Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Web Database**

A database for *Plasmodium falciparum* protein models

Ramasamy Gowthaman[1], Duraisamy Sekhar[2], Mridul Kumar Kalita[1] and Dinesh Gupta[1*]

[1]Structural and Computational Biology Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), Aruna Asaf Ali Marg, New Delhi 110067, India; [2]Current address: Dana-Farber Cancer Institute, Harvard Medical School, Dana-830, 44-Binney street, Boston, MA-02115, USA;

Dinesh Gupta* - Email: dinesh@icgeb.res.in; Phone: +91 11 2617 3184; Fax: +91 11 2616 2316; * Corresponding author

received September 21, 2005; revised October 3, 2005; accepted October 3, 2005; published online October 5, 2005

Abstract:

There is an urgent need for developing alternate strategies to combat Malaria caused by *Plasmodium falciparum (P. falciparum)* because of growing drug resistance and increased incidents of infection in humans. 3D models of *P. falciparum* annotated proteins using molecular modeling techniques will enhance our understanding about the mechanism of host parasite interactions for the identification of drug targets and malarial vaccine design. Potential structural templates for *P. falciparum* annotated proteins were selected from PDB (protein databank) using BLASTP (basic local alignment search tool for proteins). This exercise identified 476 *Plasmodium* proteins with one or more known structural templates (≥ 40% identity) for further modeling. The pair-wise sequence alignments generated for protein modeling were manually checked for error. The models were then constructed using MODELLER (a comparative protein modelling program for modelling protein structures) followed by energy minimization in AMBER force field and checked for error using PROCHECK.

Availability: http://bioinfo.icgeb.res.in/codes/model.html

Keywords: molecular modeling; protein templates; comparative modeling; profile based alignment for modeling

Background:

The *P. falciparum* genome has been published. **[1]** However, known structures for *P. falciparum* proteins are limited. **[2]** 3D structural information for Plasmodium proteins is critical for the identification of suitable drug targets. The use of protein modeling techniques produces modeled structures for annotated Plasmodium proteins. Nonetheless, automatic modeled proteins are error prone and further validation is required. Therefore, modeling of Plasmodium proteins is performed in a semi-automated manner with a series of manual error checking. A database containing the Plasmodium protein models is described in this article.

Methodology:

Plasmodium protein sequences

The annotated protein sequences for the *P. falciparum* genome were obtained from PlasmoDB (http://PlasmoDB.org). **[3]**

Molecular modeling:

Modeling of target proteins was performed using MODELLER (a comparative protein modelling program for protein structure modeling). **[4]**

Plasmodium proteins (target) and template selection for modeling:

P. falciparum annotated proteins with known structural templates in PDB (protein databank) were identified using BLASTP (basic local alignment search tool for proteins). This exercise selected 476 *Plasmodium* proteins with one or more known structural templates (≥ 40% identity) for modeling.

Alignment of target and template sequences:

The alignment of target and template sequence was performed using align2D (a sequence alignment module in MODELLER). Align2D performs global alignment of sequences using dynamic programming. This program uses variable gap penalty for structural loops and core regions using information derived from template structures. **[4]** Subsequently, the alignment was manually checked for mismatch.

Energy refinement:

The modeled structures were energy minimized in AMBER force field for 500 steps using steepest descent and conjugate gradient minimization algorithms. **[5]** The stereo-chemical qualities of the generated protein models were assessed using PROCHECK. **[6]** Models were then superimposed with the templates using SWISS-PDB viewer and RMSD (root mean square deviation) was calculated for each model. **[7]** The modeled structures thus constructed were stored in a public domain database and the models can be viewed visualization tools such as RasMol **[8]** and SwissPDBviewer and MDL/chime. **[9]**

Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Web Database**

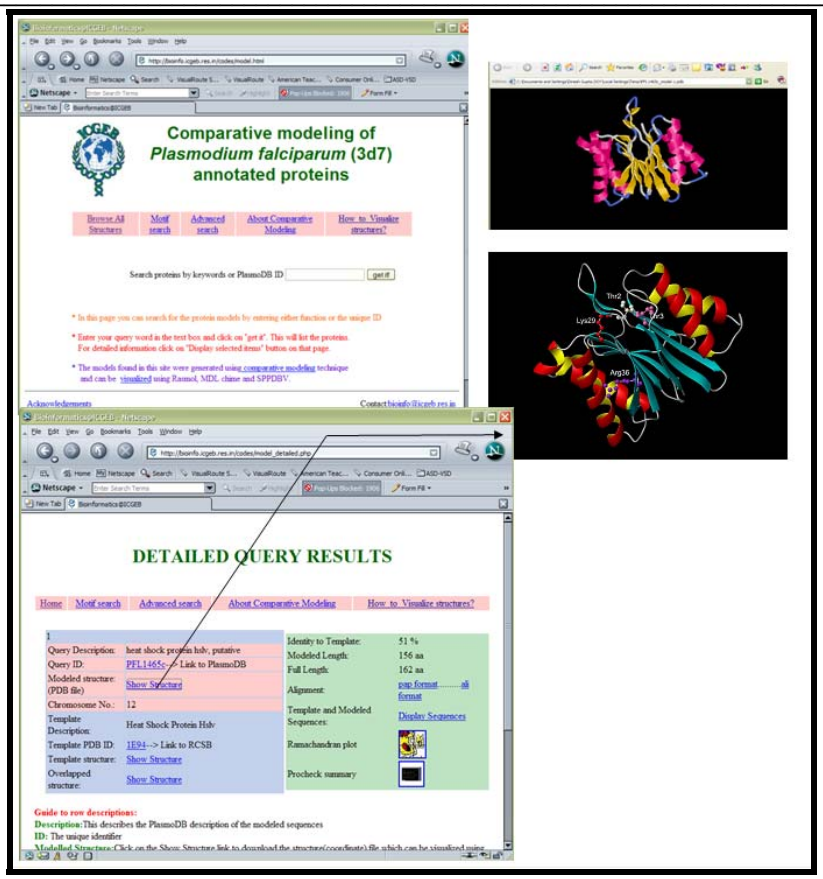

Figure 1: A database for *Plasmodium falciparum* protein models. The database snapshot and visual models are shown

Utility to the biological community:
This database is a collection of *P. falciparum* protein models. All the models were manually curated and verified. Comparison of host and parasite protein structures will provide insights into host-parasite interaction for the identification of potential drug targets.

References:

[1] M. J. Gardner, *et al*., *Nature*, 419:498 (2002) [PubMed: 12368864].
[2] http://www.rcsb.org/pdb
[3] J. C. Kissinger, *et al., Nature,* 419:490 (2002) [Pubmed: 12368860]
[4] A. Sali, *et al., J. Mol. Biol.,* 234:779 (1993) [Pubmed: 8254673]
[5] D. A. Case, *et al*., AMBER (Assisted Molecular Building with Energy Refinement) University of California (1999)
[6] http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html
[7] http://www.expasy.org/spdv/
[8] http://www.bernstein-plus-sons.com/software/RasMol_2.7.2.1.1/README.html
[9] http://www.mdlchime.com/

Edited by M.K. Sakharkar
Citation: Gowthaman *et al.,* Bioinformation 1(2): 50-51 (2005)