J Mol Model (2015) 21:37
DOI 10.1007/s00894-015-2586-4

ORIGINAL PAPER

Homology modeling and molecular dynamics simulation of N-myristoyltransferase from *Plasmodium falciparum*: an insight into novel antimalarial drug design

Paulomi Paul · Abhishek Chowdhury · Anupam Das Talukdar · Manabendra Dutta Choudhury

Received: 13 December 2014 / Accepted: 19 January 2015

Abstract Malaria is an infectious disease caused by parasites of the genus *Plasmodium*. It leads to approximately 1 million deaths per annum worldwide, with an increase number of 6.27 million deaths in 2012 alone. Validation of new antimalarial targets is very important in the context of the rise in resistance to current drugs. One such putative target is the enzyme N-myristoyltransferase (NMT), which catalyzes the attachment of the fatty acid myristate to protein substrates (N-myristoylation) for activation. Reports suggests that NMT is an essential and chemically docile target in malaria parasites both in vitro and in vivo, and the selective inhibition of N-myristoylation leads to irreversible failure to form an inner membrane complex—an essential subcellular organelle in the parasite life cycle. In this work, we modeled the three-dimensional structure of *Plasmodium falciparum* NMT (PfNMT) using Modeler 9.0 taking *Plasmodium vivax* NMT (PvNMT) as the template. The novelty of the work lies in the selection of template as the similarity of PfNMT with PvNMT was 80.47 %, whereas earlier similar work showed template similarity with *Candida albicans* NMT (CaNMT) and *Saccharomyces cerevisiae* NMT (ScNMT) to be less than 50 %. The generated structure was then validated using various programs such as PROCHECK, RAMPAGE server, CHIMERA and the stability of the model was checked by Gromacs 5.0.

Keywords *Plasmodium falciparum* NMT · pfNMT · N-myristoyltransferase · Malaria · *Plasmodium falciparum* · Homology modeling · Molecular dynamics

Introduction

Malaria is one of the most important global infectious diseases, affecting hundreds of millions of people each year. According to reports, in Africa a child dies from malaria almost every minute. The latest estimates released in December 2013 suggest that there were about 207 million new cases of malaria in 2012 and estimated deaths of 6.27 million [1]. The rise in cases of malaria may be due to the resistance of parasites to current drugs. Hence, there is an urgent need to chalk out new targets for the design of novel drugs. Literature surveys suggest that myristoyl-CoA protein N-myristoyltransferase (NMT) is a vital target in the malaria parasite but to date there are no reports of the development of any drugs targeting NMT.

The versatile parasite *Plasmodium falciparum* is responsible for causing malaria. The parasite breaks down hemoglobin in the host's red blood cells (RBC; hemolysis) and hence is very dangerous. Our work in searching for new drug targets is based on blocking the precursor proteins responsible for hemolysis in order to prevent growth of the parasite and thus avoid host death. NMT is a cytosolic enzyme that is ubiquitous in eukaryotes [2]. *P. falciparum* NMT (PfNMT) is responsible for the sexual blood stages of the parasite and is essential for transmission. NMT follows a bi–bi catalytic reaction involving the co-translational transfer of the rare

Electronic supplementary material The online version of this article (doi:10.1007/s00894-015-2586-4) contains supplementary material, which is available to authorized users.

P. Paul · A. Chowdhury (✉) · A. D. Talukdar · M. D. Choudhury
Bioinformatics Centre, Assam University, Silchar, Assam 788011, India
e-mail: abhishek@bioinfoaus.ac.in

Published online: 07 February 2015

SeqA	Name	Length	SeqB	Name	Length	Score
1	se1	410	2	4a95	384	80.47
1	se1	410	3	1iyl	392	38.78
1	se1	410	4	2p6g	455	40.73

Fig. 1 Sequence alignment of *Plasmodium falciparum* N-myristoyltransferase (NMT) (pfNMT; se1) with *Plasmodium vivax* NMT (PvNMT; 4a95), and NMT from *Candida albicans* (CaNMT; 1IYL) and *Saccharomyces cerevisiae* (ScNMT; 2P6G) using Clustal X 2.1 showing high similarity of PfNMT with PvNMT

cellular fatty acid myristate (C14:0) to the N-terminal glycine residue of myristoyl-CoA [3–5]. N-terminal myristoylation leads to major changes in essential properties of the protein like lipophilicity and thus promotes protein interactions with hydrophobic domains and membranes [6–9]. Inhibition of PfNMT may result in prevention of formation of the inner membrane complex that is an essential sub-cellular organelle in the parasite life cycle. Surprisingly, the three-dimensional (3D) structure of PfNMT protein is not available in any popular database, e.g., Protein DataBank (PDB). Hence, in this work, we aimed to design the 3D structure of PfNMT by homology modeling in order to analyze the characteristics and functional aspects of the protein to help the process of novel drug design. Similar work was reported by Sheng et al. [10], who modeled PfNMT using NMT from *Candida albicans* (CaNMT) and *Saccharomyces cerevisiae* (ScNMT) as templates. In the present work, PfNMT was modeled using *Plasmodium vivax* NMT (PvNMT), which is phylogenetically much closer to PfNMT than CaNMT and ScNMT (Fig. 1). The modeled structure was validated using various computational tools, and its stability was checked by molecular dynamics (MD) simulation.

Table 1 Parameters of the template selected for model generation showing highest quality. *PDB* Protein database, *NMT* N-myristoyltransferase

PDB ID	Identity	Scan method	Method	Resolution	Similarity	Coverage	Description
4A95	80.47 %	HHblits	X-ray	1.55 Å	0.56	0.94	*Plasmodium vivax* NMT

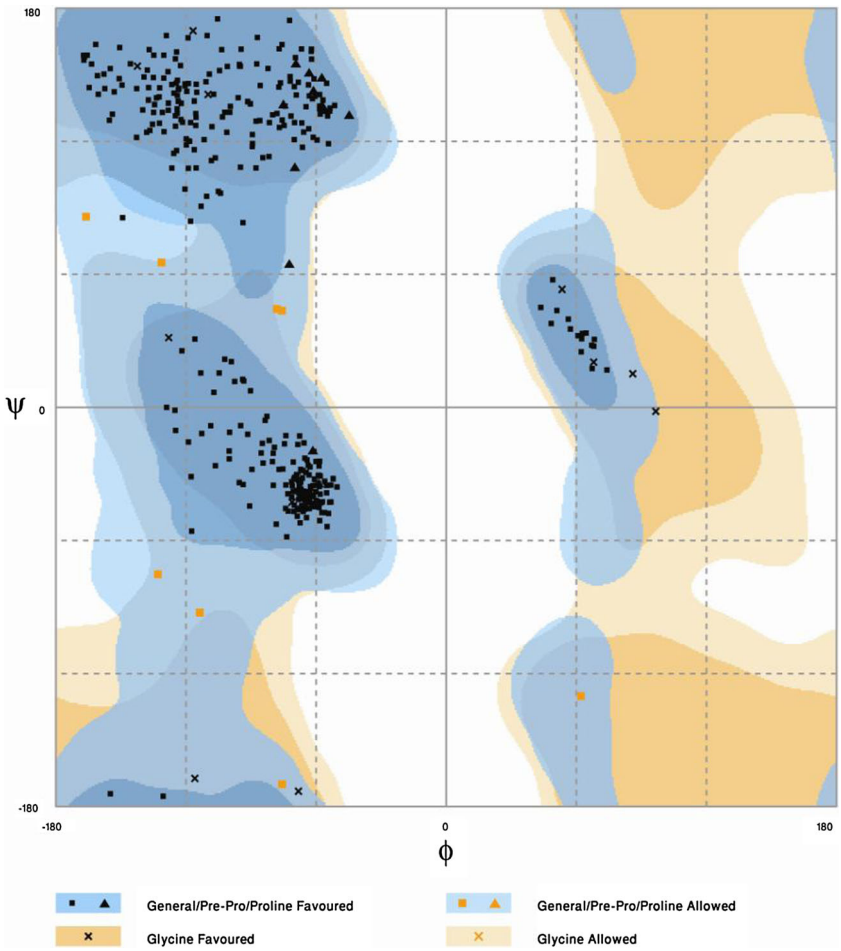

Fig. 2 Ramachandran plot of the modeled pfNMT protein generated by RAMPAGE, showing 97.9 % residues in the favorable range and the remaining 2.1 % in the allowed range

Methods

Sequence retrieval

The first step in homology modeling is to retrieve the amino acid sequence of a protein. The amino acid sequence of pfNMT was retrieved from the NCBI protein database (accession ID: AAF18461.1) in FASTA format and crosschecked with Uniprot.

Template selection

A template search with Blast and HHBlits (http://toolkit.tuebingen.mpg.de/hhblits) was performed against the SWISS-MODEL template library (SMTL). The templates with the highest quality were then selected for model building. The target sequence was searched with BLAST [11] against the primary amino acid sequence contained in the SMTL.

Model building

Models were built based on the target–template alignment using Promod-II [12]. Conserved coordinates between the target and the template were copied from the template to the model. Insertions and deletions were remodeled using a fragment library. Side chains were then rebuilt. Finally, the geometry of the resulting model was regularized using the CHARMM force field. Loop modeling with ProMod-II [12] did not give satisfactory results, hence an alternative model was built with MODELLER 9v9 [13]. The model includes a homology-based ab initio modeling.

Model quality estimation

The global and per-residue model quality was assessed using the QMEAN scoring function [14]. The generated models were then validated by PROCHECK [15]. The Ramachandran plot was generated with the RAMPAGE online portal [16].

Ligand modeling

Ligands present in the template structure were transferred by homology to the model when the following criteria were met: (1) the ligands were annotated as biologically relevant in the template library, (2) the ligand was in contact with the model, (3) the ligand did not clash with the protein, (4) the residues in

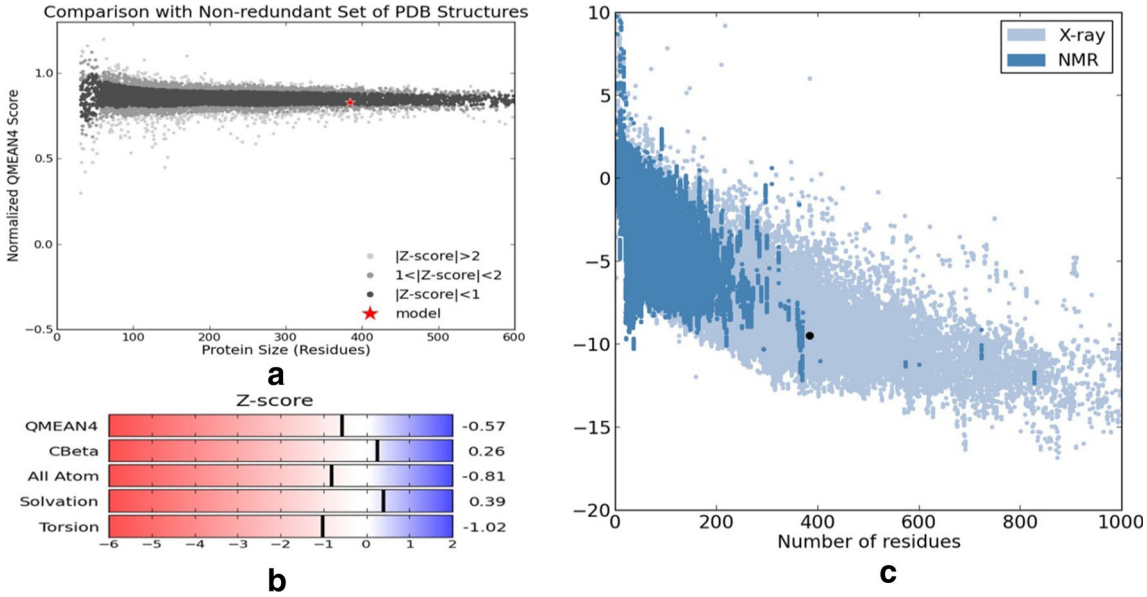

Fig. 3 Z-score of generated model showing good **a** QMEAN, **b** CBeta and **c** solvation values

contact with the ligand were conserved between the target and the template. If any of these four criteria was not satisfied, that ligand was not included in the model. The model summary includes information on which ligands were not included and why.

Oligomeric state conservation

The homo-oligomeric structure of the target protein was predicted based on the analysis of pairwise interfaces of the identified template structures. For each relevant interface between polypeptide chains, the QscoreOligomer [17] was predicted by similarity to target and frequency of observation. The oligomeric state of the target was predicted to be the same as in the template when the QscoreOligomer was predicted to be higher than or equal to 0.5.

Molecular dynamics simulation

MD simulations applied Gromacs 5.0 (http://www.gromacs.org/) to the best model in order to check stability. Energy profile, density, pressure, etc., were calculated after a 10-ns run in the simple point charge (SPC) water model.

Results

Primary amino acid sequence

The sequence retrieved from the NCBI protein database is as follows:

MNDDKKDFVGRDLYQLIRNAKDKIKIDYKFWYT-
QPVPKINDEFDENVNEPFISDNKVEDVRKEEYKLP-

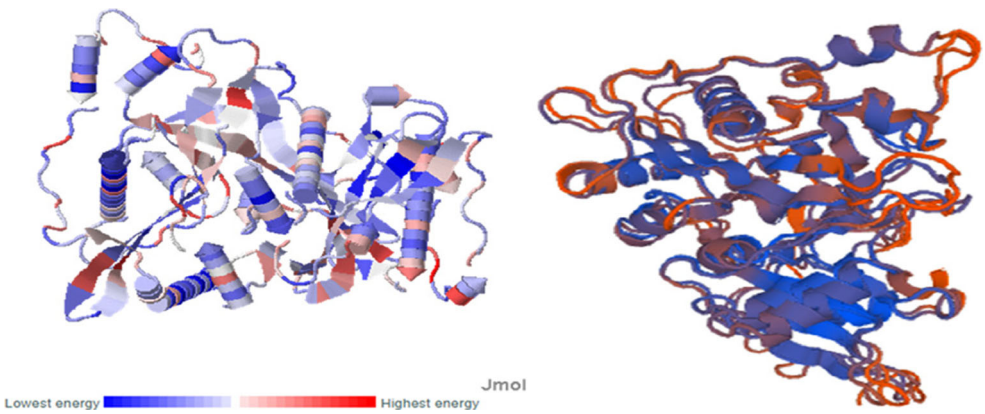

Fig. 4 Energy minimization of modeled pfNMT and structural changes after minimization

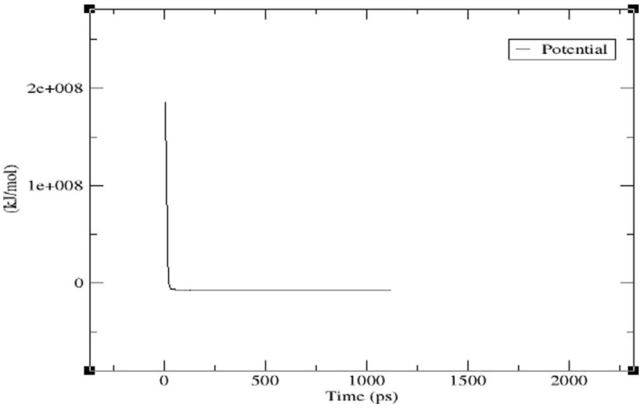

Fig. 5 Steepest descent minimization curve showing decline in potential energy using Gromacs 5.0. The potential energy declined and remained constant after 1,124 steps with potential energy of -7.508×10^{6}

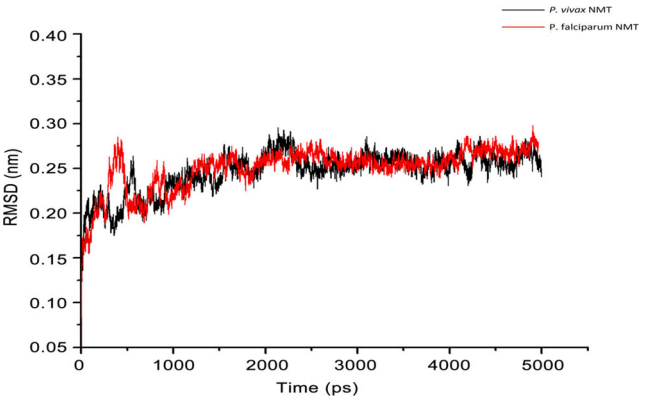

Fig. 7 Root mean square deviation (RMSD) curve of template and selected NMT model at 300 K in Gromacs 5.0 showing fluctuation from 0.23 to 0.27. The Gromacs manual states that such fluctuation is permissible for this protein at 300 K and proves the stability of the model

SGYAWCVCDITKENDRSDIYNLLTDNYVEDDDNV-
FRFNYSSEFLLWALSSPNYVKNWHIGVKYESTNK-
LVGFISAIPIDMCVNKNIIKMAEVNFLCVHKSLRSK-
RLAPVLIKEITRRINLESIWQAIYTAGVYLPKPISTAR
YFHRSINVKKLIEIGFSCLNTRLTMSRAIKLYRIDDT-
LNIKNLRLMKKKDIDGLQKLLNEHLKQYNLHAIF-
SKEDVAHWFTPIDQVIYTYVNEENGEIKDLISFYSL-
P S K V L G N N K Y N I L N A A F S F Y N I T T T T
TFKNLIQDAICLAKRNNFDVFNALEVMDNYSVFQ-
DLKFGEGDGSLKYYLYNWKCASCHPSKIGIVLL

Template selection

A total of 45 templates was found. An initial HHblits profile was built using the procedure outlined in [18], followed by one iteration of HHblits against NR20. The profile obtained was then searched against all profiles in the SMTL. A total of 363 templates was found (Supplementary Table S1). NMT of *P. vivax* (PDB ID: 4A95) was selected as the template for modeling due to its high resemblance to the target and query coverage (Table 1). In contrast, the query coverage with CaNMT (PDB ID: 1IYL) was very low with low sequence identity (Supplementary Table S1) and ScNMT was not listed within a suitable range.

Model building

Nine models were built and hence model quality estimation was necessary to select the best model based on stability.

Model quality

The Ramachandran plot showed that 97.9 % of the residues are in favorable range, and 2.1 % are in allowed range (Fig. 2). The Ramachandran plot of the model generated in a similar work [10] showed only 83.3 % in favorable regions, 15.4 % in allowed regions and 1.3 % in disallowed regions. This proved that the present model is a much better fit to that of the experimental template structure.

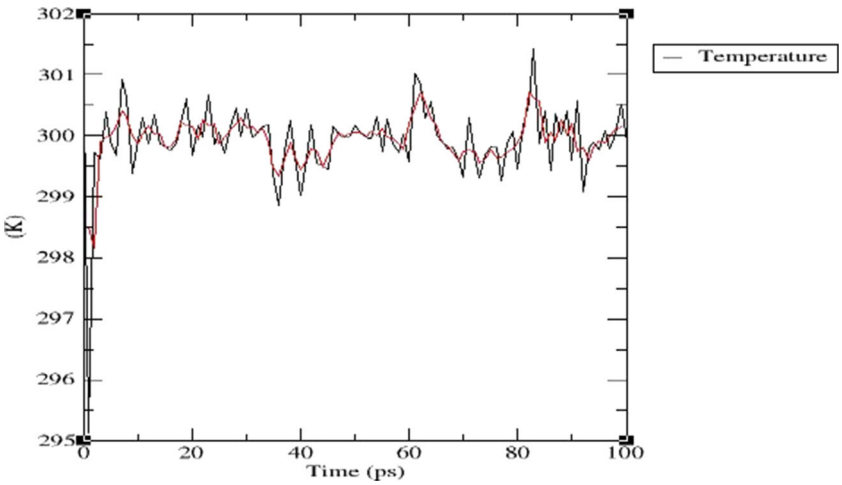

Fig. 6 Temperature curve showing the rise in system temperature, becoming stable with a running average of 300 K

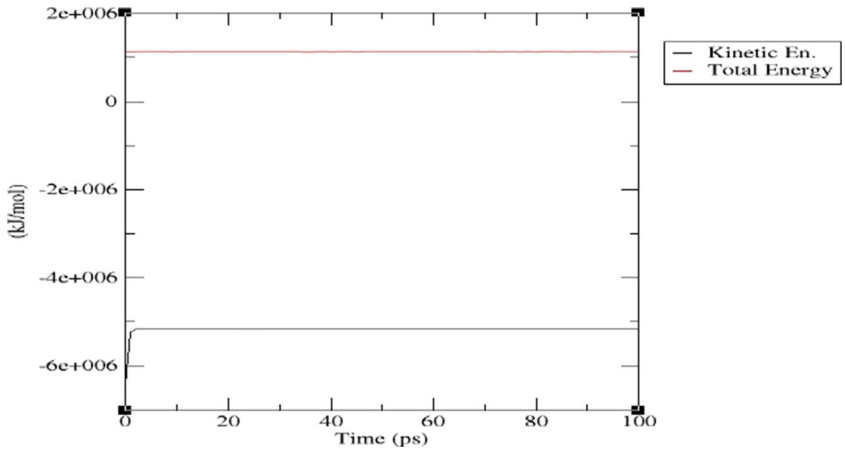

Fig. 8 Energy curve (300 K) using Gromacs 5.0 with constant total energy and kinetic energy with an initial rise in kinetic energy showing the minimized and equilibrium stage of the protein

The QMEAN (−0.57), CBeta (0.26), and Solvation (0.39) Z-score were found to be in a minimized state as all values are at average Z=0, but torsion had a score of −1.02, which showed that the model was still in a random state and required energy minimization for stabilization (Fig. 3).

MD simulation

MD simulation using Gromacs 5.0 confirmed the high stability of the selected model. Energy minimization was performed using the steepest descent method and was as shown in Fig. 4. The energy, pressure, density etc. were found to be optimum as per the Gromacs manual and are given graphically (Fig. 5).

Discussion

It is evident from Table S1 (see Supplementary File) and the Ramachandran plot (Fig. 2) that the best model created using template 4A95 employing the Modeller program (Modeller 9v9) has 97.9 % residues in the favorable range and the remaining 2.1 % in the allowed range, whereas the Ramachandran plot of earlier work [10] showed only 83.3 % in favorable, 15.4 % in allowed and 1.3 % in disallowed regions. The Ramachandran lot in the latter work was found to carry 1.3 % residues in disallowed regions and there were fewer favorable region residues (83.3 %) than in the present model (97.9 %). This again proved the validity of the current model and its superiority over the earlier study [10].

Comparison with earlier work [10] proved the novelty and quality of the current model as the template selected in this work is highly similar to that of PfNMT so far as phylogeny and sequence similarity is concerned (Supplementary Figure S1). The sequence similarity search (Table 1) result showed that the PfNMT sequence is highly similar to that of the selected template PvNMT (Score 80.47) compared to that of templates selected in similar earlier work [10]: CaNMT (Score 38.78 %) and ScNMT (40.73 %). The high similarity in query and template sequence justifies the novelty of the model and its resemblance to that of the experimental structure as *Plasmodium vivax* is very close to *P. falciparum* in a phylogram.

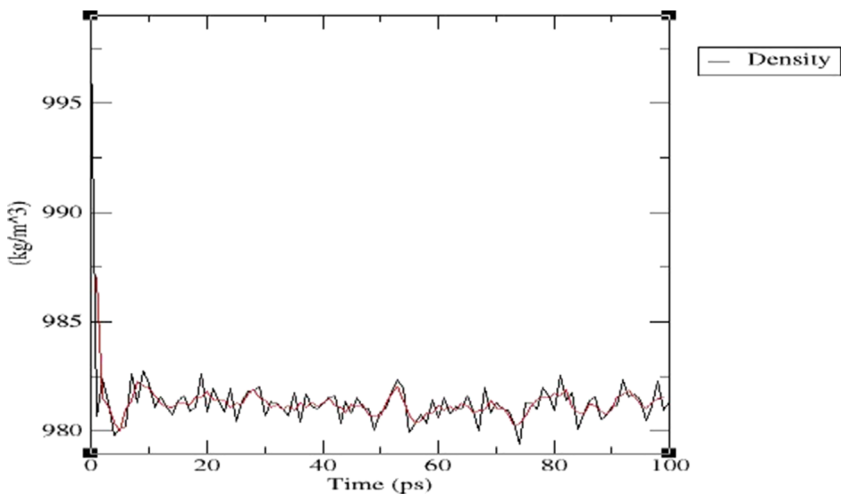

Fig. 9 Density curve (300 K) using Gromacs 5.0 showing density average at 981.307 kg/m^3, which is very close to 1,000 kg/m^3 showing high resemblance (P<0.0001), i.e., the system is well equilibrated in terms of density

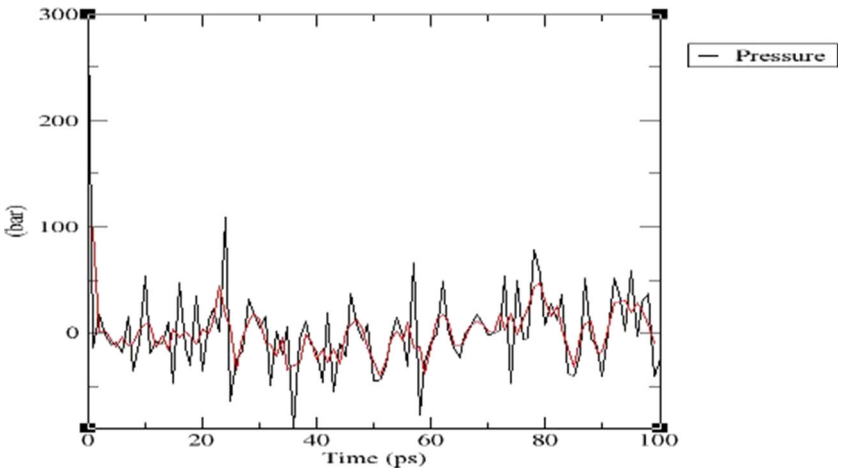

Fig. 10 Pressure curve (300 K) using Gromacs 5.0 showing pressure variation in the 100-ps run with an average at 0.83 bar with an expected value near 1.0 bar, which proves that the system is well equilibrated in terms of pressure ($P<0.0001$)

The WhatIf report [19] revealed three errors in side chain planarity for the selected model. The errors were observed in the case of Asp16A, Asp158A and Tyr66A. None of the aromatic rings were out of plane in any connection. After fixing these side chain errors, the Ramachandran plot did not show any significant difference in terms of residue torsion angles. Hence, we conclude that those amino acid residues do not contribute to the conformation of the modeled protein.

The QMEAN score (Fig. 3) proved that the modeled structure is organized randomly and hence minimization was necessary in order to stabilize the structure. The minimization of the protein, with potential energy below zero after 1,124 descent steps (Fig. 5), proved that the protein was stabilized easily in the SPC water model. The temperature of the system was then measured to be 299.7 K (Fig. 6), which is very close to normal temperature (300 K). This means that the system does not collapse after protein simulation and proves the flexibility of the protein model.

Comparison with template structure in terms of root mean square deviation (RMSD) again showed the close resemblance of the modeled structure with that of the template *Plasmodium vivax* NMT (Fig. 7).

The energy profile showed that kinetic energy increased at early stages in the 100-ps run, showing that the protein was in motion but gradually became stable (Fig. 8). The average density value was very close to the experimental value of the SPC water model and was found to be significant at $P<0.0001$ (Fig. 9). The pressure curve showed an average pressure of 0.83 bar, which was found to be significant at $P<0.0001$ (Fig. 10). This proved that the system was well equilibrated and stabilized in terms of density and pressure.

The work is novel in the sense that it was based on homology modeling with NMT from *P. vivax*, which is very similar to *P. falciparum* as seen in the BLOSUM matrix. The model quality was found to be very high and to resemble closely that of the experimental model of PvNMT as seen in RMSD analysis.

Conclusions

Plasmodium falciparum NMT is a vital target for the design of novel antimalarials. From the results of this study, it can be concluded that the modeled structure of pfNMT will open new avenues of drug design targeted against the dreaded parasite *P. falciparum*. The structure validation and molecular simulation showed that the structure may closely resemble the experimental structure. The modeled structure may be used in molecular docking studies for novel drug design.

Acknowledgments The authors acknowledge the support of Bioinformatics Center (DBT-BIF), Assam University, Silchar for use of computational tools and the DBT e-Library Consortium (DeLCON) facility for literature search. The corresponding author, A. Chowdhury is in receipt of fellowship from DBT sponsored Biotech Hub, Assam University, Silchar.

Conflict of interest The authors declare no conflict of interest. The authors follow Ministry of Science & Technology, Govt. of India Open Access Policy.

References

1. "http://www.who.int/mediacentre/factsheets/fs094/en/" retrieved 28 July 2014
2. Devadas B, Freeman SK, Zupec ME, Lu HF, Nagarajan SR, Kishore NS et al (1997) Design and synthesis of novel imidazole-substituted dipeptide amides as potent and selective inhibitors of *Candida albicans* myristoylCoA: Protein N-myristoyltransferase and identification of related tripeptide inhibitors with mechanism-based antifungal activity. J Med Chem 40:2609–2625
3. Bhatnagar RS, Schall OF, Jackson-Machelski E, Sikorski JA, Devadas B, Gokel GW et al (1997) Titration calorimetric analysis of AcylCoA recognition by myristoylCoA: protein N-myristoyltransferase. Biochemistry 36:6700–6708
4. Rudnick DA, McWherter CA, Rocque WJ, Lennon PJ, Getman DP, Gordon JI (1991) Kinetic and structural evidence for a sequential ordered Bi-Bi mechanism of catalysis by *Saccharomyces cerevisiae* myristoyl-CoA: protein N-myristoyltransferase. J Biol Chem 266: 9732–9739

5. Farazi TA, Waksman G, Gordon JI (2001) Structures of *Saccharomyces cerevisiae* N-myristoyltransferase with bound myristoylCoA and peptide provide insights about substrate recognition and catalysis. Biochemistry 40:6335–6343
6. Boutin JA (1997) Myristoylation. Cell Signal 9:15–35
7. Farazi TA, Waksman G, Gordon JI (2001) The biology and enzymology of protein N-myristoylation. J Biol Chem 276: 39501–39504
8. Gordon JI (1990) Protein N-myristoylation: simple questions, unexpected answers. Clin Res 38:517–528
9. Han KK, Martinage A (1992) Post-translational chemical modification(s) of proteins. Int J Biochem 24:19–28
10. Sheng C, Ji H, Miao Z, Che X, Yao J, Wang W, Dong G, Guo W, Lü J, Zhang W (2009) Homology modeling and molecular dynamics simulation of N-myristoyltransferase from protozoan parasites: active site characterization and insights into rational inhibitor design. J Comput Aided Mol Des 23(6):375–389
11. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402
12. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:2714–2723
13. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815
14. Benkert P, Biasini M, Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27:343–350
15. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291
16. Lovell SC, Davis IW, Arendall WB III, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC (2002) Structure validation by Calpha geometry: phi, psi and Cbeta deviation. Proteins Struct Funct Genet 50:437–450
17. Mariani V, Kiefer F, Schmidt T, Haas J, Schwede T (2011) Assessment of template based protein structure predictions in CASP9. Proteins 79(Suppl 10):37–58
18. Remmert M, Biegert A, Hauser A, Soding J (2012) HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods 9:173–175
19. Vriend G (1990) WHAT IF: a molecular modeling and drug design program. J Mol Graph 8:52–56