Journal of Computational Intelligence in Bioinformatics
ISSN: 0973-385X Volume 3 Number 1 (2010) pp. 21–30

http://www.ripublication.com/jcib.htm

Homology Modeling of Leishmania CRK3

Sen ParthoSarathi[1]* and Dubey Rounak[2]

[1]National Institute of Pharmaceutical Education and Research,Kolkata, INDIA.
[2]Indian Institute of Chemical Biology, Kolkata, INDIA.
**Corresponding author Email: partho.phd@gmail.com*

Introduction

Several syndromes are subsumed under the term Leishmaniasis ranging from the self-healing cutaneous lesions to the potentially fatal visceral form caused by over 20 different Leishmania species. Leishmaniasis has traditionally been classified in three different clinical forms, visceral leishmaniasis (VL), cutaneous leishmaniasis (CL), and mucocutaneous leishmaniasis (MCL), which have different immunopathologies and degrees of morbidity and mortality. VL caused by Leishmania donovani is fatal if untreated, whereas CL caused by other leishmania species such as Leishmania major, Leishmania mexicana etc. frequently self cures within 3–18 months. About two million cases reported annually for leishmaniasis and it is alarming that about 350 million people live in endemic areas under the risk of infection(Vannier-Santos, et al., 2002). *Leishmaniasis* is found in at least 88 countries but more than 90% of cases are observed in underdeveloped or developing countries such as Brazil, Bangladesh, India and Sudan. The relevance of this parasitic disease is further stressed out by the rise of Leishmania/HIV co-infection in many parts of the world including European countries such as Spain, Italy, France and Portugal where up to 9% of the AIDS patients suffer from visceral Leishmaniasis(Desjeux, 1998). *Leishmania* spp. is trypanosomatid digenetic parasites which infect humans and other mammals. In an invertebrate phlebotomine sand fly host the parasites proliferate in the promastigote (flagelate) form in the insect gut whereas in a vertebrate host the protozoa are obligate intracellular parasites in the amastigote (non-motile) form infecting mostly mononuclear phagocytes. Thus, there appears to be an important link between the developmental life cycle of this parasite and the regulation of its cell cycle which provides rational for targeting cell cycle of leishmania parasites for treatment of leishmaniasis(Chang, 1983).

The pentavalent antimonials have been recommended for the treatment of *Leishmania*sis for over 50 years(Croft, et al., 2003). Other drug used in the treatment of *Leishmania*sis includes the diamidine pentamidine and Amphotericin B. However, the use of these drugs has been limited due to toxicity, resistance and also, in the case of Amphotericin B, the route of administration as it is via slow parenteral infusion

over several hours. Newer drugs, such as the lipid formulations of amphotericin B (AmBisome, Amphocil and Abelcet), have been effective in the treatment of visceral *Leishmania*sis(Murray, 2004). Unfortunately, the prohibitive cost of the new formulations of this drug means that this treatment is unavailable to the majority of patients with visceral *Leishmania*sis. An exciting new development has been the discovery of miltefosine (Impavido), an alkylphospholipid, has shown efficacy as oral treatment for visceral *Leishmania*sis in India. However, miltefosine is not without problems: it is teratogenic and has a low therapeutic index.

As we known that current chemotherapy for *Leishmania*sis, have many drawbacks there is not only an urgent need for new antileishmanial drugs based on existing targets but also to validate new drug targets which would help in development of new antileishmanial drugs, overcoming the drawbacks of current ones.

During investigations into the cell and life cycles of these parasites, "*K.M. Grant*" has isolated two cdc2-related kinase genes CRK1 and CRK3 from *Leishmania Donavani*. The encoded enzymes are homologous to the cyclin-dependent kinase (CDK) family of serine/threonine protein kinases, which are ubiquitous in eukaryotes, many of which play essential roles in the regulation and co-ordination of the cell cycle(Grant, et al., 1998). Both these genes have been shown by genetic manipulation to be essential to the parasite, implying that inhibitors of the encoded proteins would have a deleterious effect upon the parasite(Doerig, et al., 2002). Review of literature showed within the CDK family, five CDKs are involved in the regulation of the cell cycle, which is turned on and off at precise points during the cell cycle. A prerequisite for CDK activity is the binding of the relevant cyclin partner to the kinase subunit to form a heterodimers. Furthermore, activity can be up or down regulated by phosphorylation at two distinct, conserved phosphorylation sites (Thr- 161 and Tyr-15). The crucial role of CDKs in the regulation of cell division and the high incidence with which their activity is abnormally regulated in human cancers suggest that CDKs would be good targets for new anticancer agents. Several CDK inhibitors have been identified which exhibit specificity for the ATP-binding pocket of these CDKs and block their phosphorylation(Walker, 1998). Several lines of evidence indicate that CRK3 is the most likely candidate for the functional CDK1 (cdc2) homologue in *Leishmania* and also bears high sequence identity to human CDKs. As the cell cycle of *Leishmania* is closely regulated, as in other eukaryotes, and integrated with its differentiation between the various life cycle stages targeting leishmania CRK3 provides a new way for treatment of leishmaniasis by arresting leishmania cell cycle. CRK3 is active in the two proliferative life cycle stages of the parasite (the insect stage promastigote and the mammalian stage amastigote), Attempts to generate a null mutant resulted in a dramatic change in the parasite's ploidy to avoid loss of this essential gene (this phenotype is widely interpreted to mean that the gene is essential to the organism.). Rational for targeting leishmania CRK3 is same as that of targeting human CDKs for cancer treatment in which cytotoxic drugs would have more effect on tumor cells, which are rapidly dividing, than on normal cells, most of which do not divide. Base on these rational CRK3 inhibitors would have more effect on leishmania parasites which are rapidly dividing than normal human cells thus overcoming toxic effects which would result from inhibiting human CDKs.

Since the CRK3 bears significant similarity to human CDKs, the inhibitors for targeting leishmania CRK3 should be selective for parasites. This study carried out with aim to understand important structural difference between human CDKs, leishmania CRK3 and to provide standard structural model of leishmania CRK3 for designing and virtual screening of new inhibitors. The three dimentional structure for leishmania CRK3 is not available. We presented here homology modeling of the leishmania CRK3 followed by comparative cavity depth analysis of model CRK3 structure and human CDKs. The docking analysis performed using 20 reported inhibitors for leishmania CRK3(Grant, et al., 2004).

Aims & Objectives

The proposed work concentrates on *Leishmania* cyclin dependent kinase2 related protein.

The aim is to utilize *in silico* techniques to develop homology model of *Leishmania* CRK3 protein for screening and studying CRK3 inhibitors interaction with protein which will help in development of specific inhibitors against *Leishmania* CRK3 protein.

Objective

- To predict the structure of CRK3 (cdc2 related kinase) using MOE Package by CCG.
- To perform docking analysis with different class of CRK3 inhibitors.

Homology Modeling of Leishmania CRK3

CRK3 is a cdc2-related kinase 3, belongs to protein kinase superfamily which are primarily involved in a regulation of cell cycle. Leishmania donovani CRK3 is composed of about 311 amino acids having conserved protein kinase core. leishmania parasites which are rapidly dividing than normal human cells thus overcoming toxic effects which would result from inhibiting human CDKs.

Since the CRK3 bears significant similarity to human CDKs, the inhibitors for targeting leishmania CRK3 should be selective for parasites. This study carried out with aim to understand important structural difference between human CDKs, leishmania CRK3 and to provide standard structural model of leishmania CRK3 for designing and virtual screening of new inhibitors. The three dimentional structure for leishmania CRK3 is not available. We presented here homology modeling of the leishmania CRK3 followed by comparative cavity depth analysis of model CRK3 structure and human CDKs. The docking analysis performed using 20 reported inhibitors for leishmania CRK3(Grant, et al., 2004).

In order to find out homologues sequences to leishmania CRK3, which would used as a templates while model building , NCBI BLASTP search against protein data bank using leishmania CRK3 as query sequence were carried out. Number of

sequence homologues to CRK3 was obtained. Percentage sequence identity and similarity together with E-value (expectation value) for top two hits from NCBI BLASTP search is shown in Table 1. The Evalue represents a number of different alignments with scores equivalent to or better than scores that are expected to occur in a random database search. Generally lower E –value indicates that alignment is more realistic and not by chance. In Table 1 both protein sequence shows same identity and similarity of 58% and 75% respectively with query sequence. The multiple sequence alignment for query sequence and two hits were shown in Figure 1 below.

Table 1: Top two hits from blastp search using leishmania CRK3 query sequence.

PDB code	Protein name	% identity	%similarity	E-value
1GZ8	Human Thr160-Phospho Cdk2-Cyclin A F82h-L83v- H84d Mutant	58	75	1e-100
2IW8	Human Cyclin Dependent Kinase 2 Complexed with butoxypurine	58	75	1e-100

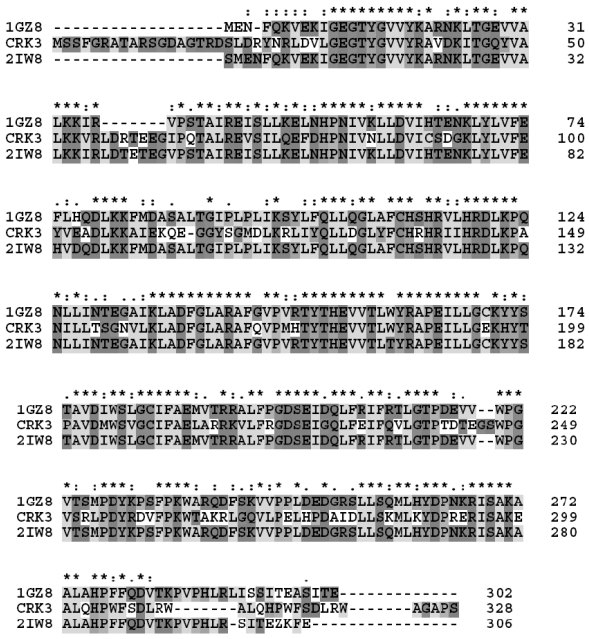
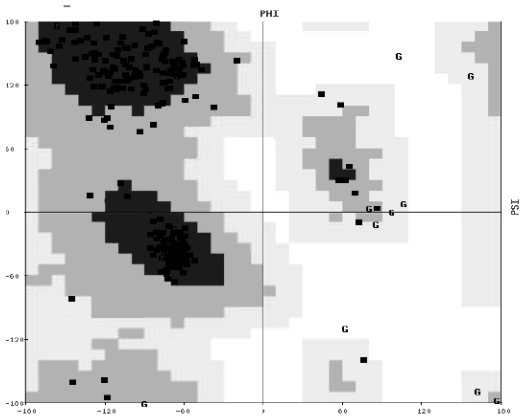


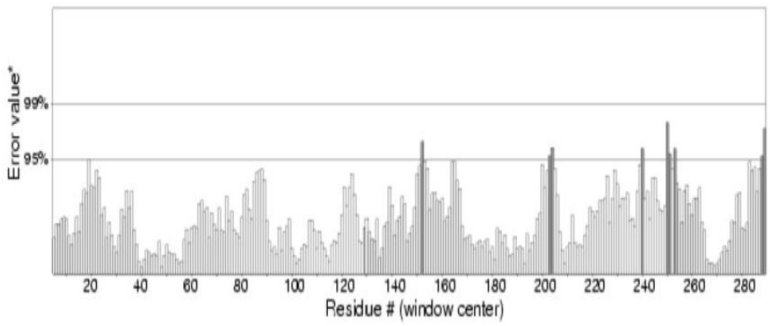
Figure 1: Multiple sequence alignment of leishmania CRK3 with templates 1GZ8 and 2IW8. Colors indicate amino acids with their similar characteristic, stars indicating identical amino acids while colon for similar amino acids and single dot for nearly similar ones.

1GZ8 is Human Cyclin Dependent Kinase 2 Complex with the inhibitor 2-Amino-6-(3'- Methyl-2'-Oxo) butoxypurine with high crystal structural resolution of 1.40 Å. Based on the high resolution crystal structure 1GZ8 and its ligand butoxypurine inhibitors similarity with the known reported inhibitors for leishmania CRK3 have directed to use 1GZ8 as a primary template for homology model building. 2IW8 protein was used as secondary template to model CRK3 sequence, aligning with gap in 1GZ8, as shown in Figure 1.

The initial crude structure of leishmania CRK3 obtained from homology model development module in MOE software (Chemical Computing Group Inc.) was further subjected to ligand modeling followed by loop modeling in MODELLER 8v2(Marti-Renom, et al., 2000).The resulting CRK3 homology model validated using tools like PROCHECK program (Laskovaski, 1993), ERRAT plot and Prosa II (Sippl, 1993). The backbone dihedral distribution of all the amino acids residues (Ramchandran plot) were 90.7 in core region, 7.8 in allowed region and 1.6 in generously allowed region (Figure 2a). This indicated that backbone dihedral angles, phi and psi, in the model were reasonably accurate. Homology model of CRK3 protein shows good ERRAT plot (Figure2b) with overall quality factors of 96.84 (Figure 2b). ERRAT Plot gives a measure of 'structure error' at each residue in the protein and also calculates an overall score for structural quality.



(2a)



(2b)

Figure 2: Ramchandran (2a) and Errat plot (2b) for CRK3 homology model.

For evaluating quality of protein folds CRK3 homology model was evaluated using Prosa II program. This program calculates the knowledge-based mean fields to judge the quality of protein folds and has been widely used to measure the stability of a protein conformation. Prosa II was used to calculate energy of both solved crystal structure of 1GZ8 and homology model of leishmania CRK3. Figure 3 shows Prosa II energy profile of the homology-modeled CRK3 in comparison to that of the X-ray structure of 1GZ8. In general folding energy of proteins show negative values and since this values correspond with the stability and nativity of the molecules. We note that the Prosa energy of leishmania CRK3 remains negative for all amino acid, indicating the acceptability of the homology-modeled structure. It can seen from figure that in most part of the sequence there is good correlation between model CRK3 and native 1GZ8.

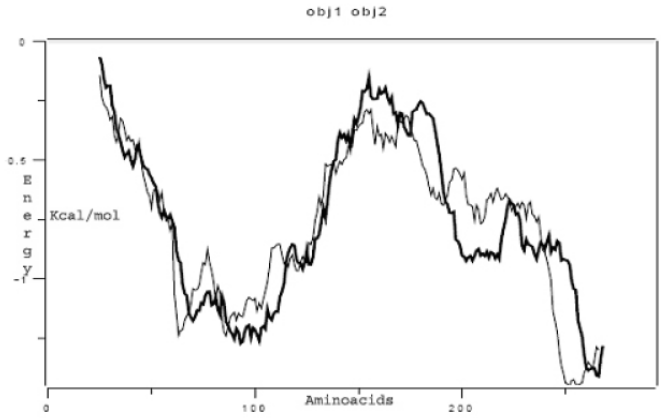


Figure 3: Prosa II energy profile of the homology-modeled CRK3 (heavy line) in comparison to that of the Xray structure of 1GZ8 (light line).

Figure 4 shows the structure of leishmania CRK3 obtained after performing homology modeling with the X-ray crystal structure of human CDK2 as a template.

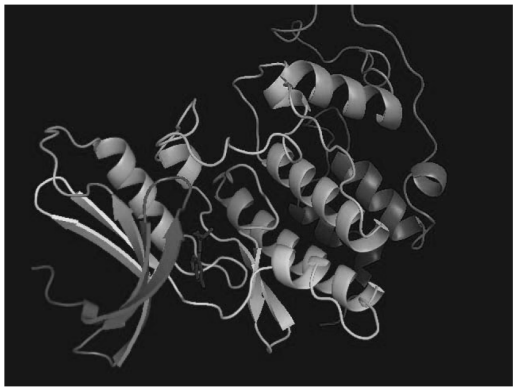


Figure 4: Ligand-supported homology model of leishmania CRK3, ligand shown in red color Capp's sticks Model.

The leishmania CRK3 is folded into the bilobal structure typical for most of protein kinase, with smaller N-terminal domain consisting of five antiparallel ß-strands and a single α-helix.

The larger C-terminal domain consists primarily of α-helix. It contain a pseudo-4-helical bundle, a small ß-ribbon and two additional α-helix these two domains are linked by a short hinge region and ATP binding site located in a deep cleft between them.

As seen from sequence alignment and model structure of leishmania CRK3, the most of the residues in the active site are conserved in both leishmania CRK3 and human CDK2.

The binding site for ATP as well as all the inhibitors which have their crystallographic structure determined in complex with CDK2 is located between the N- and C-terminal domains. Here we used butoxypurine inhibitor which is present in human CDK2 for ligand modeling in leishmania CRK3. The superposition of active site residues within 3.5 Å region surrounding butoxypurine inhibitor in both leishmania CRK3 and human CDK2 were shown in Figure 4.

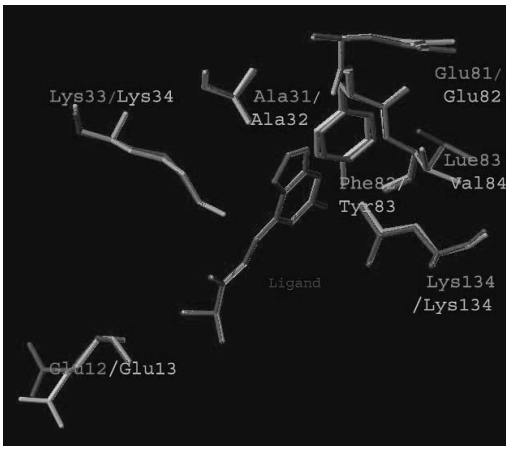
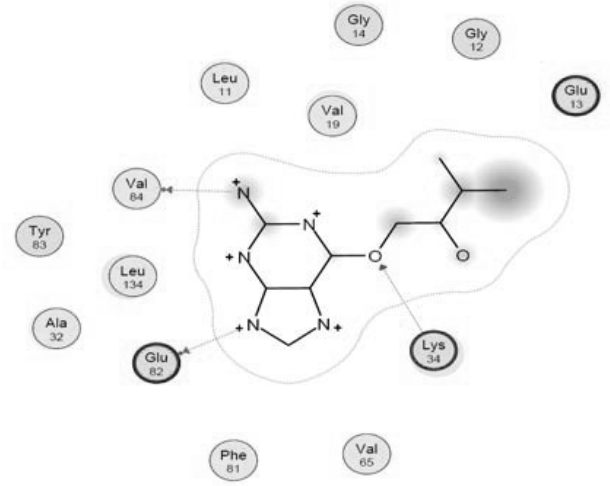


Figure 4: Superposition of active site residue within 3.5 region surrounding butoxypurine inhibitors in model CRK3 (yellow) and crystal structure of 1GZ8 (purple).

Most of the residues are conserved but few important variations do occurs. Phe82 in human CDK2 active site replaced by Tyr83 in leishmania CRK3 while Leu83 by Val84. These relative differences in the active site should undoubtedly affect the steric and electrostatic interactions with inhibitors, which would be employed for development of leishmania specific CRK3 inhibitors. The model butoxypurine inhibitor in leishmania CRK3 shows same interaction as in human CDK2. Study on CRK3-butoxypurine complex shows that purine ring bind in the ATP binding cleft forming a triplet of hydrogen bonds between residues in the hinge (Glu82, Val84) and N2, N3 and N9 atoms of the inhibitor. The purine ring is sand witched between N and C-terminal domain forming number of hydrophobic contacts.



5a 5b

Figure 5: Interactions of butoxypurine inhibitor in CRK3 (5a) and 1GZ8 (5b) active sites

Experimental Section

Homology modeling. Leshmania donovani CRK3 sequence was retrieved from UniProtKB/TrEMBL database (primary accession number 015851). In ordered to find out homologues sequence with known three dimentional structure BLASTP search were carried out against protein data bank (PDB) using leishmania CRK3 as query sequence. The top two hits from BLASTP viz. 1GZ8 and 2IW8 has a 58% sequence identity with leishmania CRK3. Sequence alignment was done using 'Roundrobin iterative refinement option' and 'tree based alignment approach' in MOE-Align module. BLOSUM62 matrix was used and all the templates were aligned against leishmania CRK3 sequence. Both templates shows good alignment with leishmania CRK3 with few gap in 1GZ8 sequence aligning with CRK3. Since the crystal structure resolution of 1GZ8 is good and also the butoxypurine inhibitor which is present in 1GZ8 bears significant identity to the known reported inhibitors for leishmania CRK3 we used it as primary templates while 2IW8 used as secondary template to model those regions in CRK3 which are align with gap in primary template.

All computations and molecular modeling of leishmania CRK3 were carried out on a Silicon Graphics Fuel Work station with IRIX 6.5 operating system using the MOE-03 (Molecular Operating Environment), MODELLER and SYBYL7.1 molecular modeling packages.

Homology modeling module of the MOE was used for comparative protein modeling.

In alignment shown Met1 in 1GZ8 aligns with Ser1 in query sequence and Glu2 aligns with Leu2 in CRK3. Ideally Met1, Glu2 in 1GZ8 should align with Leu2 and Glu3 of leishmania CRK3 respectively. To overcome this we given residue constrain so that metheonine1in 1GZ8 aligns with Lue2 in CRK3 and Glu2 with Asp3.

Initially crude homology model for CRK3 obtained by using Homology modeling module of MOE. Employing Coarse model refinement 10 independent models were constructed by the Boltzmann- weighted randomized modeling procedure. Each of these intermediate models was subjected to a sufficient degree of minimization to relieve any serious steric clashes. The 11th model was generated as an average of all these 10 models. The model with the best packing quality was chosen and subjected to minimization steps.

Generated model from MOE further subjected to ligand modeling in MODELLER which is a command-line tool and has no graphical interface; need to be provided with python script file containing MODELLER commands. For modeling ligand based on template it can be assumed that the ligand binding modes are similar in the target and the template protein.

Accordingly, ligands are then transferred among targets structures keeping their orientation as a restraint for the subsequent modelling process. Having placed the ligand in a near-native orientation into the consensus binding-site of the modelled protein, new models are generated by additionally incorporating information about this ligand. During this modelling step, the ligands are kept fixed in space. The presence of the ligand is included into the homology modelling process in terms of user-defined restraints. Interactions between butoxypurine inhibitor and active site amino acids residues as stated previously given as restraints during modeling protein-ligand complex. Having generated 10 ligand-supported models their quality was assessed using Ramachandran plot in PROCHECK validation package.

The best ligand-supported model based on PROCHECK validation further subjected to loop modeling in MODELLER using ERRAT plot as guidance for selecting a region for loop modeling. In ERRAT plot those regions which have an error values cutoff greater than 95% were selected for loop modeling one by one. After each round of loop modeling in MODELLER 10 different structures were generated which are again validated based on PROCHECK and ERRAT plot evaluation and again loop modeling carried out. Finally model which shows good PROCHECK verification and ERRAT plot subjected to native protein fold evaluation using Prosa II 2003.

Conclusions

Leishmania CRK3 is a potential target for treating visceral leishmaniasis. In ordered to understand structural features, we carried out homology modeling of leishmania *donovani* CRK3. Developed model showed good overall structural quality and is validated using different validation tools like PROCHECK, ERRAT plot. In addition model showed good native protein folding as assessed by using Prosa II program. The superposition of active site of 1GZ8 and CRK3 showed minor variations like Phe82 to Tyr83 and Lue83 to Val84 substitutions, which could be exploited in the future design of target specific inhibitors. Thus developed CRK3 model can be used for screening a chemical library for their CRK3 inhibitory activity.

Future Perspectives

In future we try to exploit specific structural features of leishmania CRK3 in comparison with human CDK2 (1GZ8) using different tools like molecular electrostatic potential and cavity depth analysis. Further in order to show acceptability of developed homology model for screening CRK3 specific inhibitors, and to understand its key interactions docking analysis would be carried out.

References

[1] M.A. Vannier-Santos, A. Martiny and W. de Souza. (2002) Cell biology of Leishmania spp.: invading and evading, *Curr. Pharm.Design*. 8, 297-318.

[2] P. Desjeux (1998) UNAIDS. Leishmania and HIV in Gridlock. WHO/CTD/LEISH/98.9 Add. I, UNAIDS/98.23. 1998. Geneva: WHO, 1998.

[3] K.P. Chang. (1983) Cellular and molecular mechanisms of intracellular symbiosis in leishmaniasis, *Int. Rev. Cytol. Suppl.* 14, 267-305.

[4] S.L. Croft and G.H. Coombs. (2003) Leishmaniasis-current chemotherapy and recent advances in the search for novel drugs, *Trends Parasitol.* 19, 502-508.

[5] H.W. Murray. (2004) Progress in the treatment of a neglected infectious disease: visceral leishmaniasis, *Expert. Rev. Anti. Infect Ther.* 2, 279-292.

[6] K.M. Grant, P. Hassan, J.S. Anderson and J.C. Mottram. (1998) The crk3 Gene of Leishmania mexicana Encodes a Stage-regulated cdc2-related Histone H1 Kinase That Associates with p12 cks1, *J Biol Chem.* 273, 10153-10159.

[7] C. Doerig, L. Meijer and J.C. Mottram (2002) Protein kinases as drug targets in parasitic protozoa, *Trends Parasitol*. 18, 366-371.

[8] D.H. Walker. (1998) Small-molecule inhibitors of cyclin-dependent kinases: molecular tools and potential therapeutics, *Curr. Top. Microbiol. Immunol.* 227, 149-165.

[9] K.M. Grant, M.H. Dunion, V. Yardley, A.L. Skaltsounis, D. Marko, G. Eisenbrand, S.L. Croft, L. Meijer and J.C. Mottram. (2004) Inhibitors of Leishmania mexicana CRK3 Cyclin-Dependent Kinase: Chemical Library Screen and Antileishmanial Activity, *Antimicrob. Agents. Chemother.* 48, 3033-3042.

[10] M.A. Marti-Renom, A.C. Stuart and A. Fiser. (2000) Comparative Protein Structure Modeling OF Genes AND Genomes, *Annu. Rev. Biophys. Biomol*. Struct. 29, 291 325.

[11] R. Laskovaski, MacArthur, M., Moss D. and Thornton, J. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Cryst.* 26, 283-291.

[12] M.J. Sippl. (1993)Recognition of errors in three-dimensional structures of proteins, Proteins: *Struct. Funct. Genet.* 17, 355–362.