PROTEIN KINASES AS DRUG TARGET AGAINST *TRYPANOSOMA BRUCEI*

Verma Divya

Department of Botany, Kalindi College (University of Delhi), New Delhi, India
Corresponding author: divyaverma@kalindi.du.ac.in

ABSTRACT

Protein kinases are the most studied proteins as drug targets against deadly disease 'sleeping sickness. The crystal structures for the *Trypanosoma brucei* protein kinase-A catalytic subunit isoform-1(PKAC1) and cell division-related protein kinase-2 (CDK2) are still not known. Therefore, homology models were constructed for the two proteins, based on their known amino acid sequences. The catalytic sites of both the proteins were then compared with their respective human homologs. Except for some conformational differences, the active site of TbrPKAC1 was found to be quite similar to that of the human homolog. Therefore, TbrPKAC1 cannot be considered as a very good drug target. Whereas, in the case of TbrCDK2, along with huge conformational differences, some important differences in the structure and nature of the binding site were also noticed when compared to their human homolog. Virtual screening was performed for TbrCDK2 and selected hits were analysed for the ligand-protein interactions. This analysis showed many important variations in TbrCDK2 from human homolog, which can be further explored as potential drug target.

Keywords: Human African Trypanosomiasis, Protein kinases, Drug target, Homology modelling, Docking.

1. INTRODUCTION

HAT (Human African Trypanosomiasis) or sleeping sickness is one of the "neglected tropical disease" [1], caused by Kinetoplastid, *Trypanosoma brucei* (Tbr) [2] and mostly transmitted by tsetse fly (*Glossina* genus) [3, 4]. There are two forms of HAT; the East African variant caused by *Tb rhodesiense* and the West African variant caused by *Tb. gambiense*. In both the forms, the first stage haemolymphatic phase, is characterised by fever, headaches, joint pains and itching; whereas in the second stage, *i.e.* in the neurological phase, the parasites reach central nervous system and cause characteristic signs and symptoms of the disease like confusion, sensory disturbances and poor coordination [5].

According to WHO Fact sheet [6], HAT threatens millions of people every year in sub-Saharan African countries, predominantly in rural populations. Unfortunately, the limited number of available drugs are old, expansive and complicated to administer; and are toxic at their therapeutic doses [7]. Therefore, new, effective, non-toxic, and affordable methods to diagnose and treat patients are direly needed.

Trypanosoma gene sequencing [8] and improvement of molecular and computational tools has opened new aspects for the identification and validation of novel drug targets for the research community [9]. Protein kinases are one of such targets which play crucial role in the regulation of majority of cellular functions like cell signalling, and cell cycle by modifying specific residue of other proteins through phosphorylation. In recent years, specific protein kinases have been targeted for the development of potential drugs for the treatment of different diseases like cancer, cardiovascular disease, malaria, trypanosomiasis, leishmaniasis etc [9-12]. In this research two specific *Trypanosoma brucei* protein kinases, viz. Protein Kinase-A Catalytic Subunit Isoform-1 (PKAC1) of *Tb. brucei* strain 927/4 GUTat10.1 with gene identifier Tb09.211.2410 and Cell Division-Related Protein Kinase-2 (CDK2) of *Trypanosoma brucei* TREU 927 with NCBI Reference Sequence XP_822746.1, were studied as possible drug targets.

2. BIOLOGY OF TARGETS

Protein kinases (PK) play an important role in signal transduction, environmental cues transmission and coordination of intracellular processes. In eukaryotes, the protein kinases are classified on the bases of the amino acid sequence of their catalytic domains [13]. According

to genome analysis, there are 156 ePKs (eukaryotic PKs) and 20 aPKs (atypical PKs) in *Tbr*. The ePKs which require phosphorylation in the activation loop between sub domains 7 and 8 for activation are marked by an RD motif within subdomain 6. In *T. brucei*, out of 156 ePKs, 130 ePKs are RD kinases, which suggest the importance of phosphorylation in cell regulation of Tbr [14]. In the present paper following two kinases were studied:

2.1. Protein kinase-A catalytic subunit isoform-1 (PKAC1): Gene identifier Tb09.211.2410

Protein kinase-A (PKA) is also known as cAMP-dependent protein kinase, as its activity depends upon the cellular levels of cyclic AMP (cAMP). The tetrameric holoenzyme consists of two regulatory and two catalytic subunits. The cAMP binds to the regulatory subunits and activates the catalytic subunits. The activated catalytic subunits dissociates from the holoenzyme and phosphorylates the serine/threonine residues of a wide range of substrate proteins involved in different physiological and developmental processes [15]. According to Siman-Tov *et al.*[16], this protein shows characteristic similarities to the mammalian counterpart and has the potential to get associated with mammalian regulatory subunit. The enzyme has also been found to be involved in parasite differentiation [17].

The cAMP-dependent protein kinase are included in the AGC group of kinomes, which is relatively poorly represented in Tbr as compared to human (half of that of human AGC kinase, after the normalisation of kinome size) [14]. In Tbr there are three AGC kinases that are related to PKA and are found to be activated by cGMP rather than cAMP.

2.2. Cell division-related protein kinase-2 (CDK2): NCBI Reference Sequence XP_ 822746.1

Cell Division-related Protein Kinase 2 is also known as cyclin-dependent kinase. It is a serine/threonine kinase which complex with the regulatory protein cyclin and phosphorylate substrate proteins involved in cell cycle regulation. In *Tbr* these are essential for transition through the G1/S phase and G2/M phase checkpoint of the cell cycle and are responsible for parasite growth and survival [18]. CDKs are the most studied kinase family as a drug target and have been considered as potential drug target for kinetoplastids [14, 19].

Cyclin-dependent kinase is included in CMGC group of kinomes, which is relatively well represented in *Tbr* with 11 members of CDKs [14]. Highly polarised cell with a

large cytoskeleton, a single mitochondria and complex association of cell cycle with life cycle differentiation makes the cell division a highly complex process. So, CDKs play a crucial role in the life cycle of *Tbr* and are essential for the survival of the parasite.

Biological importance of the two kinases suggests that these are essential for the parasite and can be used as a potential drug target against *T. brucei*.

3. SELECTION OF TEMPLATES

Till date, three dimensional structures for both the kinases of Tbr *i.e.,* PKAC1 (TbrPKAC1) and CDK2 (TbrCDK2) have not been reported. The nucleotide sequence of TbrPKAC1 was obtained from the NCBI database using the gene identifier Tb09.211.2410; which was then used to retrieve the amino acid sequence 'Q38DR5' (Q38DR5_TRYB2) from UniProtKB_ Nucleotide database [20, 21] by using blosum62 matrix through blastx (blastx BLASTX 2.2.29+). Further, the available three-dimensional structures of the sequences, homologous to the TbrPKAC1, were obtained from UniProtKB_PDB through blastx using blosum 62 matrix. The blastx result showed that the cAMP-dependent protein kinase catalytic subunit alpha, from *Mus musculus* (Mouse), with UniProtKB ID 'P05132' (KAPCA_MOUSE), was having maximum e-value of $440e^{-123}$, score of 915 and sequence identity of 54.7%; and hence was selected. The available X-ray crystallographic structure of P05132 (KAPCA_MOUSE) from RCSB-PDB database were analysed on the bases of resolution, R-value, B-factor, Ramachandran plots and sequence completeness. The 3D X-ray structure of P05132 (KAPCA_MOUSE) of *Mus musculus*, with PDB-ID 4O22 [22], having Resolution of 1.70 Å, R-Value of 0.187 (obs.), R-Free of 0.219 and the Ramachandran plot showing 97.5% (353/362) of all residues in favoured regions and 99.7% (361/362) of all residues in allowed regions, was selected to be used as template for homology modeling of TbrPKAC1.

For TbrCDK2, the amino acid sequence was obtained from the NCBI database by using NCBI Reference Sequence XP_822746.1. The available three-dimensional structures of the sequences, which were homologous to the TbrCDK2, were obtained from UniProtKB_PDB by performing blastx (BLASTX 2.2.29+) using blosum62 matrix. The blastx result showed that the Cell division Control Protein 2 Homolog from *Plasmodium falciparum* with UniProtKB ID 'Q07785' (CDC2H_PLAFK) was having maximum e-value of $410e^{-78}$, score of 600 and

sequence identity of 55.4%; and hence was selected. The known PDB structures of Q07785 (CDC2H_PLAFK) from RCSB-PDB database were analysed on the bases of resolution, R-value, B-factor, Ramachandran plots and sequence completeness. The 3D X-ray structures of Q07785 (CDC2H_PLAFK) of *Plasmodium falciparum*, with PDB-ID 1OB3 [23], having Resolution of 1.90 Å, R-Value of 0.193 (obs.), R-Free of 0.231 and the Ramachandran plot showing 96.6% (505/523) of all residues in favored regions and 100.0% (523/523) of all residues in allowed regions; was selected to be used as template for homology modeling of TbrCDK2.
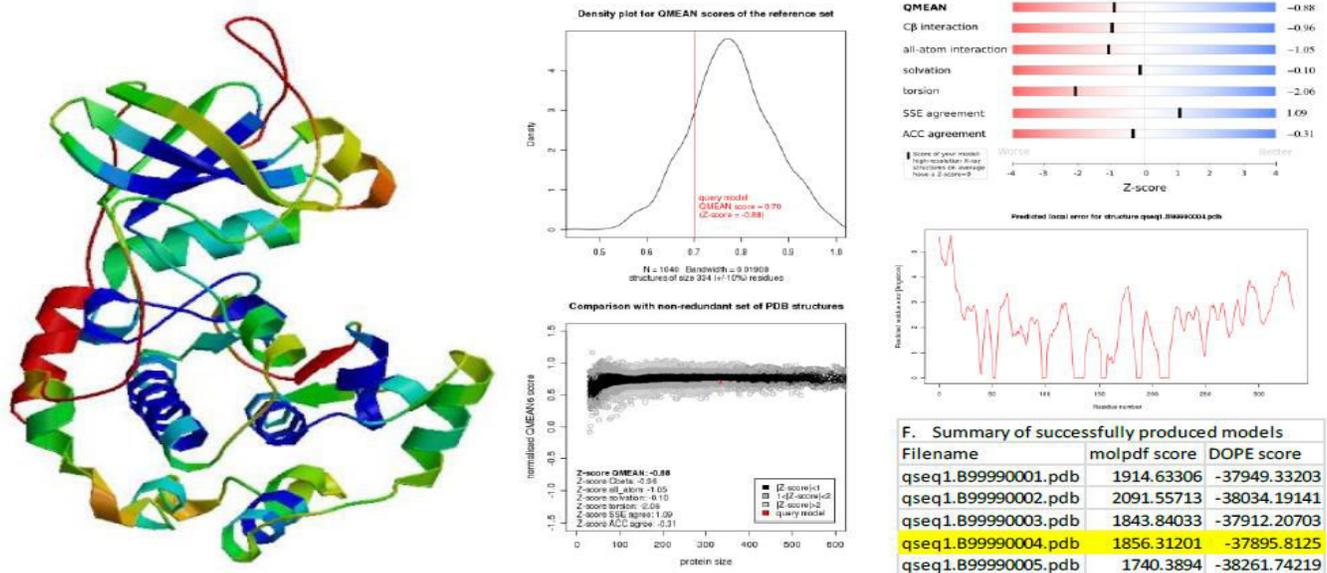

Fig. 1: Selected model of TbrPKAC1. (A) PDB structure colored according to B-Factor; (B) Density plot for QMEAN scores of the reference set; (C) Comparison with non-redundant set of PDB Structures (Query model as Red Cross); (D) QMEAN Z-score sliders; (E) Energy profile; (F) Comparative mol pdf and DOPE scores from the Modeller log file.
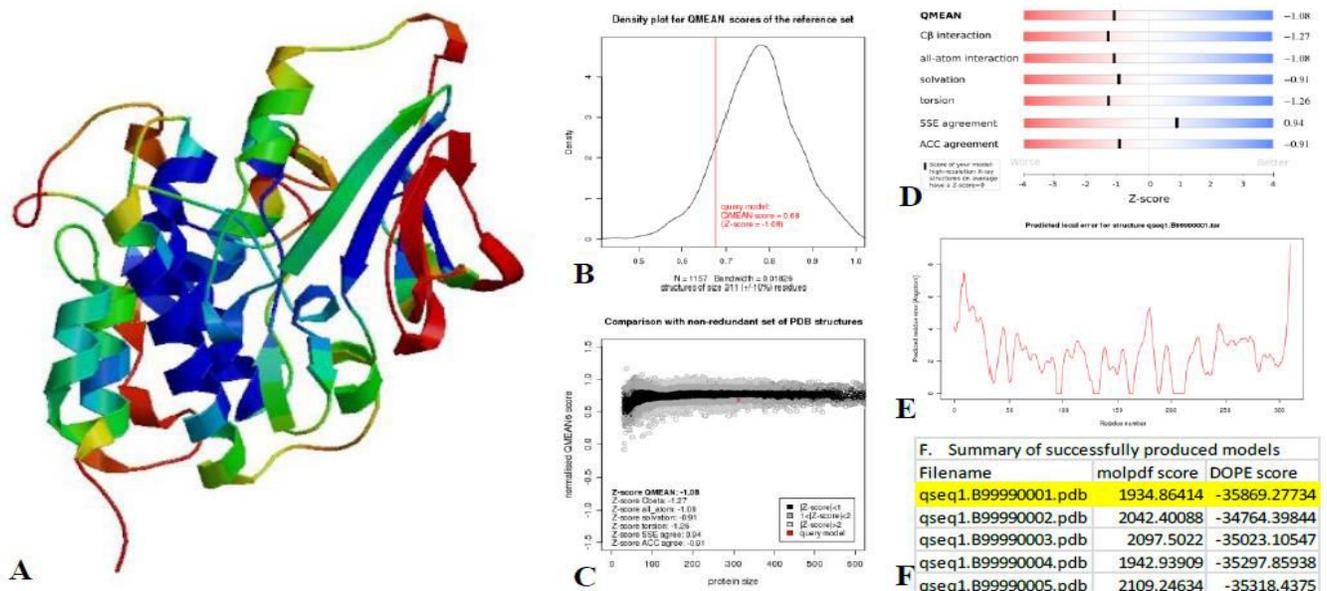

Fig. 2: Selected model of TbrCDK2. (A) PDB structure colored according to B-Factor; (B) Density plot for QMEAN scores of the reference set; (C) Comparison with non-redundant set of PDB Structures (Query model as Red Cross); (D) QMEAN Z-score sliders; (E) Energy profile; (F) Comparative mol pdf and DOPE scores from the Modeller log file.

4. HOMOLOGY MODELLING AND QUALITY ASSESSMENT

The python based programme modeller 9.15 [24, 25] was used for the homology modelling. It automatically calculates a model of all non-hydrogen atoms with the help of the aligned sequence in question with the known related structures and satisfies spatial restraints during comparative protein structure modelling [26, 27]. For TbrPKAC1, five different models of the Uniprot/Swiss-Prot Identifier: Q38DR5 were constructed using PDB structure '4O22' as a template. Similarly, for TbrCDK2, five different models of the NCBI Reference Sequence: XP_822746.1 were constructed using PDB structure '1OB3' as a template. The quality of the models thus obtained was estimated using QMEAN [28, 29] (Figure 1a-e and 2a-e), Molprobity [30, 31] and molpdf and DOPE scores from the log files of the modeller (Fig. 1f and 2f). Fourth model of TbrPKAC1 (Fig. 1a) and first model of TbrCDK2 (Fig. 2a) were selected for further studies.

5. ANALYSIS OF THE POCKET

The modelled structure of TbrPKAC1 and TbrCDK2 thus obtained were then compared to their respective human homologs. The human homologs for both the proteins were obtained from UniProtKB_PDB through blastx. For TbrPKAC1, the PDB structure 4WB5 [32] of the isoform 2 of cAMP-dependent protein kinase catalytic subunit alpha of *Homo sapiens* (HssPKAC1) showing e-value of $1.7e^{-117}$, score of 913, sequence identity of 56.0%, resolution of 1.64Å, R-Value of 0.165 (obs.), R-Free of 0.196 and the Ramachandran plot showing 98.2% (378/385) of all residues in favoured regions and 100.0% (385/385) of all residues in allowed regions of the Ramachandran plot, was selected. Similarly, for TbrCDK2, the PDB structure 2CCH [32] of Cyclin-dependent kinase-2 of *Homo sapiens* (HssCDK2) showing e-value of $36e^{-57}$, score of 458, sequence identity of 66.4%, resolution of 1.70Å, R-Value of 0.150 (obs.), R-Free of 0.182 and the Ramachandran plot showing 98.2% of all residues in favoured regions and 99.7% of all residues in allowed regions of the Ramachandran plot, was selected. The comparative structures of both the proteins with their respective human homologs are illustrated in the figure 3.

The binding sites of the modelled proteins i.e., TbrPKAC1 and TbrCDK2 were also analysed and then compared to the binding sites of their respective Human Homologs. So, a grid of site points was created in the binding pockets of TbrPKAC1 and TbrCDK2 using a high throughput in-silicon screening program LIDAEUS (Ligand Discovery at Edinburgh University) [33, 34] (Fig. 4A, C). Each grid point is associated with calculated properties, including van der Waals interaction energy, H-bonding capacity, and the extent to which it is buried. The site points are the subsets of grid points with their associated properties. For TbrPKAC1, the grid generation was based upon the known PKAC-ADP complex (PDB id: 2o21) of *Mus musculus* [22] (Fig. 4B), whereas for TbrCDK2, it was based upon CDK-Indirubin-5-Sulphonate Enzyme Inhibitor complex (PDB id: 1V0O) of *Plasmodium falciparum* [23] (Fig. 4D). The site points thus generated were used for the detailed analysis of the binding sites of the modelled proteins.

In TbrPKAC1, the side chains of the amino acid residues lys-54, glu-73, glu-109, asp-148, asn-153, thr-165, asp-166 and the peptide back bone of the amino acid residues leu-31, gly-32, thr-33, gly-37, val-86, glu-103, val-105, gly-108, glu-152, leu-156, val-164, phe-167, tyr-312 form the active sites of the binding pocket. Whereas, in the human homolog (HssPKAC1),the side chains of the amino acid residues lys-72, glu-91, glu-127, asp-166, asn-171, thr-183, asp-184 and the peptide back bone of the amino acid residues leu-49, gly-50, thr-51, gly-55, val-104, glu-121, val-123, gly-126, glu-170, leu-174, val-182, phe-185, tyr-330 form the active sites of the binding pocket. Thus, the comparative study of the binding pocket of the modelled protein and its human homolog do not show significant difference in the active residues of the binding pocket except a difference in the orientation of the residues (Figure 4E).

In TbrCDK2, the side chains of the amino acid residues thr-33, lys-52, tyr-101, asp-105, lys-108, asp-163 and the peptide back bone of the amino acid residues leu-29, glu-31, thr-33, ala-50, leu-97, glu-100, val-102, asp-103, his-104, ala-149, asp-163 form the active sites of the binding pocket. Whereas, in the human homolog (HssCDK2) the side chains of the amino acid residues thr-14, lys-33, phe-82, asp-86, lys-89, asp-145 and the peptide back bone of the amino acid residues ile-10, glu-12, thr-14, ala-31, leu-78, glu-81, leu-83, his-84, gln-85, gln-131, asp-145 form the active sites of the binding pocket (Fig. 4F). Thus, the comparative study of the binding pocket of the modelled protein and its human homolog was not only showing differences in the

orientation of the residues, but significant variations were also observed in the type and nature of residues present in the binding pocket. It was observed that HssCDK2 was having ile-10, phe-82, leu-83, his-84, gln-85 and gln-131 residues respectively in its binding pocket as compared to leu-29, tyr-101, val-102, asp-103, his-104 and ala-149 residues of the binding pocket of TbrCDK2. The polar residue tyr-101, which is one of the active residues of the binding pocket in the parasitic protein, has some acidic properties and can act as both H-bond acceptor and H-bond donor, whereas the human homolog has hydrophobic residue phe-82, which is neither H-bond acceptor or H-bond donor.

Similarly, the val-102 residue in TbrCDK2 is replaced by more bulky group leu-83 in HssPKAC1; the negatively charged polar residue asp-103 in TbrCDK2is replaced by a positively charged polar residue his-84 in HssPKAC1; the positively charged polar residue his-104 in the parasitic protein is replaced by an uncharged polar residue gln-85 in the human homolog; and the hydrophobic residue ala-149 in TbrCDK2is replaced by a more bulky, uncharged, polar residue gln-131 in HssPKAC1. These variations in the binding pocket of TbrCDK2, in comparison to that of the human homolog, make it a potential drug target.
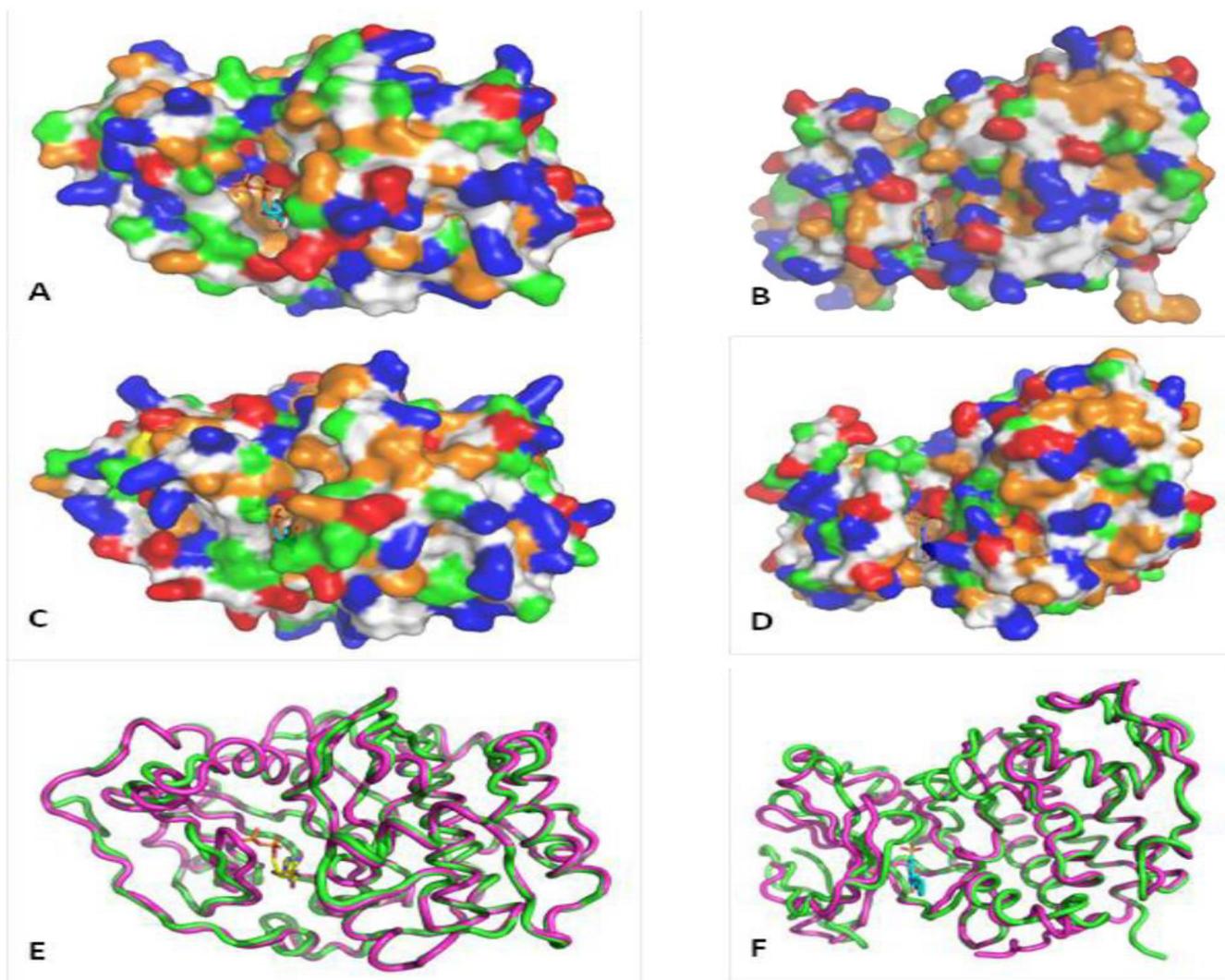

Fig. 3: Comparison of the modelled structures with their human homologs. Surface view of (A) TbrPKAC1, (B) TbrCDK2, (C) HssPKAC1 (D) HssCDK2 (colored according to hydrophobicity: acidic, red; basic, blue; polar, green; non-polar, orange; cysteine residue, yellow). Ribbon view of alignment of (E) TbrPKAC1 (green) with HssPKAC1 (magenta), and (F) TbrCDK2 (green) with HssCDK2 (magenta).
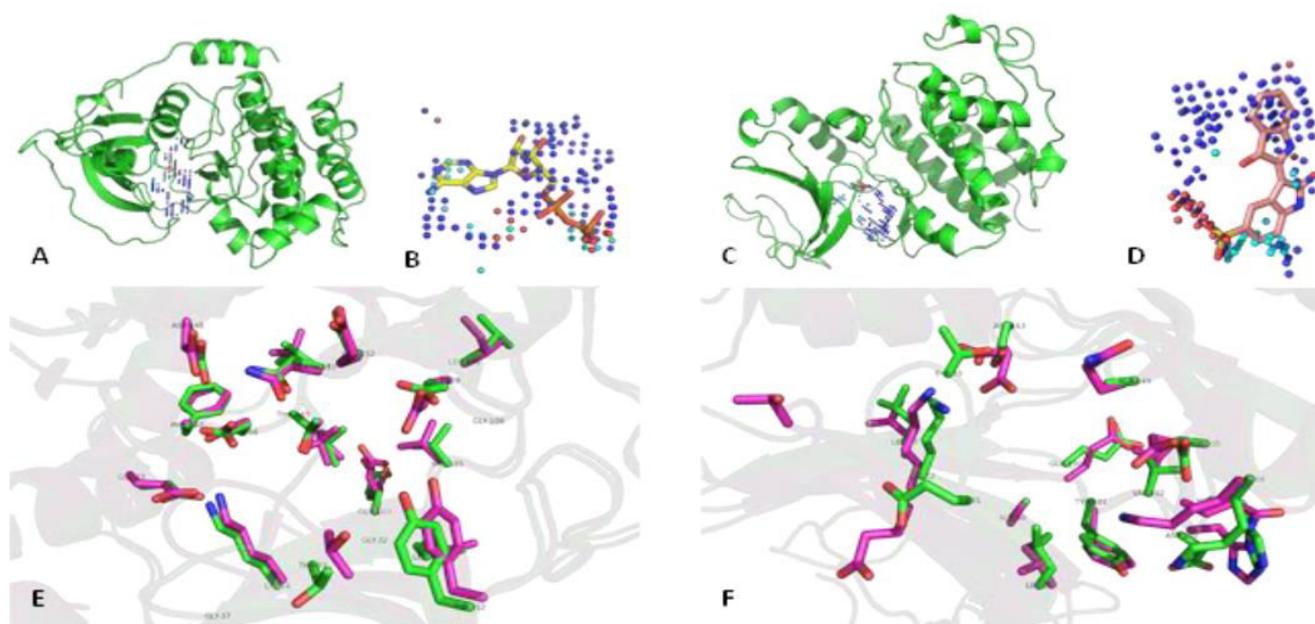
Fig. 4: **(A) Grid of site points in the binding pocket of TbrPKAC1 and (B) the enlarged view of the site points grid with ADP of PKAC-ADP complex (C) Grid of site points in the binding pocket of TbrCDK2 and (D) the enlarged view of the site points grid with the Indirubin–5–Sulphonate of CDK- Enzyme Inhibitor complex. (E) Binding pocket of TbrPKAC1 (green) compared with the binding pocket of HssPKAC1 (magenta). (F) Binding pocket of TbrCDK2 (green) compared with the binding pocket of HssCDK2 (magenta).**

6. DOCKING AND LIGAND ANALYSIS

The high throughput in-silicon screening program LIDAEUS (Ligand Discovery At Edinburgh University) was further applied for ligand screening [33, 34]. It docks selected molecules from a small molecule database, EDULISS (Edinburgh University Ligand Selection System) having data-mining and pharmacophore searching capabilities, into the grid of the site points [35]. It matches atoms of the molecule to site points, explores the binding pocket thoroughly, identifies appropriate poses, and screens the hits. During the docking procedure, the atomic properties of the potential ligand are matched with the calculated values of the site points. All the interaction between the ligand and the protein at various poses are checked to avoid severe protein-ligand clashes. During this process all the ligands are scored and ranked according to the enthalpy of interaction and other properties like van der Waals interaction energy, H-bonding capacity, and the extent to which it is buried. Finally, a rigid-body energy minimization of the top hits provides the ranked list of hits. Because of considerable difference in the TbrPKAC1 and its human homolog, the in-silicon screening program LIDAEUS was run to screen top 20 ligands from the small molecule database EDULISS. Out of these 20 ligands, hit-3 and hit-7 (Figure 5A, B), which were showing interaction with one of the active polar residues tyr-101 of the binding pocket of the parasitic protein, were selected for further in-silico analysis. The interactions of these two hits with the parasitic protein and the human homolog were then compared.

The hit-3 was showing polar interactions with side chains of the amino acid residues thr-33, tyr-101, lys-107, lys-108, lys-147 and the peptide back bone of the amino acid residues val-102 of TbrCDK2. However, when compared to human homolog, phe-82 of HssCDK was a non-polar, hydrophobic residue as compared to tyr-101 of TbrCDK2, which is a polar residue having acidic properties and can act as both H-bond acceptor and H-bond donor; thr-14 of HssCDK was showing significantly large difference in the orientation (of about 10Å) than thr-33 of TbrCDK2; lys-89 of HssCDK was showing a considerable difference in the orientation (of about 3.6Å) than lys-108 of TbrCDK2; and a very negligible difference in orientation of lys-88 and lys-129 of HssCDK was observed as compared to lys-107 and lys-147 of TbrCDK2 respectively (Fig. 5C).
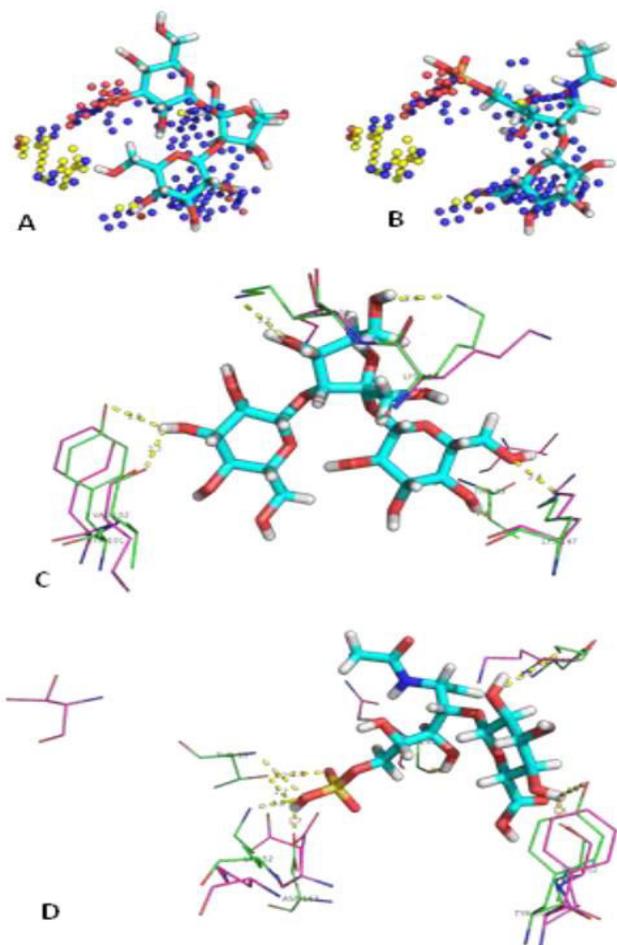
Fig. 5: Orientation of (A) Hit-3 and (B) Hit-7 with respect to the grid of site points in TbrCDK2. (C) Polar interactions of the Hit-3 with the active residues in the binding pocket of TbrCDK2 (green) compared with that of HssCDK2 (magenta), (D) Polar interactions of the Hit-7 with the active residues in the binding pocket of TbrCDK2 (green) compared with that of HssCDK2 (magenta). (H-bond acceptor, red; H-bond donor, blue and hydrophobic, yellow)

The hit-7 was showing polar interactions with side chains of the amino acid residues thr-33, lys-52, tyr-101, lys-108, asp-163and the peptide back bone of the amino acid residues thr-33, val-102 and ala-149 of the parasitic protein, TbrCDK2. When compared to human homolog, phe-82 of HssCDK is a non-polar, hydrophobic residue as compared to tyr-101 of TbrCDK2, which is a polar residue having acidic properties and can act as both H-bond acceptor and H-bond donor; thr-14, lys-33 and lys-89 were showing a considerable difference in the orientation (of about 3Å

to 4Å) than thr-33, lys-52 and lys-108 of TbrCDK2 respectively; gln-131 and asp-143 of HssCDK were showing a very small difference in orientation (of about 1Å to 2Å) than ala-149 and asp-163of TbrCDK2 respectively; and no apparent differences in orientation were observed between val-102 and leu-83 of HssCDK and TbrCDK2 respectively (Fig. 5D).

7. DISCUSSIONS AND FUTURE PERSPECTIVE

In the present study, the structure of Protein kinase A catalytic subunit isoform 1 (PKAC1) and Cell division-related protein kinase 2 (CDK2) of *Trypanosoma brucei* were modelled and then compared to their respective human homologs. The structures of TbrPKAC1 and HssPKAC1 are quite similar and the active residues of their binding pockets are also highly conserved. However, a significant difference has been observed in the structures of TbrCDK2 and HssCDK2. Some striking variances are observed between the binding residues of TbrCDK2 and HssCDK2. The binding pocket of TbrCDK2 has Ala (hydrophobic), Leu (hydrophobic), His (basic) and Asp (polar) residue; whereas, the binding pocket of the human homolog has Gln (polar), Ile (hydrophobic), Gln (polar) and His (basic) respectively. Though, the other residues in the binding pockets in TbrPKAC1 and human homolog are identical, yet a great difference has been observed in the conformation of the two pockets. For further analysis, docking was performed on TbrPKAC1 and selected hits were used to compare the protein-ligand interactions of the parasite and the host. Due to the presence of great variations in the type of active residues as well as the differences in the orientation of the residues in the binding pockets of the parasitic protein and the host protein, remarkable differences are observed in the protein-ligand interactions. These differences suggest TbrPKAC1 as a potential drug target, which can be further explored.

8. REFERENCES

1. Feasey N, Wansbrough-Jones M, Mabey DCW, Solomon AW. *British medical bulletin*, 2010; **93**:179-200.
2. Simarro P, Franco J, Diarra A, Postigo J, Jannin J.

Research & Reports in Tropical Medicine, 2013; **4**:1-6.

3. Van den Bossche P, Rocque SD La, Hendrickx G, Bouyer J. *Trends in Parasitology*, 2010; **26**(5):236-243.

4. Leak SGA. In: *CABI Publishing. Oxford and New York*. 1998, p 568.

5. Kennedy PG. *Annals of Neurology*, 2008; **64**(2):116-126.

6. WHO. 2015.

7. Malvy D, Chappuis F. *Clinical Microbiology and Infection*, 2011; **17**(7):986–995.

8. Berriman M, Ghedin E, et al. *Science (New York, N.Y.)*, 2005; **309**(5733):416-422.

9. Naula C, Burchmore R. *Expert Review of anti-infective therapy*, 2003; **1**(1):157-65.

10. Doerig C, Meijer L, Mottram JC. *Trends in Parasitology*, 2002; **18**(8):366-371.

11. Verica P, Margaret MH. *Drugs*, 2013; **73**(2):101-115.

12. Cohen P, Alessi DR. *ACS Chem Biol*, 2013; **8**(1):96-104.

13. Manning G, Whyte DB, Martinez R, Hunter T SS. *Science*, 2002; **298**(5600):1912-34.

14. Parsons M, Worthey E a, Ward PN, Mottram JC. *BMC genomics*, 2005; **6**:127.

15. Hanks SK, Quinn AM. *Protein Phosphorylation Part A: Protein Kinases: Assays, Purification, Antibodies, Functional Analysis, Cloning, and Expression*. Elsevier, 1991

16. Siman-Tov MM, Ivens AC, Jaffe CL. *Gene*, 2002; **288**(1):65-75.

17. Heath S, Hieny S, Sher A. *Molecular and biochemical parasitology*, 1990; **43**(1):133-141.

18. Tu X, Wang CC. *Journal of Biological Chemistry*, 2004; **279**(19):20519–20528.

19. Doerig C. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 2004; **1697**(1-2):155-168.

20. Magrane M, Consortium UP. *Database*, 2011; **2011**:1-13.

21. The Uniprot Consortium. *Nucleic Acids Research*, 2014; **43**(D1):D204-D212.

22. Gerlits O, Das A, Keshwani MM, Taylor S, Waltman MJ, Langan P, Heller WT, Kovalevsky A. *Biochemistry*, 2014; **53**(19):3179-3186.

23. Holton S, Merckx A, Burgess D, Doerig C, Noble M, Endicott J. *Structure*, 2003; **11**(11):1329-1337.

24. Eswar N, Webb B, Marti-renom MA, Madhusudhan MS, Eramian D, Shen M, Pieper U, Sali A. *Comparative protein structure modeling using Modeller*. 2006

25. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, A S. *Annu Rev Biophys Biomol Struct*, 2000; **29**:291-325.

26. Sali A, Blundell TL. *J. Mol. Biol.*, 1993; **234**:779-815.

27. Fiser A, Do RK, Sali A. *Protein Science*, 2000; **9**:1753-1773.

28. Benkert P, Tosatto SC, Schomburg D. *Proteins*, 2008; **71**(1):261-77.

29. Benkert P, Biasini M, Schwede T. *Bioinformatics*, 2011; **27**(3):343-350.

30. Chen VB, Bryan Arendall W, et al. *Acta Crystallogr D Biol Crystallogr*, 2010; **66**(1):12-21.

31. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC. *Nucleic Acids Research*, 2007; **35**:375-383.

32. Cheung J, Ginter C, Cassidy M, Franklin MC, Rudolph MJ, Robine N, Darnell RB, Hendrickson W a. *Proceedings of the National Academy of Sciences*, 2015; **112**(5):1374-1379.

33. Wu SY, McNae I, Kontopidis G, McClue SJ, McInnes C, Stewart KJ, Wang S, Zheleva DI, Marriage H, Lane DP, Taylor P, Fischer PM, Walkinshaw MD. *Structure*, 2003; **11**(4):399-410.

34. Taylor P, Blackburn E, Sheng YG, Harding S, Hsin K-Y, Kan D, Shave S, Walkinshaw MD. *British journal of pharmacology*, 2008; **153 Suppl**:S55-S67.

35. Hsin KY, Morgan HP, Shave SR, Hinton AC, Taylor P, Walkinshaw MD. *Nucleic Acids Research*, 2010; **39**(SUPPL. 1):1-7.