown to be important in the binding of both ATP and CDK inhibitors (Fig. 1) [20]. Furthermore, the residues essential for stabilization of the structural changes associated with CDK2 [13] activation (Lys33, Arg50, Arg126, and Arg159) are conserved in P*f*PK5. We proceeded to assign the structural coordinates of the P*f*PK5 taking care to generate loops that were acceptable sterically and that gave reasonable phi-psi angles for the spliced regions. A ribbon representation of the final model is shown in Fig. 1B.

For model validation, we extracted ATP from the CDK2 crystal structure and superimposed the ATP molecule into the putative P*f*PK5 binding site. Next, we subjected the ATP-P*f*PK5 structure to a series of energy minimization steps (described in detail in the Methods section). Subsequent to the refinement of the model, we were able to identify key interactions between P*f*PK5 and the purine group of ATP that act to stabilize the interaction; indeed six potential hydrogen bonds were easily identified between ATP and residues Glu80, Leu83, Asp85 and Gln129 (Fig. 2A). Additionally, there are at least three potential hydrogen bonds stabilizing the interactions of the phosphate groups (Table 1). From the wealth of available 3D structural data available for CDK2, the interactions important for ATP binding are clear (reviewed in Gray et al. [20]). Key interactions occur between the purine rings of ATP and residues Glu81, Leu83, Asp86 and Gln131 of CDK2. The residues associated with the interactions are identical to those we identified between ATP and P*f*PK5 when one takes into account the alignment of P*f*PK5 and CDK2 (see Fig. 1). The consistency of the ligand-macromolecular interactions for P*f*PK5 and CDK2 suggests that our model can provide insights into binding modes of small molecule inhibitors to the parasitic kinase.

Table 1
Residues involved in ligand binding

	P*f*PK5	CDK2
ATP	Glu80	Glu81
	Leu82	Leu83
	Asp85	Asp86
	Gln129	Gln131
(Phosphate Groups)	Thr14	Thr14
	Lys32	Lys33
	Lys127	Lys129
	Asn130	Asn132
	Asp143	Asp 145
HYM	Glu80	Glu81
	Leu82 (2)	Leu83 (2)
	Asp143	Asp145
I3M	Glu80	Glu81
	Leu82 (2)	Glu83
	Gln129	**Ile10**
	Asn130	
PURA	Glu80	Glu81
	Leu82 (2)	Leu83 (2)
	Asp85	Asp86

Residues in bold differ between kinases. Numbers in parentheses indicated more than one bond associated with the residue.

The crystal structures of CDK2 complexed with a vast number of purine analogues have been solved. The general conclusion from these experiments is that while no one ligand orientation predominates, consistent trends appear with a number of amino acid residues of CDK2 (e.g. Glu81, Leu83 and Asp86) playing essential roles. We proceeded to examine the association between PfPK5 and hymenialdisine (HYM), indirubin-3′-monoxime (I3M) and purvalanol A (PURA) for two reasons: (1) these compounds have been shown to directly inhibit P*f*PK5 activity, [9] and (2) these compounds (or analogues) have been crystallized with CDK2 thus enabling a direct comparison between structures

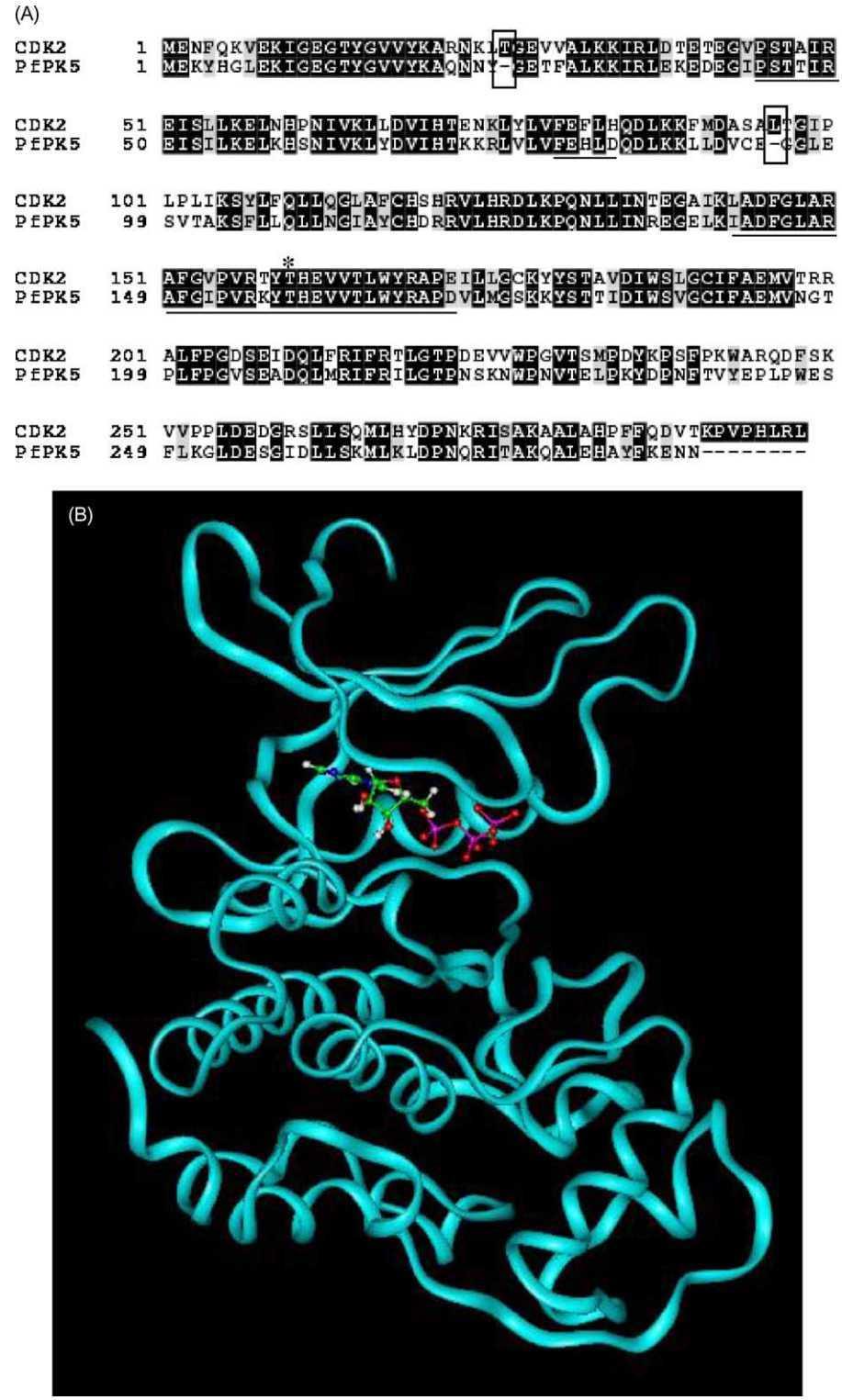

Fig. 1. (A) Sequence alignment of P*f*PK5and CDK2. Global alignment of CDK2 and P*f*PK5 using a sequence identity matrix. Residues that are identical between the two sequences are shaded with a black background, while similar residues are shaded with a gray background. The gaps are indicated with (—) and located in boxes. The symbol (∗) denotes the Thr160 phosphorylation site. The PSTAIRE region, CDK2 residues 80–84 and the activation domain are underlined. (B) Representation of P*f*PK5. Visual representation of the three-dimensional structure as obtained by homology modeling. Alpha carbons are depicted by a shaded ribbon and ATP fills the binding site.

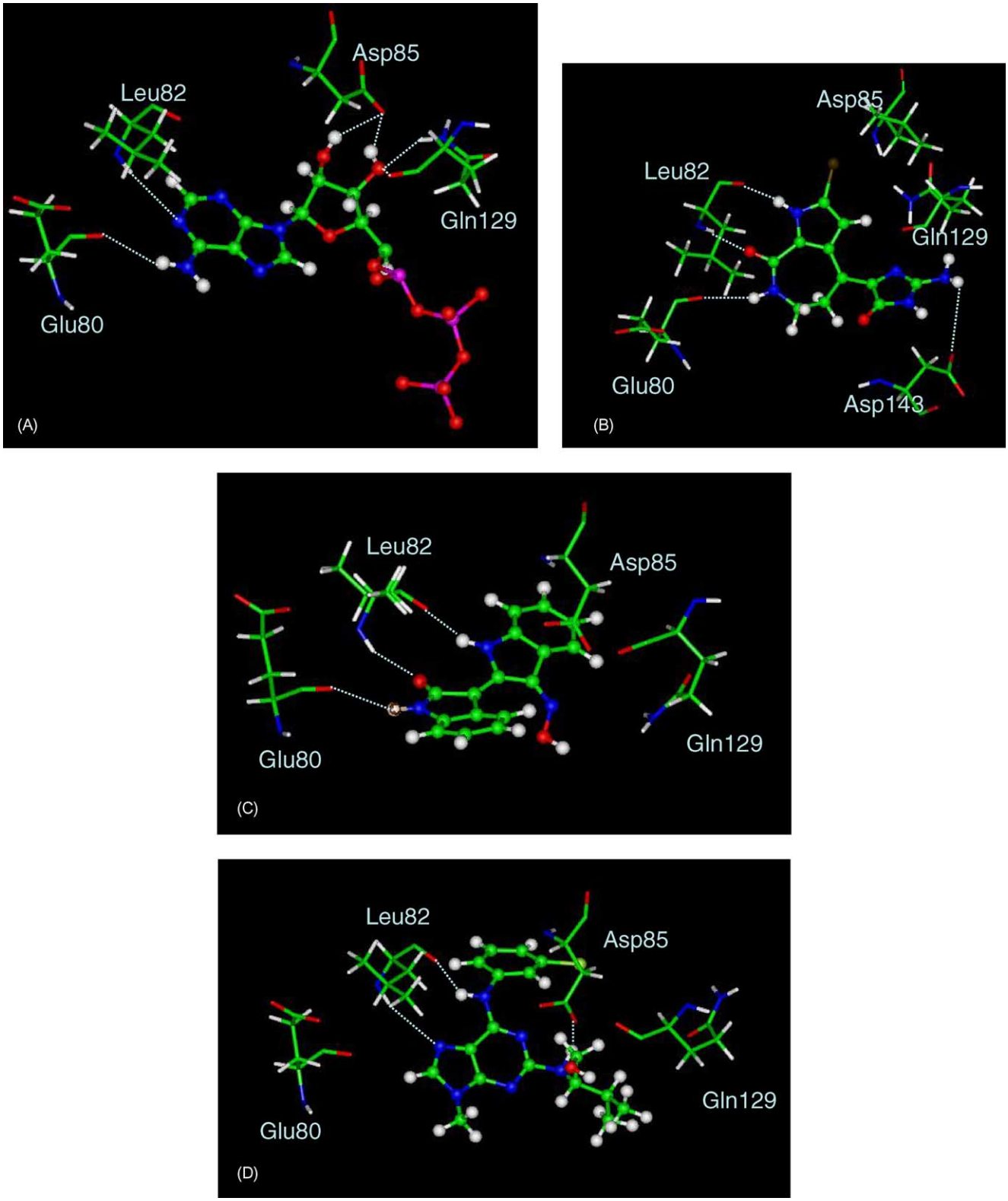

Fig. 2. Binding orientation of HYM, I3M and PURA. Suggested hydrogen bond interactions between ATP and P*f*PK5 (A), and the consensus orientation of 25 individual dockings obtained from GOLD for HYM (B), I3M (C) and PURA (D) to P*f*PK5.

[1,21,22]. To determine the orientations of HYM, I3M and PURA in the ATP binding site of P*f*PK5 without bias of the relative orientation of these inhibitors (or analogues) with CDK2, we applied a ligand-docking algorithm to explore ligand conformation and to satisfy ligand-binding requirements. We systematically docked HYM, I3M and PURA into the ATP binding pocket of P*f*PK5 using GOLD [15]. One advantage of GOLD over many other docking algorithms is that it allows for both unconstrained ligand flexibility and partial flexibility of the binding pocket, thus affording a more realistic environment for ligand. Remarkably, each docking analysis converged to a clear consensus orientation (Fig. 2B-D). Moreover, while there was high consensus for each inhibitor, as expected, there was no individual consensus orientation for all three inhibitors (Compare Fig. 2B-D). Next, we compared the binding orientations of the ligands bound to CDK2 with the results of our docking analysis. Details of the potential hydrogen bonds that stabilize each interaction are given in Table 1. Our results are highly consistent with the available crystallography data [1,21,22] and demonstrate a conservation of interactions between individual inhibitors bound to the two kinases.

While many of the interactions important for stable association between ligands and receptor are maintained between the P*f*PK5 and CDK2, we were also able to identify some interactions that offer avenues for potential inhibitor specificity. First in terms of sequence analysis, while the degree of sequence homology between the two kinases is sufficiently high to permit use of CDK2 as a structural template for building the P*f*PK5 homology model, there still exists a sufficient number of key residue substitutions located within and near the corresponding ligand-binding pockets of P*f*PK5 and CDK2 to provide clues for conferring inhibitor selectivity. For example, residues 80–84 in CDK2 have been shown to be important in the binding of both ATP and CDK inhibitors. There are two amino-acid substitutions within the aligned region of P*f*PK5: specifically, the substitution of a histidine for Phe81, and an aspartic acid for His84. Not only would these substitutions sterically alter inhibitor association, but they also exhibit significant differences in electrostatic properties and hydrogen-bonding potential (Fig. 3). Similar analyses for CDKs 1, 4 and 6 revealed possible strategies for conferring both individual and multiple selectivity for P*f*PK5. For example, Lys9 and Gln129 of P*f*PK5 align with glutamic acids in the CDK4 sequence. It is noteworthy that the notion of conferring selectivity via specific and single residue changes within the active site has previously been exploited for conferring inhibitor selectivity amongst the mammalian CDKs (viz. CDK2 and CDK4) [23–25].

We are also able to utilize our models to glean structural data concerning the binding of ATP competitive ligands and to delineate specificity. For example, the structure of CDK2 with indirubin-3′-monoxime bound suggests a hydrogen bond interaction between the ligand and Ile10. In the P*f*PK5-I3M model, Ile10 is at a distance of >5 Å, thus no such interaction is suggested. It is also interesting that the binding pocket for *PfPK5* is smaller (surface area = 576.1 Å^2; volume = 762.9 Å^3) than the CDK2 binding site (surface area = 639.7 Å^2; volume = 1004.6 Å^3), which might provide another strategy for designing *PfPK5*-selective ligands.

The results of our structural modeling and docking studies provide a wealth of structural information and, more importantly, details of the ATP-binding site of P*f*PK5. The consensus docking mode of each inhibitor is in agreement with structural data of the inhibitors bound to CDK2, and further identifies several residues that appear essential for ATP and inhibitor binding: Glu80, Leu82, and Asp85. Efforts are currently underway to further delineate the structural prerequisites for molecular recognition and tight binding P*f*PK5 inhibitors and, thereby, to guide the structure-based design of a new generation of antimalarial drugs for use with strains of P. *falciparum* resistant to traditional therapeutic agents.

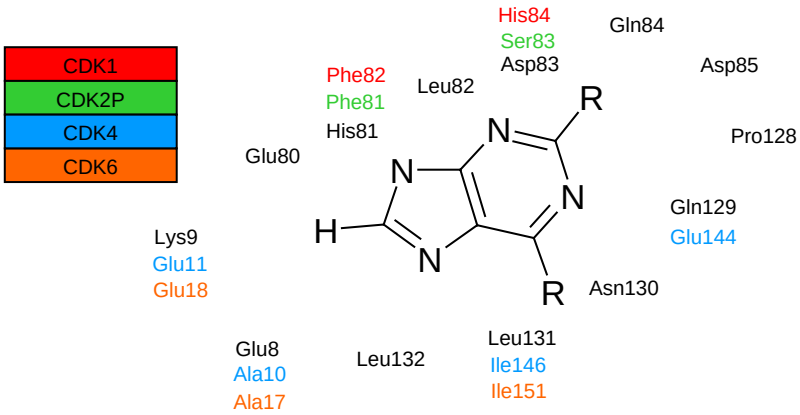

Fig. 3. Similar and dissimilar contact residues. Illustration of similar and dissimilar contact residues in the respective ligand binding pockets of P*f*PK5 (black), CDK1 (red), CDK2 (green), CDK4 (blue), and CDK6 (orange). Note that the variations at residues 8, 9, 81, and 131 (numbering P*f*PK5) provide a basis for conferring inhibitor selectivity for P*f*PK5 over the mammalian CDKs.

References

[1] N.S. Gray, et al., Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors, Science 281 (1998) 533–538.

[2] World Health Organization, 2002.

[3] P.B. Ross-Macdonald, et al., Isolation and expression of a gene specifying a cdc2-like protein kinase from the human malaria parasite *Plasmodium falciparum*, Eur. J. Biochem. 220 (1994) 693–701.

[4] Z. Li, et al., Influence of human p16(INK4) and p21(CIP1) on the in vitro activity of recombinant *Plasmodium falciparum* cyclin-dependent protein kinases, Biochem. Biophys. Res. Commun. 288 (2001) 1207–1211.

[5] B. Kappes, C.D. Doerig, R. Graeser, An overview of Plasmodium protein kinases, Parasitol. Today 15 (1999) 449–454.

[6] C. Doerig, J. Endicott, D. Chakrabarti, Cyclin-dependent kinase homologues of *Plasmodium falciparum*, Int. J. Parasitol. 32 (2002) 1575–1585.

[7] R. Graeser, R.M. Franklin, B. Kappes, Mechanisms of activation of the cdc2-related kinase PfPK5 from *Plasmodium falciparum*, Mol. Biochem. Parasitol. 79 (1996) 125–127.

[8] Y. Gu, J. Rosenblatt, D.O. Morgan, Cell cycle regulation of CDK2 activity by phosphorylation of Thr160 and Tyr15, EMBO J. 11 (1992) 3995–4005.

[9] K. Le Roch, et al., Activation of a *Plasmodium falciparum* cdc2-related kinase by heterologous p25 and cyclin H. Functional characterization of a *P. falciparum* cyclin homologue, J. Biol. Chem. 275 (2000) 8952–8958.

[10] C.J. Sherr, Cancer cell cycles, Science 274 (1996) 1672–1677.

[11] L. Harmse, et al., Structure-Activity Relationships and Inhibitory Effects of Various Purine Derivatives on the in vitro growth of *Plasmodium Falciparum*, Biochem. Pharmacol. 62 (2001) 341–348.

[12] R. Graeser, et al., *Plasmodium falciparum* protein kinase 5 and the malarial nuclear division cycles, Mol. Biochem. Parasitol. 82 (1996) 37–49.

[13] A.A. Russo, P.D. Jeffrey, N.P. Pavletich, Structural basis of cyclin-dependent kinase activation by phosphorylation, Nat. Struct. Biol. 3 (1996) 696–700.

[14] R. Rodriguez, et al., Homology modeling, model and software evaluation: three related resources, Bioinformatics 14 (1998) 523–528.

[15] G. Jones, et al., Development and validation of a genetic algorithm for flexible docking, J. Mol. Biol. 267 (1997) 727–748.

[16] J. Sachs, P. Malaney, The economic and social burden of malaria, Nature 415 (2002) 680–685.

[17] J. Pines, The cell cycle kinases, Seminars Cancer Biol. 5 (1994) 305–313.

[18] S.S. Taylor, E. Radzio-Andzelm, Three protein kinase structures define a common motif, Structure 2 (1994) 345–355.

[19] S.S. Taylor, et al., Structural framework for the protein kinase family, Ann. Rev. Cell Biol. 8 (1992) 429–462.

[20] N. Gray, et al., ATP-site directed inhibitors of cyclin-dependent kinases, Current Med. Chem. 6 (1999) 859–875.

[21] R. Hoessel, et al., Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases, Nature Cell Biol. 1 (1999) 60–67.

[22] L. Meijer, et al., Inhibition of cyclin-dependent kinases, GSK-3beta and CK1 by hymenialdisine, a marine sponge constituent, Chem. Biol. 7 (2000) 51–63.

[23] M. Barvian, et al., Pyrido[2,3-d]pyrimidin-7-one inhibitors of cyclin-dependent kinases, J. Med. Chem. 43 (2000) 4606–4616.

[24] K.S. Kim, et al., Thio- and oxoflavopiridols, cyclin-dependent kinase 1-selective inhibitors: synthesis and biological effects, J. Med. Chem. 43 (2000) 4126–4134.

[25] R. Soni, et al., Inhibition of cyclin-dependent kinase 4 (Cdk4) by fascaplysin, a marine natural product, Biochem. Biophys. Res. Commun. 275 (2000) 877–884.