Available online at www.elixirpublishers.com (Elixir International Journal)

Biosciences

Elixir Biosci. 76 (2014) 28414-28421

ISSN: 2229-712X

Structural Insight into molecular model of hypothetical protein from *Trichomonas vaginalis*: A Computational Approach

Nitin Kumar Verma[1*], Vibha Verma[2], R K Deshwal[1] and Neelesh Yadav[3]

[1]Faculty of Bioscience, Shri Ram College, Muzaffarnagar, Uttar Pardesh, India.

[2]Genetics and Tree Propagation, Forest Research Institute, Dehradun, Uttarakhand, India.

[3]Forest Informatics, Forest Research Institute, Dehradun, Uttarakhand, India.

ARTICLE INFO

Article history:

Received: 16 September 2014;
Received in revised form:
27 October 2014;
Accepted: 8 November 2014;

Keywords

Trichomonas vaginalis,
Asparagine synthetase,
Comparative modeling,
Modeller,
Hypothetical protein,
Molecular Dynamics Simulation.

ABSTRACT

Hypothetical proteins predicted from nucleic acid and which have not shown any chemical evidence in the genome. In this study, we elaborated the functional and structural molecular model of the hypothetical protein of *Trichomoniasis vaginalis*. Functional annotation was carried out by using Pfam, SMART, CDD and BLAST. The conceptual three-dimensional structure has been investigated, since there was no structure available in any of the databases. We predicted the structure of hypothetical protein in *Trichomonas vaginalis* by using the comparative modeling approach. Molecular dynamics simulation was used to characterize its structural and dynamic feature at 10 ns by using the GROMACS. In the end the simulated model was validated with different web servers SAVES, WHAT IF, PorSA, iPAB. We observed this hypothetical protein was involved in asparagine biosynthesis performing catalytic activity. These findings are essential in reducing the gap between the deficiencies of annotation and crucial biosynthetic pathways and may be endorse in relational drug designing of molecules of structure function studies.

© 2014 Elixir All rights reserved.

Introduction

Trichomonas vaginalis (TV) is appreciated one of the most common non-viral genitourinary sexually transmitted diseases (STDs) in the world [1]. Infection by this parasite is associated with premature rupture of membranes and premature delivery in pregnant women [1, 2, 3]. Recently, the vaginitis has been associated with an increased risk of human immunodeficiency virus (HIV) acquisition and transmission [4, 5]. This infection occurs only in men, but its prevalence and associated morbidity appear to be significantly lower than those reported for women [6, 7, 8, 9, 10]. According to a WHO estimates, it accounts for almost half of all curable STDs [11]. In spite of high prevalence, it is one of the poorly studied parasites with respect to virulence properties, pathogenesis, and immunopathogenesis. *T. vaginalis* infection is asymptomatic in about 50% of infected women and over 90% of men; thus, re-infection and exposure is problematic. Furthermore, coinfections among these three STIs (sexually transmitted infections) are commonly known [12]. It causes prostatitis in men and vaginitis, cervicitis, and urethritis in women. Additionally, *T. vaginalis* has been considered a danger marker for other sexually transmitted agents, like as *Chlamydia trachomatis* and *Neisseria gonorrhoeae* [13, 14]. It has been involved as a cofactor in the transmission of the human immunodeficiency virus [5]. The *Trichomonas vaginalis* genome size is large, ~160 Mb and was found and approximately 60,000 protein-coding genes were identified throughout the six *T. vaginalis* chromosome.

In 1995, the first genome research was started with sequencing from *H. Influenza* strain Rd KW20 which genome size was 1.8 Mb. Eight years later, 100 organisms have been sequenced so far, and sequencing of others are under process. At the beginning of the genome era incompatibility in the accuracy of genome annotation that was the subject of hot discussion. Still, so called "70% hurdle" holds, as functions of only ~50 ± 70% of the genes in any given genome can be expected with reasonable confidence. Other remaining genes are either (i) homologous to genes of unspecified function, and are typically introduced as "conserved hypothetical" genes, or (ii) do not have any unspecified homologs termed "hypothetical" or "non characterized" or "unknown" because of confusion that they encode actual proteins [15]. Hypothetical proteins are predicted proteins using, nucleic acid sequences and their experimental existence are under scrutiny. Furthermore, these proteins are defined to low identity to known, annotated proteins [16].

Three-dimensional protein structures are inestimable sources for information for the functional annotation of protein molecule. Protein structures are determined by X-ray crystallographic and Nuclear Magnetic Resource (NMR) spectroscopy. In such cases, prediction of protein structure by computational methods can frequently result in a useful method. Protein structure can be modeled either *ab initio* from sequence alone or by the comparative methods that rely on a database of known protein structure [18, 19]. On the one hand, *Ab initio* methods are largely based on the laws of physics. On the other, comparative methods, including comparative modeling and threading, are primarily based on statistical learning. Besides the significant improvements in the *ab initio* [20] and threading methods [21], comparative modeling are known to produce accurate results if a known protein structure that is sufficiently similar to the modeled sequence is available [18]. Two conditions have to be met to predict protein structure of comparative modeling [22, 23]. Firstly, the target sequence must have similarly to the template sequence of known structure Secondly, it must be possible to base on accurate alignment between the target sequence and the template structure. The whole structure prediction process consists of fold assignment, target-template alignment, model building and model evaluation.

Tele:
E-mail addresses: nittinkumarverma@gmail.com

© 2014 Elixir All rights reserved

In this study, using sequence and structure-based approaches, we were capable to predict the function of protein annotated as hypothetical in the *T. vaginalis* genome. The structural features and the structural stability were assessed by molecular dynamics (MD) simulation.

Material and Methods

Sequence retrieval and analysis of the sequence

The full length amino acid sequence of hypothetical protein (NCBI ID: EAY06873.1) was retrived from NCBI database. The full length of hypothetical protein was comprised of 318 amino acids. The InterProScan tool was used to infer the protein families, super families and the domain arrangement within the protein. Conserved domains of the hypothetical protein were explored using following: Pfam databases (http://pfam.sanger.ac.uk) [24], Simpler Modular Architecture Research (SMART) tool (http://smart.embl-heidelberg.de) [25] and Conserved Domain Database (CDD) (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) [26]. Primary structure analysis on amino acid sequences was performed *via* Protparam tool on ExPasy proteomics server and secondary structure was predicted by Jpred web server.

Molecular Homology Modeling

Homology modeling is the method when there is a well-defined relationship of homology between the sequences of a target protein and at least one experimentally definite three-dimensional structure. This computational technique is based on the assumptions that the tertiary structure of two proteins will be similar if their sequences were related, and it is the approach most likely to give an accurate result [27].

Sequence comprising the domain was used to build up the 3-D structures using the comparative protein modeling method of Modeller9v10 [28]. To search the appropriate templates, DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST) [29] search tool was used against Protein Data Bank (PDB) (http://www.rcsb.org). If the retrieval accuracy and sensitivity towards protein analysis using normal BLAST is more than of DELTA-BLAST, in those cases, we used DELTA-BLAST. To assess the accuracy of template identification process, apart from DELTA-BLAST, hypothetical protein was subject to various meta-servers like GeneSilico [30], Geno3D [31] and Pcons.net [32] in order to find reliable templates. After observation, hypothetical protein (target) shows similarity with a template protein named Asparagine Synthetase Mutant C51a, C315a complexed with L- Asparagine from *E.coli* [PDB ID: 11AS]. The target-template alignment was performed using ClustralX and edited with BioEdit. The alignment of the target-template is shown in fig. 1.

Fig 1. Sequence Alignment for TVAG (hypothetical protein) and Asparagine Synthetase Mutant from *E. coli* (11AS_A). The sequence identity between TVAG and 11AS_A is ~41%. The alignment was performed using ClustralX and edited with BioEdit.

Total five models were constructed using the program MODELLER9v10 [28]. MODELLER executes comparative protein modeling by the gratification of spatial restraints.

Analysis of the Model

We used modeller for construction of models and their respective models used for the further study. We used the model with the smallest Modeller Objective Function and PROCHECK statistics. The whole stereochemistry quality of the models TVAG protein using Ramachandran plot calculation computed with PROCHECK [33] program, available at NIH (National Institute of Health) server (http://nihserver.mbi.ucla.edu/SAVES_3/Procheck/) and the final structure was later checked by VERIFY-3D graph accessible through NIH server (http://nihserver.mbi.ucla.edu/SAVES_3/Verify-3D/). The data of all five models are in Table 1.

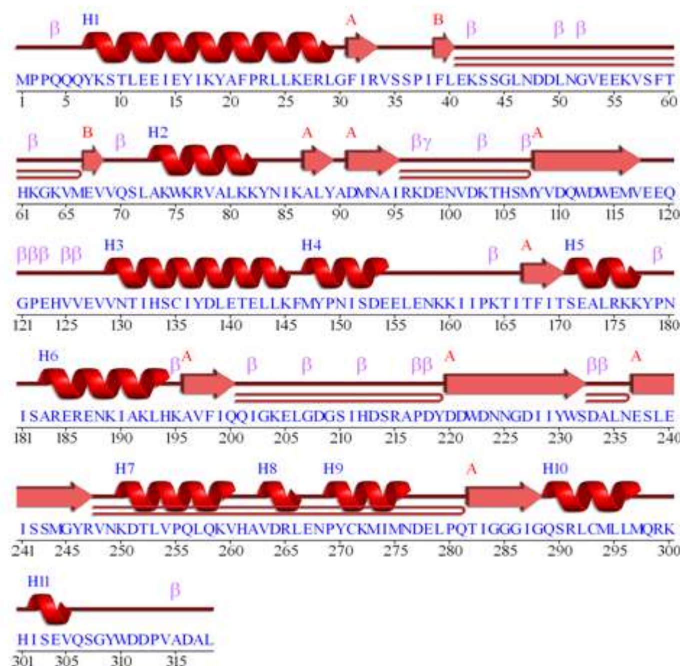

Fig 2. The graphical representation of hypothetical protein with its secondary structural elements

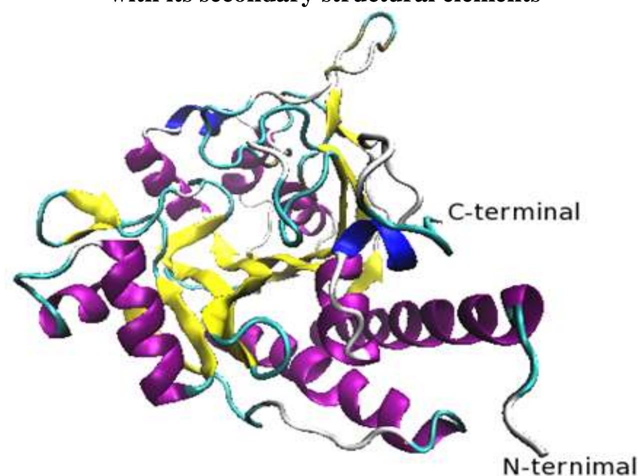

Fig 3. Homology model of hypothetical protein (gi| 121901871) from *T.vaginalis* solid ribbion representation of hypothetical protein model colored by it secondary structure elements

Energy Minimization

The best model with the lowest DOPE score was subject to energy minimization by GROMACS 4.5.6 using minimization protocol. The minimization protocol employs the steepest

descent method for the removal of bad van-der-Waals contacts from the model. In this study the calculations were done by utilizing GROMACS 4.5.6 [34] software with G43a1 force field with a flexible SPS water model in a cubic box of 1.2 Angstrom dimension.

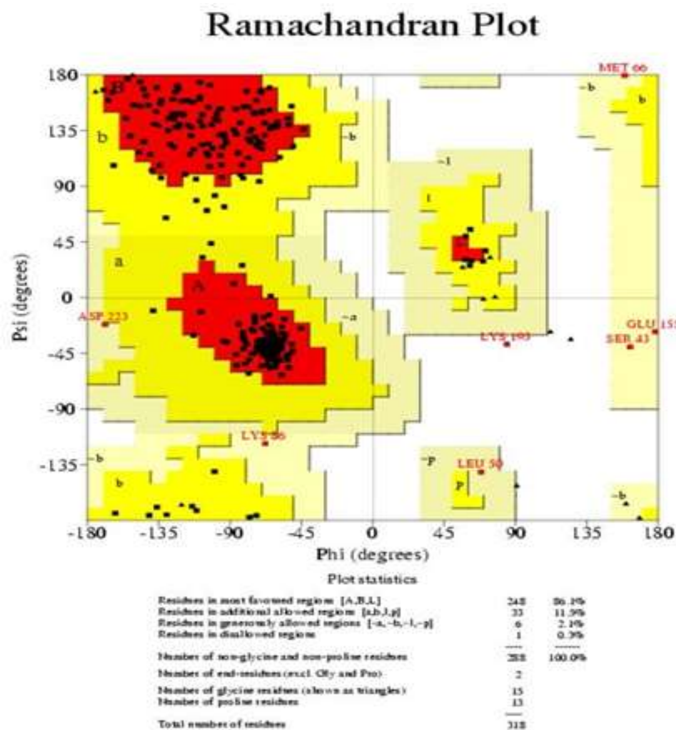

Fig 4. Ramachandran plot of the modelled hypothetical protein (gi| 121901871). The plot was calculated by PROCHECK program.

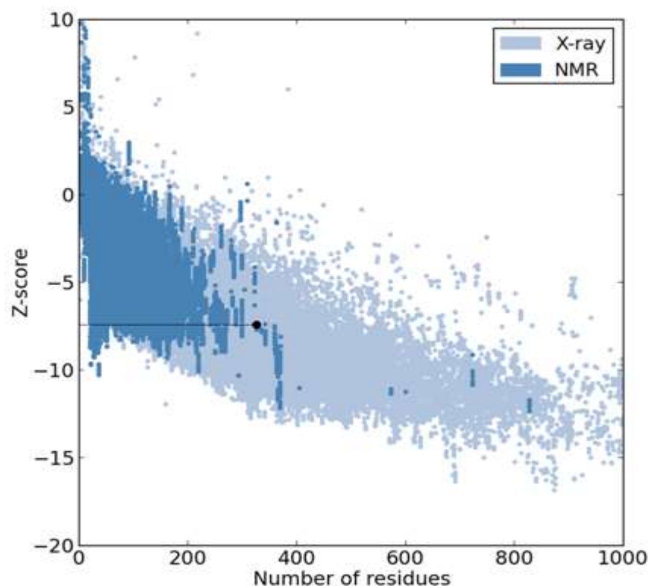

Fig 5 (A). This plot that contains the Z-score of all experimentally determined protein chains in the current PDB that have been solved by either X-ray diffraction or NMR. The Z-score of -7.43 indicated on this graph represents the overall quality of the modeled 3D structure of 11AS

Quality assessment and validation

The quality, internal consistency and reliability of the energy minimized TVAG (hypothetical protein) were evaluated by a number of computational tools. PROCHECK [33] was used to check the stereo-chemical of the model which quantifies the residues in the available zones of the Ramachandran plot provide the position of the torsion angles phi *(φ)* and psi (ψ) between C_α-C and N-C_α atoms of the residues contained in a peptide. VERIFY_3D program determines the compatibility of the atomic model (3D) with its own amino acids sequences where a high VERIFY_3D profile [34]. Protein Structure Analysis (ProSA) was employed for the refinement and validation of the modeled structure which checks the native protein folding energy of the modeled by comparing the energy of the model with the potential mean force derived from large sets of known protein structures [35].

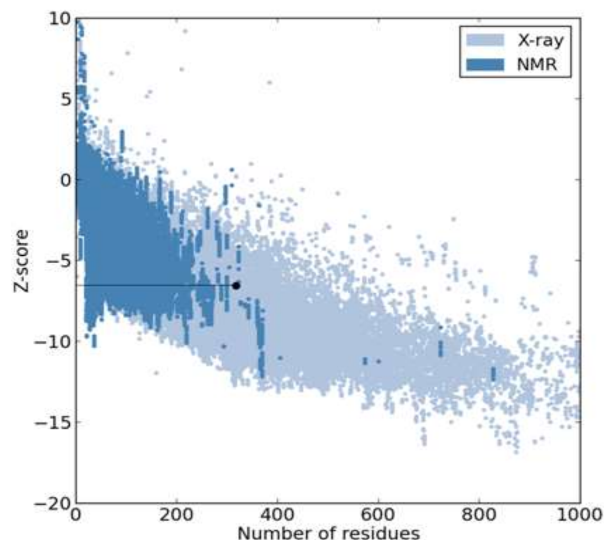

Fig 5 (B). The Z-score of -6.56 indicated on this graph represent the overall quality of the templates 3D structure (PDBID: 11AS). When compared to figure (A), it was found that the over all quality of the 3D structure to the templates and that of the target are very similar in the term of their Z-score

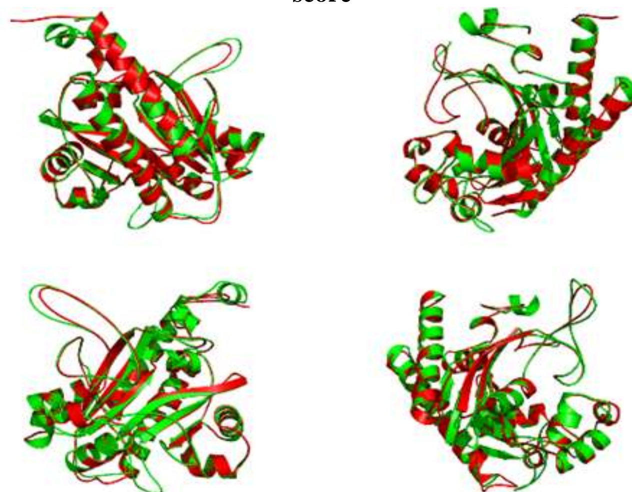

Fig 6. Superimposition of the hypothetical protein (gi| 121901871) and the template (PDBID: 11AS) by iPAB web server

Protein interaction network mapping

Protein-protein interactions were constructed from the STRING database [37] comprising known and predicted physical and functional protein-protein interactions. STRING in protein mode was used, and the only interactions with high confidence levels (>0.7) were kept. STRING quantitatively inter-grates interaction data from these origins for a huge number of organisms, and transfer information between these organisms was applicable.

Molecular dynamics simulation of TVAG

MD simulation was performed with the GROMACS (GROningen MAchine for Chemical Simulations) [38] package using the Gromos 96.1 (43A1) force field. The accurate force field is necessary for reproducing the conformational and dynamic behavior of condensed-phase system. The Gromos 96.1 force field is well parameterized for proteins. To solvate the model, it was placed in a cubic box maintaining a distance of 1.2 nm between the box edges and the protein surface. Particles mesh Ewald (PME) electrostatic and periodic boundary conditions were applied in all directions. The system was neutralized by adding 10 Na+ counter ions since the overall protein charge was negative. To get rid of the high energy interactions and steric clashes of the system, the energy of the system was minimized using steepest descent method until a tolerance of 1000 KJ/mol. All the bond lengths were constrained with the LINear Constrains Solver (LINCS) [38] method, whereas the geometry of water molecules was constrained with SETTLE algorithm [39]. The energy minimization system was treated for 400ps equilibration run. The pre-equilibrated system was consequently subjected to 10ns production MD simulation, with a time-step of 200ps at constant temperature (300K), pressure (1atm) and without any position restraints. Snapshots of the trajectory were taken on every 1ps and all the analysis of the MD simulation were performed using XMGRACE software. From the 10ns MD simulation, average structure with low RMSD value was chosen as the best model for further analysis.

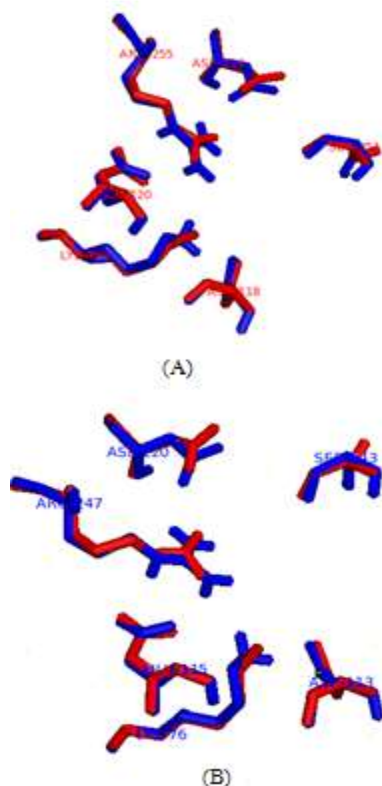

Fig 7. Stereo view of superposition of the catalytic regions between 11AS (in red color) and TVAG (in blue color)

Results and Discussion:

Sequence Analysis

Hypothetical protein (36, 7909 Kda) of *Trichomonas vaginalis* has been extracted from Genbank . The functional analysis of protein includes protein domains and family, prediction of hypothetical protein by conserved domain (CD) search and Pfam, revealed one putative domain AsnA (12-236) Table-1. A conserved domain search of hypothetical protein revealed which belong to AsnA, consist of catalysis the conservation of L-aspartate to L-asparagine in the presence of ATP and ammonia. Ammonium is a form of inorganic nitrogen derived from several metabolic pathways, and is assimilated into glutamine, glutamate, asparagine and carbamoylphosphate. These molecules play important role in nitrogen assimilation, recycling, transport and storage in organism. Ammonium assimilation into asparagine is catalyzed by ammonia-dependent asparagine synthetase by asnA (EC 6.3.1.1) in prokaryotes and eukaryotes.

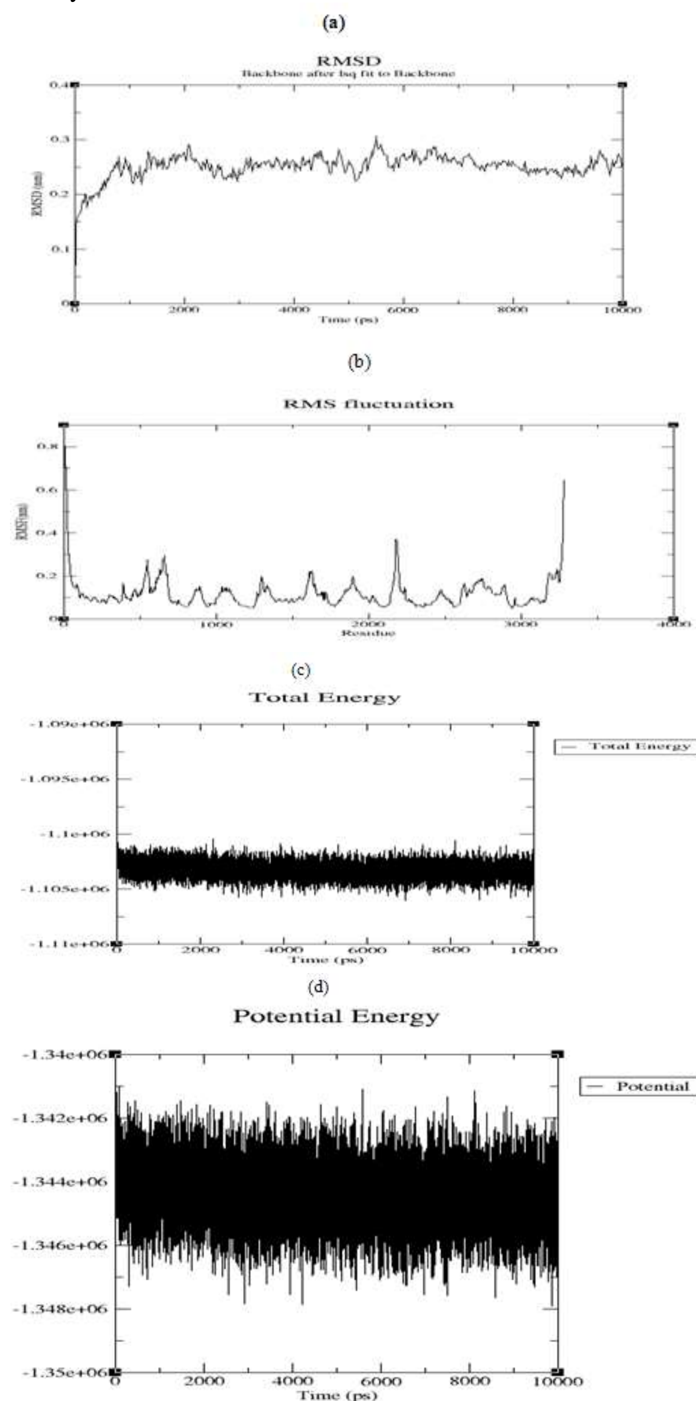

Fig 8. The RMSD, RMSF, Total energy and Potential energy graphs of the modelled hypothetical protein during MD Simulation. (a) RMSD of backbone C_α atom of the modelled hypothetical protein structure. (b) RMSF analysis of amino acid residues of the modelled hypothetical protein structure. (c) Total energy (kJ/mol) during 10 ns trajectory. (d) Potential energy (kJ/mol) during 10 ns trajectory

Table 1. Identification of Domain and Families

Sequence ID	CDD-BLAST	PFAM	PFAM Description
gi|121901871	asparagine synthetase AsnA	AsnA	Aspartate-ammonia ligase

The primary sequence analysis of hypothetical protein was calculated in Table-2 since the isoelectric point (pI), solubility is minimized and mobility in an electro focusing system is zero that's why calculated pI will be useful. Isoelectric point (pI) is the pH at which the surface of the protein is covered by the charge, but a net charge of the protein is zero. At pI, protein is stable and dense. For processing buffer system for purification by isoelectric focusing method, the computed isoelectric point (pI) will be valuable. While Expasy's Protparam computes the extinction coefficient of 276, 278, 279, 280 and 282 nm wavelengths, 280 nm has been elected since proteins absorb light strongly. Extinction coefficient of hypothetical protein at 280 nm was 53985 M^{-1} cm^{-1}. The computed extinction coefficient can help in the quantitative study of protein-protein and protein-ligand interaction in solution. The instability index provides and determine of the stability of protein in a test-tube. There are definite dipeptides, the occurrence of which is particularly divergent in the unstable protein compared with those in the stable once. This method assigned a weight value of instability, which is feasible to compute an instability index (II). A protein whose instability index is slighter than 40 is estimated as stable, a value above 40 estimates that the protein may be unstable.

Table 2. Physiochemical properties of hypothetical protein by Protparam tool

Sequence ID	Seq. Len.	MW	pI	(+)R	(-)R	EC	II	AI	Gravy
gi|121901871	318	36790.9	5.41	41	51	53985	48.98	87.64	-0.543

The instability index value of the hypothetical protein was found 48.98, which indicate that, the protein is moderately unstable. The aliphatic index (AI) is elucidated as the relative volume of a protein occupied by aliphatic side chains (A, V, I and L) is estimated as a positive factor for increase of the thermal stability of globular proteins. Aliphatic index (AI) for the hypothetical protein sequence was 87.64. The very high aliphatic index of the protein sequence indicates that these proteins may be stable for a vast temperature range. The minimal thermal stability of protein was indicative of a more flexible structure when compared to other protein. The Grand Average hydropathy (GRAVY) value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divide by the number of residues in the sequence. A GRAVY index of hypothetical protein was -0.543. This low value shows the probability of better interaction with water. Disulphide bonds play an important role in the stability and folding of proteins. Absence of disulphide bonds in that hypothetical protein of *T. vaginalis*. Disulphide bridges were found using a Cys_REC tool from softberry. The structural model of the hypothetical protein of *T. vaginalis* contains an eleven alpha-helices and four beta ribbons motif (Fig. 2).

Protein-protein interaction database have become a major resource for investigation biological networks and pathways in cells. To compute protein interaction properties of the hypothetical protein, we utilized search tool for the retrieval of interacting genes and proteins (STRING) database of physical and functional interactions. The protein with id gi|121901871 interaction using protein structural similarities permits to construct various candidate interactions with possibly significant functional relevance. For this purpose, the relation between the nine identified proteins was examined. The interaction network of genetically interacting proteins possibly related in function with *Trichomonas vaginalis* is shown in Fig. 9 and the detailed information is presented in Table-3. The blue line indicates statistically significant co-occurrence across multiple genomes. The graph of the A2EKG5 (hypothetical protein of *Trichomonas vaginalis*) network shows the identified A2EKG5-interacting proteins and phylogenomic profile of A2EKG5-related function.

Table 3. List of predicted interactive proteins with A2EKG5 (hypothetical protein) of *Trichomonas vaginalis*

Sr. No.	ID	Protein Name	AA residue	Score
1	CBK	Carbamate kinase	314	0.899
2	A2GK90	Glutaminase family protein	244	0.899
3	A2G7J5	Aspartate aminotransferase	399	0.899
4	A2FCY9	Glutamate dehydrogenase	507	0.899
5	A2F7X5	Aminotransferase, classes I and II family protein	398	0.899
6	A2EIU6	Aminotransferase, classes I and II family protein	414	0.899
7	A2EI95	Aminotransferase, classes I and II family protein	400	0.899
8	A2DLF1	Aspartate aminotransferase	398	0.899
9	A2FBPO	fatty aldehyde dehydrogenase variant form	465	0.814

Comparative modelling and energy minimization

Comparative modeling of protein provides a significant hypothesis of homology between the target and template. This approach provides reasonable results based on the assumption that the tertiary structure of the two proteins will be similar if their sequences are related. Absence of the experimentally determined three dimensional structure of hypothetical protein of *Trichomonas vaginalis* in PDB (Protein Data Bank), comparative modeling method was utilized to construct its theoretical three dimensional structures. DELTA-BLAST scanning results had revealed more identical with crystallographic structure *E.coli* K-12 (PDB ID: 11AS at 2.50 Å resolution) while the template was determined on the basis of higher sequence identity. It has been 39.80 % sequence identity, with 135 conserved residues and 54.30% sequence similarity. Comparative modeling predicts the 3-D structure of hypothetical model of a given protein sequence (target), based primarily on this alignment to the template. The resulting 3-D structure of hypothetical protein was sorted according to the scores calculated from discrete optimization, protein energy (DOPE) scoring function. The final model which has lowest root mean square deviation (RMSD), relative to the trace of the crystal structure was selected for further study (Fig. 3).

Model quality assessment

The detailed residue-by-residue stereo-chemical quality of the modelled protein structure was evaluated by the Ramachandran plot (Fig. 4) using Procheck tool. The reliability of the backbone torsion angle Φ and Ψ distribution of the protein and the template was evaluated by the Ramachandran plot in Procheck tool. The perceived Ramachandran plot (Phi-Phi) pairs had 86.1% residues in most favored regions, 11.5% core residues in additional allowed regions, 2.1% residues in generously allowed regions and 0.3% residues in disallowed regions Table-4. This value indicates a good quality model. Whereas the crystal structure of *E.coli* PDB ID: 11AS shows 84.60% residues in most favored regions. In order to

characterize the model, structural motif and mechanically important loops were assigned to build a final 3D model of hypothetical protein. The packing quality of each residue of the model was assessed by Verify_3D program where the compatibility of the model residues with their environment is assessed by a score function. Residues with a score over 0.2 should be considered reliable. As shown in Table-5; the score of the refined model maximally was above 0.2 which corresponds to the acceptable side chain environment. PorSA revealed a Z-score of -6.56 (Fig. 5-B) for modelled hypothetical protein, where the template has a Z-score of -7.43 (Fig. 4-A) reflecting the overall quality of the model.

Table 4. Comparison of Ramachandran plot statistics of TVAG model with closest homologue structure 11AS

Ramachandran Plot Statistics	Modeled structure of TVAG		Template (11AS)	
	Residue	Percentage (%)	Residue	Percentage (%)
Residue in most favored region	248	86.1	477	84.6
Residue in additional allowed region	33	11.5	83	14.7
Residues in generously allowed region	6	2.1	3	0.5
Residue in disallowed region	1	0.3	1	0.2
Number of non-glycine and non-proline residues	288	100	564	100
Number of end residues	2	-	6	-
Number of glycine residues	15	-	58	-
Number of proline residues	13	-	28	-

RMSD (root mean square distance) between the equivalent Cα atoms pair (target and template) was measured to check the degree of structural similarity. We examined the best modelled structure for fitting into the template (crystal structure), the prepare model and its closest relative was superimposed based on Cα and backbone atom pairs. A pairwise 3D alignment search of the template protein with the modelled structure through iPAB web server showed the identity of enormous 90.83% for 307 aligned residues (Fig. 6) with massive RMSD of 0.48 Å on their backbone atom. The iPAB web server's results conclude that hypothetical protein and its structural homologues share strong structural conservation and similarity in the structural folding. It also signifies that the generated model is reasonably good for further studies.

Table 5. Comparison of model validation scores from different server between modeled TVAG and its closest structural homologue 11AS (A chain)

Target/ template	Verify_3D	Errat	Prove (Z score)	ProSA
TVAG	80.88%	72.81	7.75	-6.56
11AS	92.68%	87.77	1.87	-7.43

All residues (Lys_77, Asp_118, Glu_120, Asp_219, Ser_251, and Arg_255) are already defined of the cavity in 11AS (*E.coli*) [38]. We observed in structure alignment between TVAG and 11AS, the residues position in TVAG cavity are Lys_76, Asp_113, Glu_115, Asp_220, Ser_243 and Arg_247 (Fig. 7).

Molecular dynamics simulation

The refinement of the model protein was performed through MD simulation to get an optimized and stable structure for docking with cofactor. The stability and dynamic properties of hypothetical properties of *T. vaginalis* model was observed by MD simulation run at 10 ns duration. The steepest descent energy minimization for the solvated modelled protein revealed the maximum force reached the threshold of 1000 $kJmol^{-1}nm^{-1}$ in 401 steps. The RMSD of the protein backbone atoms are plotted as a function of time to check the stability of the system throughout the simulation. Compared to the starting coordinates, the RMSD of the backbone atoms increased in the first 2 ns and then reached a plateau in the subsequent simulation time (Fig. 8 (A)). The relative flexibility of the model was also characterized by plotting the root mean-square flexibility (RMSF) related to the average structure obtained from the MD simulation trajectories. Three flexible regions have been predicted for the modeled protein structure of *T. vaginalis* considering the RMSF value and represented in (Fig. 8 (B)). One flexible region has been located in the N-terminal end, the RMSF value of this region is 0.29.u The second flexible region is the middle region of the protein, this region processes several continuous peaks. The RMSF values of the peaks in this region are 0.21, 0.19, and 0.39. The C-terminal end has a single flexible region with a dominant RMSF value of 0.69. It can be summarized from the RMSF analysis that middle portion of the modelled *T. vaginalis* (hypothetical protein) is more flexible in comparison of the N-termini and the C-termini. The MD trajectories of 10ns showed the compactness of structure through the radius of gyration of an average value of ~1.92 nm from nm to 1.9 the end of the simulation. During simulation, variation in potential energy (Fig. 8 (C)) and total energy was calculated to evaluate the stability of the model (Fig. 8 (D)).

Conclusions

In this study, we have evaluated both sequence and structural function of hypothetical protein through homology modeling and comparative genomics approach. An aggregation of Bioinformatics tools, focused not only on sequence analysis but also structural information, guided us to suggest the function of hypothetical protein in *Trichomonas vaginalis*. First of all, we suggested those functional properties of hypothetical protein (gi| 121901871) that responsible for catalytic activity in asparagine biosynthesis. There are two distinct types of asparagine synthetase (AS), AS-A and AS-B, encoded by *asnA* and *asnB* genes, respectively. This hypothetical protein (gi| 121901871) showed higher similarity with *asnA* gene. Both (*asnA* and *asnB*) type catalyzes the ATP-dependent conversion of aspartate into asparagine. Secondly, we have also constructed a 3D model of hypothetical protein using the comparative modeling approach. We determined their physiochemical characteristics of protein by different parameter like as isoelectric point, molecular weight, total number of positive and negative residue, extinction coefficient, instability index, aliphatic index and grand average hydropathy (GRAVY). This hypothetical protein was discovered at very low percentage in the amino acid cysteine and their lack presence of Disulphide Bridges are also inferred from analysis of cys_res result. In the absence of disulphide bonds, extensive hydrogen bonding is believed to be liable for the stability of proteins. The hypothetical protein model had a stable confirmation in response to the atomic flexibility. Protein-protein interactions (PPI) are necessary for almost all cellular functions. Proteins frequently interact with one another in a manually dependent way to perform an ordinary function. Computational analysis and experimental explanation of the composite networks established by respective protein-protein interactions (PPIs) are one of the vital challenges in the post-genomic era.MD simulation was carried out to distinguish its structural and impulsive features, which was an advance

validated by the SAVES, WHAT IF. The MD simulation of a hypothetical protein at temperature 300K, then β-sheets and α-helices of hypothetical protein are shown to play a role in protein stability. Data shows α-helices are more stable than β-sheets during simulation. The presented model will be potentiate and facilitate structural and functional investigation to use structure-based drug designing.

Competing Interests: The authors declare that they have no competing interests.

Acknowledgements

We are thankful to FRI University, Dehradun and Shri Ram College, Muzaffarnagar for providing the necessary facilities to this pursue the above research.

References:

[1] S.L. Cudmore, K.L. Delgaty, S.F. Hayward-mcclelland, D.P. Petrin, G.E. Garber, Treatment of Infections Caused by Metronidazole-Resistant Trichomonas vaginalis, 17 (2004) 783–793.

[2] M.F. Cotch, J.G. Pastorek, R.P. Nugent, S.L. Hillier, R.S. Gibbs, D.H. Martin, et al., Trichomonas vaginalis associated with low birth weight and preterm delivery. The Vaginal Infections and Prematurity Study Group., Sex. Transm. Dis. 24 (1997) 353–360.

[3] P.H. Hardy, J.B. Hardy, E.E. Nell, D.A. Graham, M.R. Spence, R.C. Rosenbaum, Prevalence of six sexually transmitted disease agents among pregnant inner-city adolescents and pregnancy outcome., Lancet. 2 (1984) 333–337.

[4] H. Minkoff, A.N. Grunebaum, R.H. Schwarz, J. Feldman, M. Cummings, W. Crombleholme, et al., Risk factors for prematurity and premature rupture of membranes: a prospective study of the vaginal flora in pregnancy., Am. J. Obstet. Gynecol. 150 (1984) 965–972.

[5] M. Laga, A. Manoka, M. Kivuvu, B. Malele, M. Tuliza, N. Nzila, et al., Non-ulcerative sexually transmitted diseases as risk factors for HIV-1 transmission in women: results from a cohort study, AIDS. 7 (1993) 95–102.

[6] C.C. Wang, R.S. Mcclelland, M.T. Reiner, J.M. Overbaugh, S.R. Emery, K.N. Mandaliya, et al., The effect of treatment of vaginal infections on shedding of human immunodeficiency virus type 1., J. Infect. Dis. 183 (2001) 1017–22.

[7] M.M. Hobbs, P. Kazembe, A.W. Reed, W.C. Miller, E. Nkata, D. Zimba, et al., Trichomonas vaginalis as a cause of urethritis in Malawian men., Sex. Transm. Dis. 26 (1999) 381–387.

[8] J.L. Joyner, J.M. Douglas, S. Ragsdale, M. Foster, F.N. Judson, Comparative prevalence of infection with Trichomonas vaginalis among men attending a sexually transmitted diseases clinic., Sex. Transm. Dis. 27 (2000) 236–240.

[9] J.N. Krieger, Trichomoniasis in men: old issues and new data, Sex Transm Dis. 2 (1995) 83–96.

[10] J.N. Krieger, M.R. Tam, C.E. Stevens, I.O. Nielsen, J. Hale, N.B. Kiviat, et al., Diagnosis of trichomoniasis. Comparison of conventional wet-mount examination with cytologic studies, cultures, and monoclonal antibody staining of direct specimens., JAMA. 259 (1988) 1223–1227.

[11] J.N. Krieger, C. Jenny, M. Verdon, N. Siegel, R. Springwater, C.W. Critchlow, et al., Clinical manifestations of trichomoniasis in men., Ann. Intern. Med. 118 (1993) 844–849.

[12] F. Who, Working to overcome the global impact of neglected tropical diseases First WHO report on neglected tropical diseases, World Health. 86 (2010) 186.

[13] J. Sobngwi-Tambekou, D. Taljaard, M. Nieuwoudt, P. Lissouba, A. Puren, B. Auvert, Male circumcision and Neisseria gonorrhoeae, Chlamydia trachomatis and Trichomonas vaginalis: observations after a randomised controlled trial for HIV prevention., Sex. Transm. Infect. 85 (2009) 116–120.

[14] D. Petrin, K. Delgaty, R. Bhatt, G. Garber, Clinical and microbiological aspects of Trichomonas vaginalis., Clin. Microbiol. Rev. 11 (1998) 300–317.

[15] D.E. Riley, M.C. Roberts, T. Takayama, J.N. Krieger, Development of a polymerase chain reaction-based diagnosis of Trichomonas vaginalis., J. Clin. Microbiol. 30 (1992) 465–472.

[16] S. Sivashankari, P. Shanmughavel, Functional annotation of hypothetical proteins - A review., Bioinformation. 1 (2006) 335–338.

[17] G. Lubec, L. Afjehi-Sadat, J.-W. Yang, J.P.P. John, Searching for hypothetical proteins: theory and practice based upon original data and literature., Prog. Neurobiol. 77 (2005) 90–127.

[18] D. Baker, A. Sali, Protein structure prediction and structural genomics., Science. 294 (2001) 93–96.

[19] A. Fiser, M. Feig, C.L. Brooks, A. Sali, Evolution and physics in comparative protein structure modeling., Acc. Chem. Res. 35 (2002) 413–421.

[20] P. Bradley, D. Chivian, J. Meiler, K.M.S. Misura, C.A. Rohl, W.R. Schief, et al., Rosetta predictions in CASP5: successes, failures, and prospects for complete automation. Proteins. 53 Suppl 6 (2003) 457–468.

[21] L.N. Kinch, J.O. Wrabl, S.S. Krishna, I. Majumdar, R.I. Sadreyev, Y. Qi, et al., CASP5 assessment of fold recognition target predictions., Proteins. 53 Suppl 6 (2003) 395–409.

[22] M.A. Martí-Renom, A.C. Stuart, A. Fiser, R. Sánchez, F. Melo, A. Sali, Comparative protein structure modeling of genes and genomes., Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 291–325. doi:10.1146/annurev.biophys.29.1.291.

[23] T. L. Blundell, B. L. Sibanda, M. J. Sternberg, and J. M. Thornton, Knowledge based prediction of protein structures and the design of novel molecules. Nature 326 (1987) 347–352.

[24] R.D. Finn, J. Mistry, J. Tate, P. Coggill, A. Heger, J.E. Pollington, et al., The Pfam protein families database., Nucleic Acids Res. 38 (2010) D211–D222.

[25] I. Letunic, T. Doerks, P. Bork, SMART 7: recent updates to the protein domain annotation resource., Nucleic Acids Res. 40 (2012) D302–5.

[26] A. Marchler-Bauer, S. Lu, J.B. Anderson, F. Chitsaz, M.K. Derbyshire, C. DeWeese-Scott, et al., CDD: a Conserved Domain Database for the functional annotation of proteins., Nucleic Acids Res. 39 (2011) D225–D229.

[27] R.T. Kroemer, S.W. Doughty, A.J. Robinson, W.G. Richards, Prediction of the three-dimensional structure of human interleukin-7 by homology modeling., Protein Eng. 9 (1996) 493–498.

[28] B. John, A. Sali, Comparative protein structure modeling by iterative alignment, model building and model assessment., Nucleic Acids Res. 31 (2003) 3982–3992.

[29] G.M. Boratyn, A.A. Schäffer, R. Agarwala, S.F. Altschul, D.J. Lipman, T.L. Madden, Domain enhanced lookup time accelerated BLAST, Biol. Direct. 7 (2012) 12.

[30] M.A. Kurowski, J.M. Bujnicki, GeneSilico protein structure prediction meta-server., Nucleic Acids Res. 31 (2003) 3305–3307.

[31] C. Combet, M. Jambon, G. Deléage, C. Geourjon, Geno3D: automatic comparative molecular modelling of protein., Bioinformatics. 18 (2002) 213–214.

[32] J. Lundström, L. Rychlewski, J. Bujnicki, A. Elofsson, Pcons: a neural-network-based consensus predictor that improves fold recognition., Protein Sci. 10 (2001) 2354–2362.

[33] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[34] D. Eisenberg, R. Luthy, J.U. Bowie, VERIFY3D: assessment of protein models with three-dimensional profiles, Methods Enzymol. 277 (1997) 396-404.
[35] M. Wiederstein, M.J. Sippl, PorSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (2007) W407-W410.
[34] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible, and free., J. Comput. Chem. 26 (2005) 1701–1718.
[35] L.J. Jensen, M. Kuhn, M. Stark, S. Chaffron, C. Creevey, J. Muller, et al., STRING 8--a global view on proteins and their functional interactions in 630 organisms., Nucleic Acids Res. 37 (2009) D412–D416.
[36] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[37] S. Miyamoto, P.A. Kollman, Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models, J. Comput. Chem. 13 (1992) 952–962.
[38] Oda, J., Nakatsu, T., & Kata, H., Crystal structure of asparagine synthetase reveales a close evolutionary relationships to class II aminoacyl-tRNA synthetase. Nature Publishing Group, 5(1) (1997).