Print ISSN: 0976-2876 Online ISSN: 2250-0138

Available online at: http://www.ijsr.in

INDIAN JOURNAL OF SCIENTIFIC RESEARCH

DOI:10.32606/IJSR.V12.I1.00010

Received: 08-05-2021 Accepted: 23-07-2021 Publication: 31-08-2021

Indian J.Sci.Res. 12 (1): 55-60, 2021 Original Research Article

IN SILICO ANALYSIS OF *Leishmania donovani* GLYCOLYTIC ENZYME

ANKITA SRIVASTAVA[a1] AND A.A. MAHDI[b]

[ab]Department of Biochemistry, King George's Medical University, Lucknow, Uttar Pradesh, India

ABSTRACT

Pyruvate kinase (PK) is an important enzyme of glycolytic pathway and significant difference between *Leishmania* and human PK makes it an ideal drug target for structure based drug design against *Leishmania*. A homology modeling was performed in order to create three dimensional structure of *Leishmania donovani* PK using MODELLER v9.9. Further quality of structure was validated using PROCHECK and ERRAT programs which confirm that the structure is reliable. The PDB file of verified model was submitted to Protein Model Database (PMDB; ID PM0079479. This model provides the information about binding/active regions, helpful in drug designing against *L. donovani*.

KEY WORDS: *Leishmania*, Pyruvate Kinase, Drug Target, Modeling, Drug Design, Protein, PyMOL

Pyruvate kinase is a putative protein of molecular weight 366kDa. The parasitic protozoa *L. donovani* belong to the order kinetoplastida. This zoonotic organism is the causative agent of deadly *Leishmania*sis, also known as Kala-azar. The Leishmaniasis is a wide spectrum of vector born disease with great epidemiological and clinical diversity. It is caused by more than 20 species of protozoan parasite (Mishra, 2009). Transmission occurs by the bite of insect vector, sandfly and around 30 species of sand fly are known to be vector for *Leishmania*. The annual global prevalence of all forms (i.e. cutaneous, mucocutaneous and visceral) of Leishmaniasis is nearly 10 million and approximately 350 million people are at risk (Desjeux, 2004). Additionally, in the developing countries like India, cases are often being reported from newer (non-endemic) areas and disease is occupying pandemic status due to population migration to non endemic regions though current statistical data are lacking (Desjeux, 2001). However, there is a gross under reporting of the cases from endemic regions and these figures may go up (Singh, 2006). Treatment options for visceral *Leishmania*sis include chemotherapy, which is not found satisfactory. The main reason for this failure is long drug course, high drug toxicity and parasitic resistance to drug such as antimonials drugs. Since the disease belongs to poor thus high cost of drugs make it unavailable in affected rural areas (Srivastava, 2012). Vaccine development has been still far from reality due to intricate mechanisms has developed for evading from immune system of host by different species of *Leishmania*. Fortunately the cell biology of *Leishmania* is very unusual and become a subject for interrogation in order to achieve effective remedy for VL. An ideal target for investigation in *Leishmania* is glycolytic pathway. Ten enzymes of glycolysis are responsible for catalyzing glucose into pyruvate and produce ATP. In *Leishmania* parasite pyruvate kinase (MW 366Kda) is the chief enzyme of the major energy generating pathway i.e Glycolysis. It catalyzes a phospho enol pyruvate to ADP and thus forming pyruvate and ATP. Recently the crystal structure of *L. mexicana* enzyme was determined (Crowther, 2001). This enzyme is homotetramer with each monomer possessing four domains. Human possess four isoenzymes of PK based on the location in the human body. The structural modeling using bioinformatics tools provides us about each and every aspects of secondary and tertiary structure of protein. For a protein to be a drug target, it is needed to know the active regions so that any small ligand molecule can effectively bind it.

MATERIALS AND METHODS

Template Selection

The protein sequence of *L. donovani* PK was obtained from NCBI (http://www.ncbi.nlm.nih.gov/) having a length of 499 amino acids. For selection of templates based on sequence similarity and identity, a BLAST search was executed using this sequence.

Sequence Alignment

Based on the sequence similarity and identity, PK sequence of various species of *Leishmania* along with *Trypanosoma* and human was aligned to observe the conservation and variations among various species. This was performed using CLUSTALX2 software.

[1]Corresponding author

Homology Modelling

A 3D structure of the target sequence i.e *L. donovani* PK was built using offline software MODELLER v9.9. It generates 5 similar iterative structures of the *L. donovani* PK which was based on its template structure and the alignment input file (filename.ali) in the PIR format. The best model was picked by choosing the lowest MODELLER dope score value mentioned in the PDB file of model.

Validation of Homology Model

The protein 3D model constructed by homology modelling technique could be prone to errors and thus imprecise. Therefore validating the quality of overall structure is important. It needs a careful investigation of the stereo chemical parameters and accuracy of folds at different levels. To structure was evaluated using various programs such as PROCHECK and ERRAT using online server SAVES (http://nihserver.mbi.ucla.edu/SAVES/). PROCHECK checks the stereochemical quality of protein structure by analyzing residue-by-residue geometry and overall geometry and a Ramachandran plot is produced. ERRAT is used for statistical analysis of non-bonded interactions between different atom types. A plot is generated using the value of the error function versus position of a 9-residue sliding window which is calculated by comparing the statistics from highly refined structures (Luthy *et al.*, 1992). The three dimensional structure of PK was viewed and analyzed using PyMol software (Colovos and Yeates, 1992).

Structure Scanning

The structure was further scanned for presence of any unusual patches using Hotpatch web gateway version 4 (http://hotpatch.mbi.ucla.edu/).

RESULTS

Alignment and Homology Modeling of PK

The details of sequence alignment of *L. donovani* and *H. sapiense* are given in the figure 1. On comparing these *Leishmania* PK with Human PK, most of the amino acid residues are more or less conserved. However, variation in amino acids residues at some places makes the parasitic PK dissimilar from Human PK.

The input amino acid sequences of *Leishmania* PK generate five structures having their dope score. Among five, the best structure was chosen based up on the dope score. The details of all five structures were given in table 1. As per the values in the table, B99990003.pdb was found the best structure. The three dimensional view was shown in figure 2.

Structure Validation

The PROCHECK analysis shows that the 94% residues (420) are in most favored region including 5.2% (23) in additionally allowed region. Moreover 0.2% residue (1) is found in generously allowed region. No residue is found at disallowed region (figure 3 and figure 4).

Table 1: The dope scores and mol pdf values of all five structures generated by Modeller

Structures files	Mol. pdf	Dope score	GA341 score
B99990001.pdb	2242.64062	-56733.19141	1.0
B99990002.pdb	2432.85767	-57059.92579	1.0
B99990003.pdb	2270.05176	-57213.60156	1.0
B99990004.pdb	2293.74976	-56856.50391	1.0
B99990005.pdb	2318.88818	-57035.29297	1.0

```
                 ....|....|    ....|....|    ....|....|    ....|....|    ....|....|    ....|....|
                      5              15              25              35              45              55
L.donovani       ----------- --MSQLAHNL TLSIF----- ---------- ---------E PVANHRATRI
H.sapiens        MEGPAGYLRR ASVAQLTQEL GTAFFQQQQL PAAMADTFLE HLCLLDIDSE PVA-ARSTSI
Clustal Co                   ::**:::*   ::*                             * ***  *:* *

                 ....|....|    ....|....|    ....|....|    ....|....|    ....|....|    ....|....|
                      65              75              85              95             105             115
L.donovani       VCTIGPSTQS VEALKGLIQS GMSVARMNFS HGSHEYHRTT INNVRQAAAE LGVN------
H.sapiens        IATIGPASRS VERLKEMIKA GMNIARLNFS HGSHEYHAET IANVREAVES FAGSPLSYRP
Clustal Co       :.****:::* ** ** :*:: **.:**:*** *******  * * ***:*. . :. .

                 ....|....|    ....|....|    ....|....|    ....|....|    ....|....|    ....|....|
                     125             135             145             155             165             175
L.donovani       IAIALDTKGP EIRTGQFVGG ---EAVMERG ATCYVTTDPA FADKGTKDKF YIDYQNLSKV
H.sapiens        VAIALDTKGP EIRTGILQGG PESEVELVKG SQVLVTVDPA FRTRGNANTV WVDYPNIVRV
Clustal Co       :********** ***** : **     *. : :* :    **.*** *  :*. :.. ::** *: :*

                 ....|....|    ....|....|    ....|....|    ....|....|    ....|....|    ....|....|
                     185             195             205             215             225             235
L.donovani       VRPGSYIYID DGILILHVQS HEDEQTLKCT VTNAHTISDR RGVNLPGCDV DLPAVSAKDC
H.sapiens        VPVGGRIYID DGLISLVVQK ISPE-GLVTQ VENGGVLGSR KGVNLPGAQV DLPGLSEQDV
Clustal Co       *  *. **** **:: * **.  . *  *     * *. .:..* :******.:* ***.:* :*

                 ....|....|    ....|....|    ....|....|    ....|....|    ....|....|    ....|....|
                     245             255             265             275             285             295
L.donovani       ADLQFGVEQG VDMIFASFIR SAEQVGEVRE ALGAKGRDIM IICKIENHQG VQNIDSIIEE
H.sapiens        RDLRFGVEHG VDIVFASFVR KASDVAAVRA ALGPEGHGIK IISKIENHEG VKRFDEILEV
Clustal Co         **:****:* **::****:* .*.:*. **   ***.:*:.*  **.*****:* *:.:*.*:*

                 ....|....|    ....|....|    ....|....|    ....|....|    ....|....|    ....|....|
                     305             315             325             335             345             355
L.donovani       SDGIMVARGD LGVEIPAEKV VVAQKILISK CNVAGKPVIC ATQMLESMTY NPRPTRAEVS
H.sapiens        SDGIMVARGD LGIEIPAEKV FLAQKMMIGR CNLAGKPVVC ATQMLESMIT KPRPTRAETS
Clustal Co       ********** **:******* .:***::*.: **:*****:* ********   :********.*

                 ....|....|    ....|....|    ....|....|    ....|....|    ....|....|    ....|....|
                     365             375             385             395             405             415
L.donovani       DVANAVFNGA DCVMLSGETA KGKYPNEVVQ YMARICLEAQ SAVNEYVFFN SIKKLQPIPM
H.sapiens        DVANAVLDGA DCIMLSGETA KGNFPVEAVK MQHRIAREAE AAVYHRQLFE ELRRAAPLSR
Clustal Co       ******::** **:******* **::* *.*:     **. **: :** .  :*: .:::   *:.

                 ....|....|    ....|....|    ....|....|    ....|....|    ....|....|    ....|....|
                     425             435             445             455             465             475
L.donovani       SAEEAVCSSA VNSVYETKAK VMVVLSNTGR SARLVAKYRP NCPIVCVTTR LQTCRQLNIT
H.sapiens        DPTEVTAIGA VEAAFKCCAA AIIVLTTTGR SAQLLSRYRP RAAVIAVTRS AQAARQVHLC
Clustal Co       .. *... .* *::.::  *  .::***:.*** **:*:::*** ...::.**     *:.**:::

                 ....|....|    ....|....|    ....|....|    ....|....|    ....|....|    ....|....|
                     485             495             505             515             525             535
L.donovani       QGVESVFF-D AEKLGHDEGK EQRVAMGVGF ATSKGYVQTG DYCVVIHADH KVKGYANQTR
H.sapiens        RGVFPLLYRE PPEAIWADDV DRRVQFGIES GKLRGFLRVG DLVIVVTGWR PGSGYTNIMR
Clustal Co       :** .::: : . :    :.  ::** :*:   .. :*::::.* *  :*: . :   .**:*  *

                 ....|
                     545
L.donovani       ILLVE
H.sapiens        VLSIS
Clustal Co       :* :.
```

Figure 1: Alignment of amino acid sequences of Pyruvate kinase (PK) of *L. donovani* and *H. sapiense.* * indicates the conserved sites; - represents a gap introduced for alignment optimization

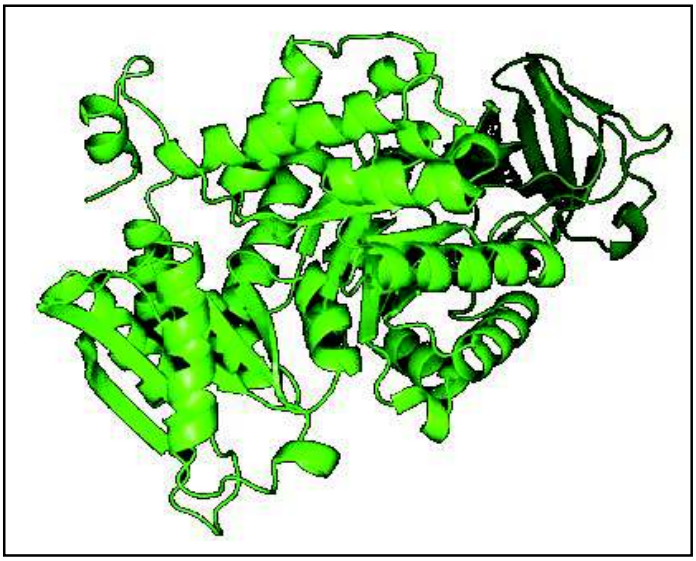

Figure 2: The three dimensional view of modeled structure of *Leishmania* PK (PyMOL view)

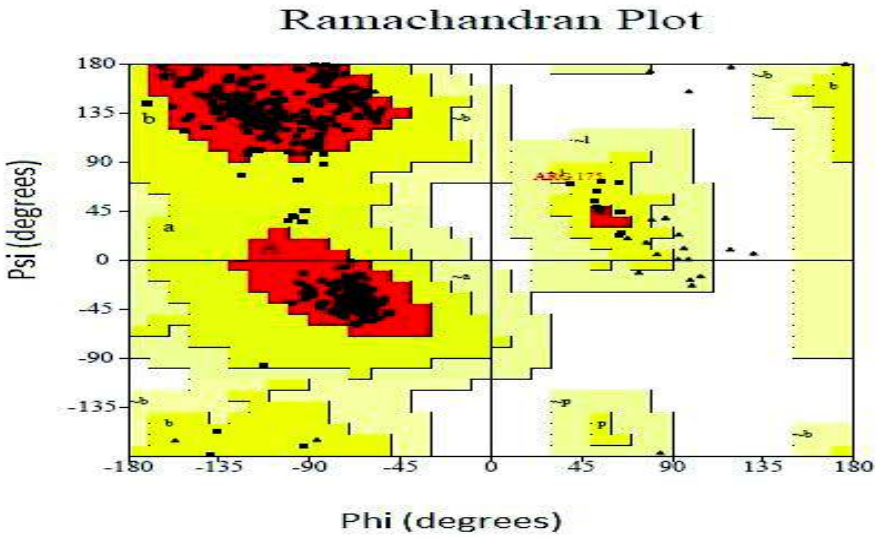

Figure 3: Ramachandran plot of *Leishmania* PK structure showing the positions of amino acids

Plot statistics		
Residues in most favoured regions [A,B,L]	420	94.6%
Residues in additional allowed regions [a,b,l,p]	23	5.2%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.2%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	444	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	37	
Number of proline residues	16	
Total number of residues	499	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20% a good quality model would be expected to have over 90% in the most favoured regions.

Figure 4: Details of PROCHECK analysis of modeled structure of *Leishmania* PK

ERRAT statistics shows that the overall quality factor for the modeled structure was found to be 88.187. The ERRAT plot is shown in figure 5 and figure 6.

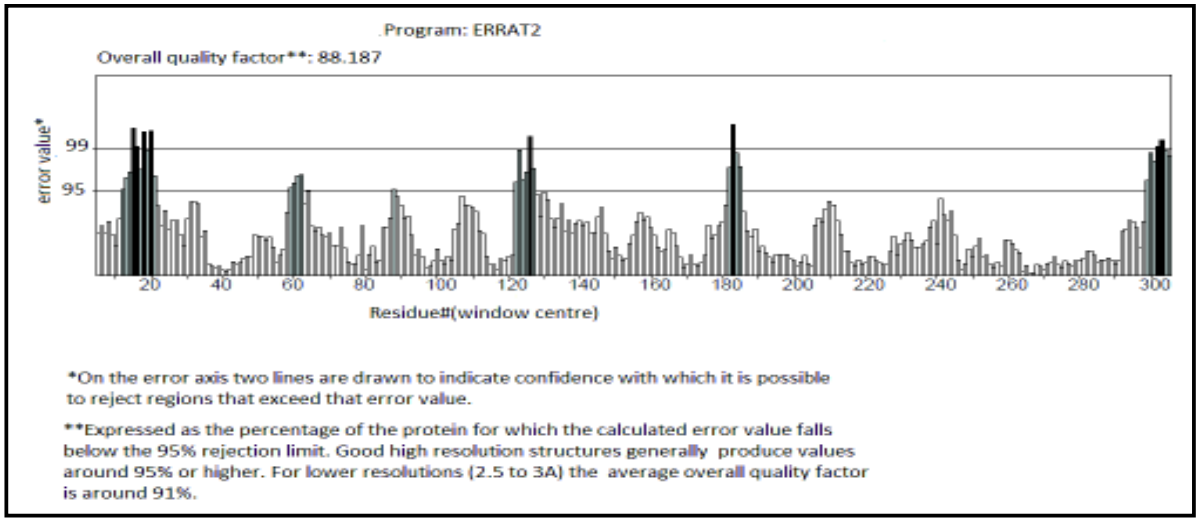

Figure 5: ERRAT plot showing the error value per residue- residue number 1-300

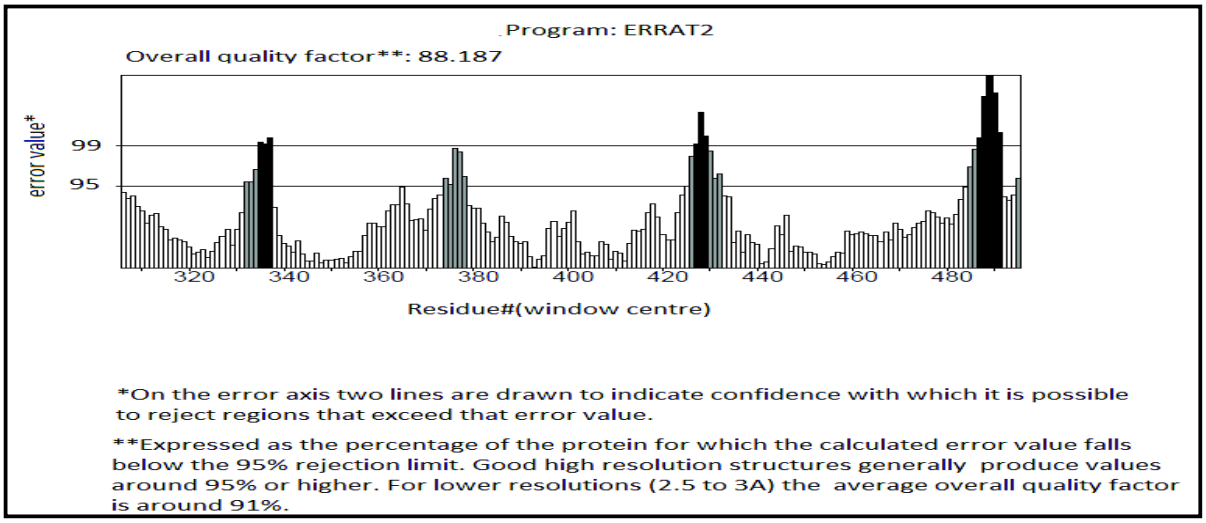

Figure 6: ERRAT plot showing the error value per residue- residue number 301-499

Further, the tertiary structure was scanned for presence of any active/binding sites. The result shows that there are 7 residues are engaged as unusual patches as metal ion binding site of large database of proteins (figure 7) whereas 12 residues are also present in a functional area as small molecule binding sites (figure 8).

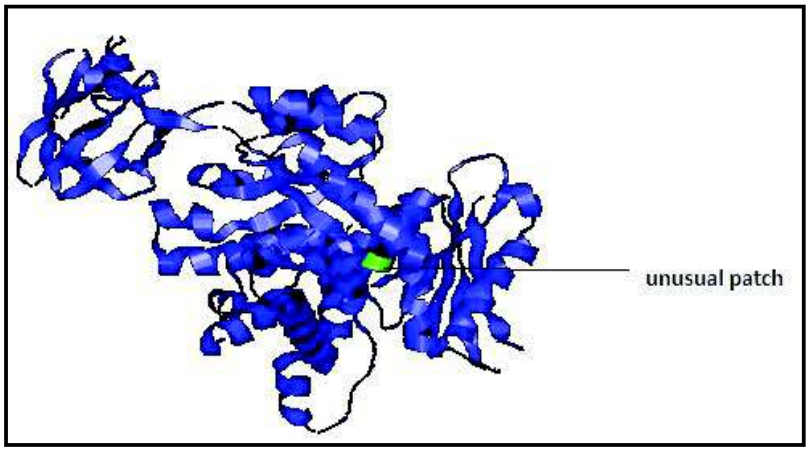

Figure 7: *Leishmania* PK structure (RasMol view) showing one unusual patch(green) i.e metal ion binding site

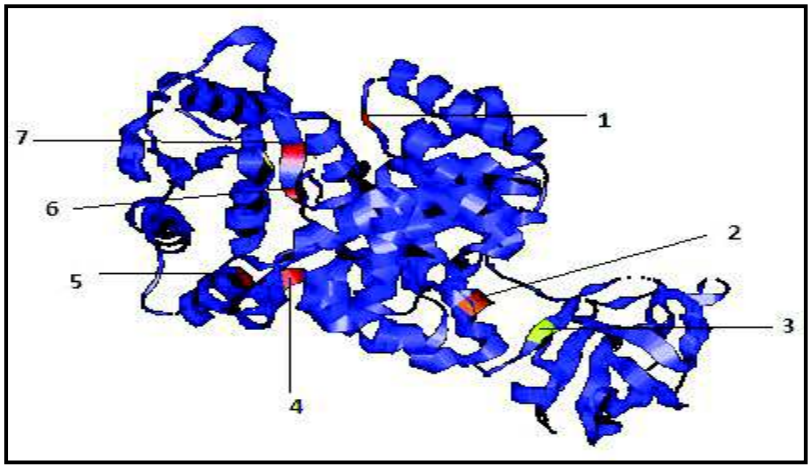

Figure 8: *Leishmania* PK structure (RasMol view) showing unusual patches(1-7) (red, yellow and orange) i.e small molecule binding site

MODEL AVAILABILITY

The PDB file of homology model of *L. donovani* PK was submitted to Protein Model Data Base (http://www.caspur.it/PMDB) and an ID has been assign to this structure. This model is now available at PMDB website with ID PM0079479.

DISCUSSION

*Leishmania*sis is a major public health problem and there are no effective vaccine is available till date. The disease control strategy is not very reliable as well as cheap to make its reach to every needy person. However, the present repertoire of drugs is very restricted and increasing drug resistance has posed a major concern. The first step in drug discovery is to identify a suitable drug target. As we have known a lot from the genome sequences of *Leishmania* which gives opportunity to identify novel drug targets that are unique to these parasites. Since PK is an important enzyme of the glycolytic pathway but difference between human PK and l*eishmania* is evident which can be observed in clustal sequence analysis of human PK and *leishmania* PK. The sites which shows * are conserved but there are multiple sites which are not conserved as shown in figure 1. These unconserved sites could be probed for antileishmanial treatment. A three dimensional modeled structure of *Leishmania* PK is produced (figure 2) and analyzed. From the PROCHECK analysis (figure 3), it is clearly evident that out of 499 amino acid residues, 420 amino acid residues are lying in favored region, 23 amino acid residues are in additionally allowed region and no amino acid is in disallowed region (figure 4). Thus stereo chemically structure is stable. Further, the ERRAT plot analysis also shows a quality factor of 88.187 (figure 5 and figure 6) which is satisfactory for the structure. The detailed surface analysis of three dimensional structure of PK by Hotpatch gateway server has shown that it has an unusual amino acid patch which is metal ion binding site and can be visualized in RasMol (figure 7). Similarly few small molecule binding sites can be viewed in RasMol view of PK (figure 8). These patches or binding sites may be a functional site and needs further probing. Thus we can conclude that the modeled homology structure is good enough to further drug designing studies. Utilization of this information can discover a new candidate drug molecule. After extensive research in this area, Pyruvate kinase has been evolved as a drug target against *Leishmania* (Crowther, 2001). Since glycolysis is the major energy generating pathway for *Leishmania* thus inhibition of PK may lead to arrest of glycolysis and subsequently cell death due to lack of energy. Thus modeling of PK would lead to further *in silico* enzyme inhibition studies.

REFERENCES

Colovos C. and Yeates T.O., 1992. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci., **2**:1511–1519.

Crowther G.J., Shanmugam D., Carmona SJ., Doyle M.A., Hertz-Fowler C., Berriman M., Nwaka S., Ralph S.A., Roos D.S., Van Voorhis W.C. and Fernan Aguero, 2010. Identification of Attractive Drug Targets in Neglected-Disease Pathogens Using an In Silico Approach, 2001. Plos Negl Trop Dis., **8**(4): 1-18.

Desjeux P., 2001. The increase in risk factors for leishmaniasis worldwide. Trans R. Soc Trop. Med. Hyg., **95**:239-243.

Desjeux P., 2004. Leishmaniasis: current situation and new perspectives. Comp Immunol Microbiol Infect Dis., **27**: 305-318.

Hooft R.W.W., Vriend G., Sander C. and Abola E.E., 1996. Errors in protein structures. Nature, **381**:272.

Laskowski R.A., MacArthur M.W., Moss D.S. and Thornton J.M., 1993. PROCHECK : a program to check the stereo chemical quality of protein structure. J. Appl. Crystallogr., **26**:283–291.

Luthy R., Bowie J.U. and Eisenberg D., 1992. Assessment of protein models with three-dimensional profiles. Nature, **356**:83–85.

Mishra B.B., Singh R.K., Srivastava A., Tripathi V.J. and Tiwari V.K., 2009. Fighting against Leishmaniasis: search of alkaloids as future true potential anti-Leishmanial agents. Mini. Rev. Med. Chem., **9**: 107-123.

Singh S.P., Reddy D.C., Rai M. and Sundar S., 2006b. Serious underreporting of visceral leishmaniasis through passive case reporting in Bihar, India. Trop Med. Int. Health, **11**: 899-905.

Srivastava A., Singh N., Mishra M., Kumar V., Gour J.K., Bajpai S., Singh S., Pandey H.P. and Singh R.K., 2012. Identification of TLR inducing Th1-responsive *Leishmania donovani* amastigote-specific antigens. Mol. and cell Biochem., **359**(1-2): 359-368.

The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.