

## Structural and Functional Insights into Fungal Uricases: A Computational Approach

S. Thatikonda<sup>a</sup>, M.N.P. Satish Kumar Machavarapu<sup>a</sup>, S. Tallapudi<sup>b</sup>, J. Babu D<sup>a</sup>, T.C. Venkateswarulu<sup>a</sup> and A.K. Nelapati<sup>a,\*</sup>

<sup>a</sup>Department of Biotechnology, Vignan's Foundation of Science, Technology and Research, Guntur, Andhra Pradesh, India

<sup>b</sup>Department of Chemical Engineering, Andhra University College of Engineering(A), Visakhapatnam, Andhra Pradesh, India

(Received 25 November 2023, Accepted 9 March 2024)

Gout is a prevalent rheumatic disorder arising from elevated uric acid levels in the bloodstream, a condition known as hyperuricemia. Uricase serves the purpose of reducing uric acid levels nevertheless, it is absent in humans as a result of evolutionary modifications. The present investigation focused on the computational analysis of uricase enzymes derived from eighty-eight fungal species and the analysis includes their structural, functional, physicochemical properties, motif search, and domain architecture analysis. In addition to this, we have performed the phylogenetic analysis of these protein sequences to study their phylogenetic relationship. Physicochemical analysis revealed that uricase spans amino acid residues 296-330, with a molecular weight of 33.03-35.05 kDa and theoretical pI values of 5.71 to 9.03. Valine was the most prevalent amino acid, averaging 8.40% across all species. These uricase were thermally stable and active in acidic and alkaline conditions. The secondary structure analysis indicated a predominance of alpha-helices and beta-sheets in uricase. *Aspergillus niger* (ABI79319.1) was chosen as the representative species, with its three-dimensional structure accurately predicted and validated. The comprehensive computational analysis conducted on the uricase protein offers the potential to identify a suitable microorganism capable of producing uricase with desirable characteristics for industrial applications.

**Keywords:** Uricase, Gout, Fungal species, Phylogenetic analysis, Secondary structure, Homology modeling

### INTRODUCTION

Uric acid (UA) is an important biological molecule in body fluids. It is the result of the purine metabolism process. Gout, hyperuricemia, and Lesch-Nyhan syndrome are all disorders characterized by abnormally high levels of uric acid in the blood. Because of its significance, UA screening is performed routinely in the clinical laboratory [1-4]. The enzyme uricase (EC 1.7.3.3) facilitates the conversion of uric acid to allantoin. It is part of the purine degradation pathway. All five types of hominoids (gibbons, orangutans, gorillas, chimpanzees, and humans) don't have active uricase, but most Old World and New World monkeys do. When

uricolytic activity goes away, uric acid levels in the blood rise. Uric acid is thought to be a significant antioxidant that may protect against environmental damage and make people live longer. The inactivation of uricase could confer evolutionary benefits in primates and potentially play a role in the emergence of humans. The processes leading to uricase inactivation vary among different hominoid lineages [5]. While humans and other higher primates have an ancestral uricase encoded in their genomes, the development of a functional enzyme is prevented by two nonsense mutations. It is a matter of heated debate as to why such genetic variations, which may give unique benefits, were permitted to proliferate throughout evolution. However, some people do acquire pathological disorders associated with elevated UA levels, and it has been shown clinically that exogenous uricase can reduce these symptoms [6-8].

\*Corresponding author. E-mail: nelapatianandkumar@gmail.com

The initial investigation into hyperuricemia treatment in humans utilizing uricase purified from beef kidneys dates back to 1957. Subsequently, in 1967-1968, researchers extracted uricase from *Aspergillus flavus* and applied it to manage hyperuricemia in both animals and humans. Due to its capacity to swiftly decrease serum uric acid levels, uricase emerged as a promising therapeutic option for addressing hyperuricemia. In 1975, the use of uricase from *Aspergillus flavus* received initial approval in France for treating hyperuricemia induced by tumor lysis syndrome. However, due to its fungal source, administering the medication often led to allergic reactions and hypersensitivity. To circumvent these adverse effects, scientists developed a recombinant version of uricase from *Aspergillus flavus* produced in *Saccharomyces cerevisiae* [9]. *Aspergillus flavus*, *Bacillus subtilis*, and porcine liver are a commercial source of uricase and are used extensively for the treatment of hyperuricemia [10]. Hyperuricemic diseases have been treated with recombinant *Aspergillus flavus* uricase (Rasburicase), but the possibility of immunogenicity limits its use. Clinical studies using PEGylated uricase, which has a long circulation half-life and a low immunogenicity profile, date back to the 1980s. Fungal uricasases are the primary focus in investigations involving PEGylation. However, thus far, there have been no commercially accessible PEGylated microbial uricasases for further clinical investigation. Long-term PEGylated uricase production is best accomplished with mammalian uricase; however, a PEGylated porcine-like recombinant uricase (Pegloticase) has been available since September 14, 2010. However, in clinical studies of Pegloticase, repeated dosing resulted in antibody development in 92% of patients and decreased urate-lowering efficacy in 58% of patients. The immunogenicity of PEGylated uricase may be the biggest barrier to their practical application [10-12].

A homo-tetramer of active uricase has been found in all species. One-third of the residues are hydrophobic, and the tetramers tend to group into octomers or larger groups. Denatured monomeric uricase is the only form that can be evaluated by SDS-PAGE and reversed phase-high performance liquid chromatography (RP-HPLC) for homogeneity. However, the contents of high molecular weight aggregates are highly immunogenic. Most published works neglect the significance of native aggregated uricase content [11,13]. Uric acid in blood and other biological fluids

was initially measured using uricase in conjunction with the 4-amino-antipyrine-peroxidase method. Uricase serves as a protein-based medication aimed at diminishing the buildup of harmful urate in addressing hyperuricemia and gout, along with its use in preventing and treating tumor lysis syndrome-induced hyperuricemia [14,15]. Uricase that has been immobilized can function as a biosensor for uric acid. Compared to other urate-lowering medications, such as allopurinol, uricase administration was found to be more effective; it decreased increased serum urate concentrations. Hyperuricemia from tumor lysis or organ transplantation is efficiently treated with rasburicase. It is also included as a component in many commercial hair dyes [16-18].

A therapeutic uricase has been previously characterized, including its crystal structure and catalytic mechanism. The functional molecule can be described as a non-covalent tetramer, wherein its four active sites are located at the interface between two subunits. Each subunit consists of two domains, or building blocks, that are physically identical and share a common structural fold known as a tunneling fold (T-fold). A porin-like structure is formed when two identical subunits join together symmetrically to form a dimer. The active enzyme consists of two dimers opposite and connected by a double crystallographic axis. The catalytic reaction requires only oxygen and water molecules; no metal ions or prosthetic groups are involved. Arginine, glutamine, and phenylalanine, all considered highly conserved active site residues, are essential for catalytic activity. The enzyme has been compared to a molecular tweezer, with uric acid playing the role of a permanent prosthetic group and molecular oxygen serving as the enzyme actual substrate [6,19,20].

In the quest to enhance the biostability of uricase from *Aspergillus flavus*, Najjari *et al.* [21] have utilized computational methods to create multiple uricase PAS fusion variants. These variants are formed by combining various PAS polypeptides, and their physicochemical properties and structural activity have been assessed through *in-silico* analysis. Tripathi *et al.* [10] directed their attention toward unraveling the structural characteristics using an immunoinformatics methodology, which sheds light on the allergenic and immunogenic attributes of the biopharmaceutical enzyme uricase derived from *Aspergillus flavus*, *Bacillus subtilis*, and mammalian origins.

Wet-lab enzyme structural and functional analysis is

time-consuming and expensive compared to bioinformatics [22]. Using the sequence, structural, functional, and evolutionary data from computational genomics and proteomics investigations, the unknown protein profiles can be understood with the help of bioinformatics web-based servers and tools [23-25]. The three-dimensional structure of a protein is essential for studying its function. Protein modeling predicts unknown protein structures. Using the protein sequences of different industrially important enzymes, *in silico* study and characterization have become more important. Determining the 3D structures of proteins experimentally presents significant challenges due to its complexity, difficulty, as well as the considerable expenses and time required. Consequently, alternative approaches must be taken into consideration. In this context, bioinformatics tools hold significant appeal and are extensively utilized for predicting 3D protein structures in numerous instances, as well as for gene analysis [26]. Until now, there has been no study on the structure and function of uricase from different fungal species. The primary goal of the current study was to use various bioinformatics tools to characterize the physicochemical features, phylogenetic relationships, domains, motifs, secondary and tertiary structure prediction, and functional analysis of eighty-eight uricase protein sequences derived from different fungal species.

## MATERIALS AND METHODS

### Sequence Retrieval

Eighty-eight amino acid sequences of uricase protein from various fungus species were obtained in FASTA format from NCBI (<http://www.ncbi.nlm.nih.gov/>). The cDNA corresponding to each amino acid sequence was generated using a reverse translation tool ([http://www.bioinformatics.org/sms2/rev\\_trans.html](http://www.bioinformatics.org/sms2/rev_trans.html)) to compare the evolutionary relatedness of the taxa through phylogenetic analysis [27].

### Multiple Sequence Alignment and Phylogenetic Analysis

Multiple sequence alignment of the selected fungal sequences was carried out using ClustalW [28,29] in MEGA 7 [30] software with default settings to assess the extent of

similarity among uricase in various fungal species. This analysis aimed to investigate the conserved residues and ascertain the sequence similarity of enzymes belonging to fungal species.

The primary focus of performing phylogenetic analysis was to study the evolutionary relationships among eighty-eight uricase sequences. Phylogenetic analysis was performed using a statistical technique called Neighbor-joining (NJ) method in MEGA 7 based on the p-distance model. Phylogenetic tree topologies were assessed using 1000 bootstrap replicates for each sequence [29,31].

### Physio-Chemical Characterization

The ExPASy ProtParam tool (<https://web.expasy.org/protam/>) was used to predict the physicochemical properties of selected fungal uricase sequences. Several different physicochemical parameters, such as the theoretical isoelectric point (pI), extinction coefficient (EC), instability index (Ii), aliphatic index (Ai), grand average hydrophobicity (GRAVY), number of negative residues (-R), and number of positive residues (+R), were computed [32-34].

### Motif Analysis

The tool MEME (Multiple EM for Motif ELICitation) (<https://meme-suite.org/meme/tools/meme>) is widely used to identify motifs in protein or DNA sequences. This tool is beneficial for doing motif database searches, motif-sequence database searches, and the identification of motifs, among other functions. The blocks also exhibit the presentation of the motifs' starting and ending points. The conserved areas, which are linked to the structural and functional features of the enzymes through evolution, are described by the location of motifs or sites discovered in the MEME [35]. The Pfam database is utilized to analyze uricase domains, with sequences uploaded to Pfam in UniProt numbers. The domain organization associated with the uricase protein was verified for all sequences [36].

### Secondary Structure Analysis

Secondary structures were predicted using online servers namely PSIPRED v3.3 (<http://bioinf.cs.ucl.ac.uk/psipred/>) CFSSP (<https://www.biogem.org/tool/chou-fasman/>) and SOPMA server (<https://npsa->

prabi.ibcp.fr/NPSA/npsa\_sopma.html) which count the number of helices, sheets, turns, and coils in the sequence. The number of secondary structural elements has a direct effect on protein folding [27,37-39].

### Phyre2 Protein Modeling, Prediction and Analysis

Phyre2, also known as the Protein Homology/Analogy Recognition Engine Version 2.0, is a computational tool widely employed for various protein structure and function predictions. Uricase from *Aspergillus niger* (ABI79319.1) as a representative of fungal species was submitted to the online tool (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) which uses the existing model to predict protein sequences. Secondary and tertiary structure, domain composition, and model quality are studied using the protein sequence for the chosen uricase [23,40].

### Homology Modeling and Model Evaluation

Protein function is defined by its tertiary structure, composed of the aggregate of the proteins secondary structure. Among the 88 different species of fungal, the uricase protein sequence from *Aspergillus niger* was selected as the reference for tertiary structure prediction. The SWISS-MODEL server (<https://swissmodel.expasy.org/>) was utilized in automated mode to produce 3D protein homology models for a total of 88 protein sequences. Among these, the protein sequence of *Aspergillus niger* (ABI79319.1) was selected based on its QMEAN value, which closely approximated zero [38]. This selection was made by identifying the best suitable template. SWISS-MODEL is a homology-modeling service that can predict 3D protein structures. It is one of the most well-known and widely used free web servers. The predicted model accuracy was evaluated and confirmed using SAVES server (<https://saves.mbi.ucla.edu/>). The predicted structure was submitted to ERRAT, and the Ramachandran plot results in the best overall quality factor. The model backbone structural regions were examined using Ramachandran plot analysis [41,42].

### Functional Analysis

Cysteine is one of a kind among the amino acids due to its reactive sulfhydryl group. Enzymes like thiol-disulfide

oxidoreductase and protein disulfide isomerase help generate disulfide bonds between proteins during protein folding. The CYS\_REC tool ([https://www.softberry.com/berry.phtml?topic=cys\\_rec&group=programs&subgroup=propt](https://www.softberry.com/berry.phtml?topic=cys_rec&group=programs&subgroup=propt)) was utilized to analyze disulfide bonds and identify the precise positions of cysteine residues [43]. Additionally, this tool performs computations on the protein sequence to determine the pairings of SS bond patterns and evaluates the occurrence of SS bonds within a protein. The selected uricase protein sequence was analyzed for functional motifs and superfamily using the Motif search tool (<https://www.genome.jp/tools/motif/>) [27]. The identification of conserved domains within the selected uricase protein was conducted utilizing the conserved domain database (CDD) (<https://www.ncbi.nlm.nih.gov/cdd/>) provided by NCBI [23]. The interaction network of *Aspergillus niger* uricase was examined using the STRING v11.0 server (<https://string-db.org/>) [38]. STRING stores information about observed and expected interactions between proteins [44]. SOSUI (<http://www.tuat.ac.jp/mitaku/sosui/>) was utilized to ascertain whether a given amino acid sequence constitutes a transmembrane helix and whether a protein is soluble or transmembrane [43]. This server examined all of the retrieved fungal uricase sequences [45]. The Peptide Cutter tool ([https://web.expasy.org/peptide\\_cutter/](https://web.expasy.org/peptide_cutter/)) identifies regions of a protein sequence that may serve as cleavage sites for proteases or other compounds [26].

## RESULTS AND DISCUSSION

The sequence of amino acids in a protein determines its structure, and this structure, in turn, governs its biochemical function. Proteins with similar amino acid sequences typically carry out similar biochemical functions, even across distantly related organisms. Consequently, the process of elucidating the function of a newly discovered protein often commences with a search for previously characterized proteins that exhibit similarity in their amino acid sequences. In the past, sequence-based inference served as the initial approach for predicting protein function, although many studies at that time refrained from explicitly correlating homology with function [46].



**Table 1.** Distribution of Motifs in Fungal Species Uricase Amino Acid Sequences and Pfam Analysis

S.No.	Source	Motif width	E-value	Motifs present in a number of sequences	Motif	Pfam
1	Fungi	50	3.8e-3259	88	QNPVTPPELFASILGTHFIEKY SHIH AAHVNIITHRWTRMTIDGKPH PHS	Uricase
2	Fungi	41	7e-2967	88	GLTVLKSTGSAFHGFVRDEY TTLPETWDRILSTDVDATWQ W	Uricase
3	Fungi	37	2.1e-2485	86	QATMYKMAEQILAAVPEVET VEYSLPNKHIFEIDLSW	Uricase
4	Fungi	29	5e-2066	88	CVLLEGIETSYTKADNSVV VATDSIKNT	Uricase
5	Fungi	29	2.0e-1775	79	KGLKNTGKDAEVYAPQSDPN GLIKCTVSR	Not found
6	Fungi	29	2.0e-1584	86	MPYLSAARYGKDNVRVYKV HRDEKTGVQT	Not found

sequences. The E-value of each of the six motifs was reduced by selecting them based on their width and frequency of occurrence. The width of the motif was specified to be between 6 and 50 (Table 1). Supplementary Table 2 shows that all six patterns (regions) were widely distributed among uricase protein sequences from various fungal organisms. Each of the six discovered motifs was unique within their respective groupings. The motifs may play a role in uricases catalytic properties and structural organization. Domain analysis results from the Pfam database suggested a single domain structure for uricase protein sequences from different organisms related to the uricases. Several authors performed similar studies with a wide variety of proteins [36,50].

### Physio-Chemical Characterization of The Retrieved Sequences

A total of 88 fungal species were characterized based on physicochemical criteria. The parameters encompass the number of amino acid residues, molecular weight, theoretical pI, extinction coefficient, instability index, aliphatic index, and GRAVY. Supplementary Table 3 displays the results of a ProtParam calculation of the various features of uricase from different fungal species. The sequences chosen for

analysis encompassed amino acid residues spanning from 296 to 330. The sequences molecular weight ranged from 33.03 to 35.05 kDa. The extinction coefficient (EC) of the aforementioned proteins fell between 22140 and 59610  $M^{-1} cm^{-1}$ . The aliphatic index ranges from 76.72 to 95.02, which indicates the thermostability of globular proteins [51]. The isoelectric point (pI) denotes the condition of a solution wherein an amino acid demonstrates an equilibrium between positive and negative charges, yielding a zero net charge. The pI values of uricase from various fungal species were calculated to be between 5.71 and 9.03. A pI number above seven indicates an alkaline protein, while a pI value below 7 indicates an acidic protein. Here, the GRAVY range was between -0.080 and -0.466. The lower range of GRAVY showed that the association between protein and water was stronger [33]. Supplementary Table 4 shows all fungal uricase protein sequence amino acid compositions. The findings of our study were supported by previous researchers who had observed comparable physicochemical features in numerous other microbial proteins [29,41].

### Secondary Structure Analysis

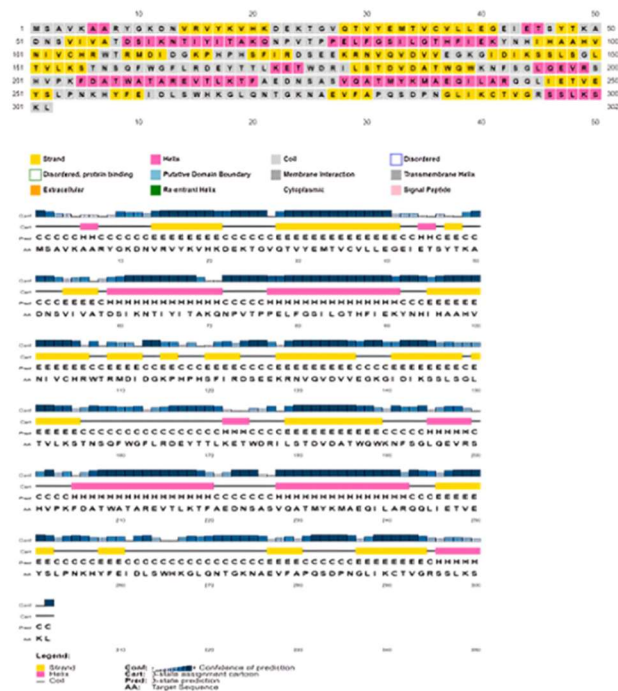
Proteins in their natural state typically exhibit highly



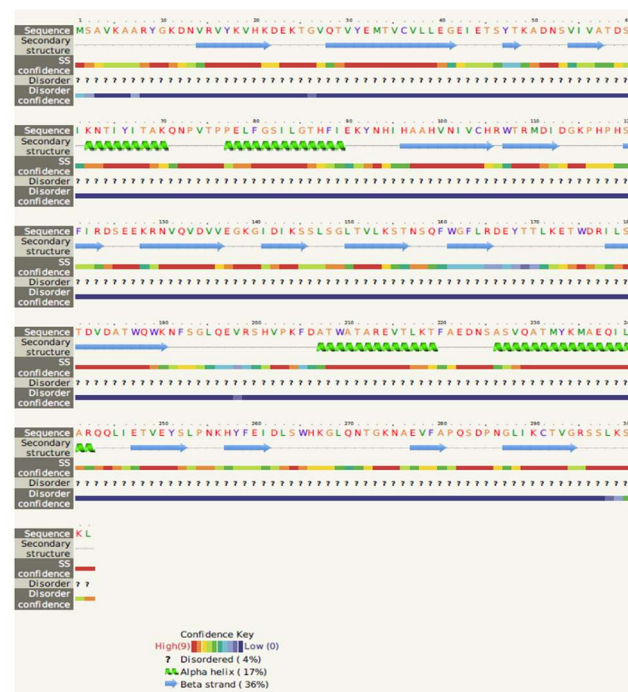
ordered patterns of structure known as secondary structure. About half of the total protein structure is comprised of secondary structure. The secondary structure holds significant importance within the hierarchical categorization of protein structure, as it serves to identify protein characteristics for fold recognition. To anticipate the protein three-dimensional structure, it is necessary to first determine its secondary structure. The significance of secondary structure in protein structure and protein folding has been the subject of several studies. The protein folding rate was investigated by [52] who discovered that topological features of the native structure significantly affect the folding rate constants for proteins of different secondary-type content. Supplemental Table 5 displays the results of an analysis of the CFSSP and SOPMA tools for predicting secondary structural elements in uricase from various fungal species. Alpha helices (34.44%), extended strands (23.51%), beta turns (5.30%), and random coils (36.75%) were the most commonly detected secondary protein elements using SOPMA. In CFSSP, the secondary structure prediction result shows that the percentage of the alpha helix (81.5) was much greater than the other conformation, such as sheet (67.5) and turn (15.2). A secondary structure that is more stable is likely to have higher amounts of alpha-helical conformation. As seen in (Fig. 2), the PSIPRED protein analysis tool was used to predict the uricase secondary structure map. There were no binding sites for disordered proteins found. The predicted secondary structure of the amino acid sequences of *Aspergillus niger* uricase (ABI79319.1) is depicted in detail in the supplementary Fig. 4. Multiple authors have performed comparable secondary structure analyses on different proteins [37,49].

### Protein Modeling, Prediction, and Analysis

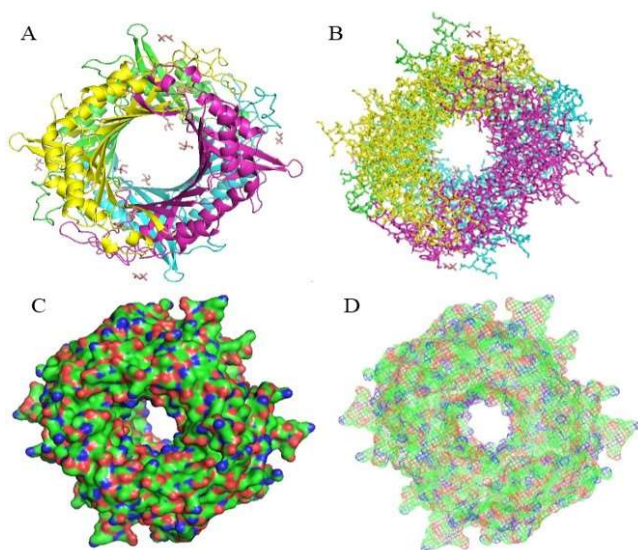
Uricase from *Aspergillus niger* (ABI79319.1) was employed for the  $\beta$  sheet model construction. The Phyre 2 tool was utilized to generate the model, which exhibited dimensions of (Å). The coordinates provided by the user are as follows: X: 49.447, Y: 61.739, Z: 78.420 (Supplementary Fig. 5). Based on the results of the BLAST search, the c1r56H template was selected as the most appropriate option for protein modeling. The crystal structure of urate oxidase from *Aspergillus flavus*, in its uncomplexed form, served as the PDB input. The highest-scoring template anticipated the



**Fig. 2.** PSIPRED map study of the secondary structure of *Aspergillus niger* (ABI79319.1).



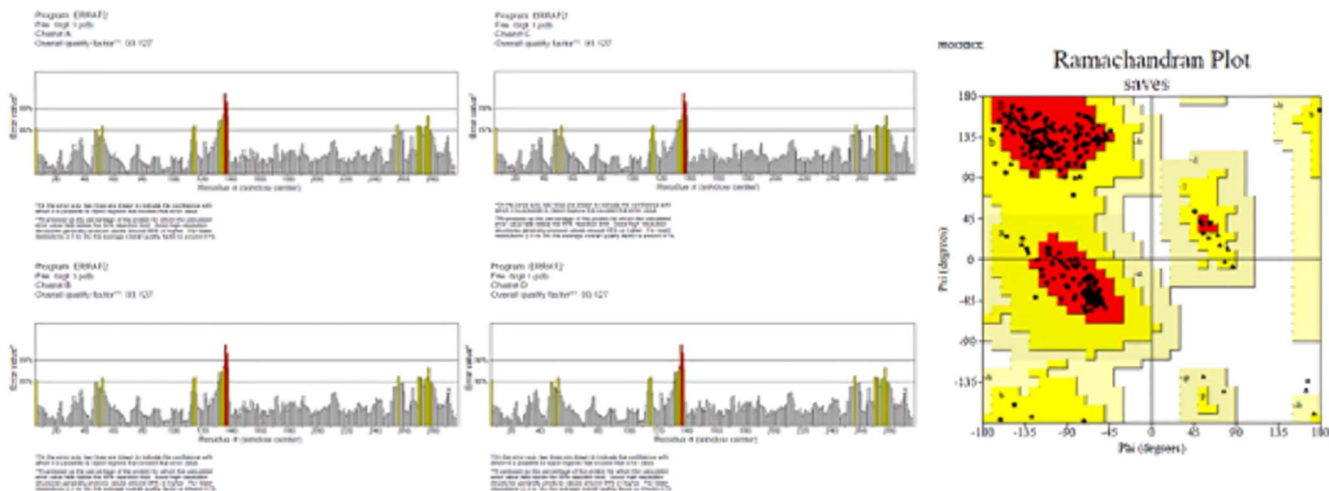
**Fig. 3.** Prediction of the secondary structure and disorder of uricase from *Aspergillus niger* (ABI79319.1) using the Pyre2 server.



**Fig. 4.** The predicted 3D model structures of the uricase protein from *Aspergillus niger* (ABI79319.1) in different views by PyMol. (A) and (B) show four separate chains of the protein (C) the surface view and (D) the mesh view.

secondary structure of uricase and modeled 97% of residues with 100.0% confidence. The analysis of the secondary structure of the uricase enzyme revealed that it possesses 4% disordered regions, 17% alpha-helix structures, and 36% beta-strand formations, as illustrated in Fig. 3.

The interesting part of comparing two uricases, such as uricase from *Bacillus fastidiosus* and the modeled *Aspergillus niger* three-dimensional structures of different microbial origins is their shape: tunnel-shaped in bacteria and barrel-shaped in fungus. A secondary structure analysis of *Aspergillus niger* uricase was conducted (Supplementary Fig. 5). Interestingly, the monomeric structure of uricase from *Aspergillus niger* exhibits a secondary structure similar to that of rasburicase from *Aspergillus flavus*, both in terms of the number and arrangement of alpha helices and beta sheets. These findings are unsurprising given the high level of identity and similarity observed between their sequences. The Phyre 2 tool facilitates tasks such as identifying and determining domain boundaries, conducting site-directed mutagenesis, and classifying proteins based on evolutionary relationships. Additionally, it may be employed for the molecular replacement of protein crystal structures. This protein bioinformatics tool is generally recognized as one of the most commonly utilized protein structure prediction servers, with over 1500 citations. It is known for its user-friendly interface and ease of use. This software can construct accurate three-dimensional protein models using distant homology detection techniques. These computational models are useful for predicting protein-ligand binding sites and assessing the effects of individual changes to amino acid sequences [23,40].



**Fig. 5.** *Aspergillus niger* (ABI79319.1) modeled uricase protein validation from the SAVES server using ERRAT and Ramachandran plot.



## Homology Modeling and Model Evaluation

The 3-dimensional structure of uricase from *Aspergillus niger* (ABI79319.1) was generated using the SWISS-MODEL system, employing the template that best matched the protein sequence. The correspondence between the chosen template (6rgt.1. A) and the target sequence of *Aspergillus niger* (ABI79319.1) is depicted in Supplementary Fig. 6, while the protein models in Fig. 4 provide a visual representation of this alignment. The template that exhibited the highest degree of similarity (6rgt.1. A) corresponded to the crystal structure of the cofactor-free *Aspergillus flavus* urate oxidase T57A variant anaerobically complexed with 9-methyl uric acid. The template exhibits an identity of 99.67% and a similarity of 0.62. It contains four molecules of 9-methyl uric acid. The resolution of the template, obtained using the X-ray diffraction method, is 1.60 Å. Furthermore, the BLAST analysis reveals that the template exhibits an oligomeric state of homo-tetramer. The QMEAN score of the constructed model was determined to be 0.90, signifying the overall quality of the entire model. The examination of the


Ramachandran plot revealed that the majority of residues, specifically 92.7%, are located inside the favorable region. Additionally, a smaller proportion of residues, accounting for 7.3%, were found within the allowed region. According to the Ramachandran plot, the best model has more than 90% of its residues in the favorable region. The uricase protein model

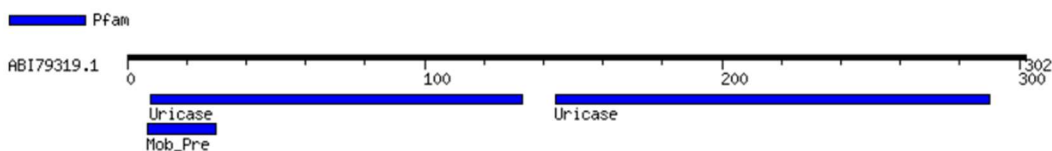
selected for analysis exhibited a maximum overall quality factor value of 93.17 in chain A, as determined by SAVES ERRAT (Fig. 5). The investigation of homology modeling and the subsequent validation of diverse proteins has been examined by numerous researchers [29,53,54].

## Functional Analysis

Protein function embodies a multifaceted concept encompassing "everything that happens to or through a protein," and is conventionally examined from biochemical, biological, and phenotypic standpoints. Understanding the function of proteins is crucial to comprehending existence at the molecular level. In the context of human disease, understanding protein function is crucial as many conditions result from changes or abnormalities in protein function [46]. The SOSUI web server assessed the solubility of all selected fungal uricase species. These results are in Supplementary Table 6. The production of disulfide bonds in proteins, which occurs through the oxidation of thiol groups present in cysteine residues, is a pivotal mechanism that contributes significantly to the protein's thermostability. Disulfide bridge locations and cysteine SS-bond states in fungal uricases were predicted using the CYS\_REC tool. The study found that uricase sequences, containing cysteine residues in positions 1-4, did not exhibit SS-bonding, which could affect their stability (Supplementary Table 7). The uricase selected from *Aspergillus niger* showed three cysteine residues at

## Result of MotifFinder

Number of found motifs: 2 



Pfam (2 motifs)

Pfam	Position(Independent E-value)	Description
Uricase	8..133(1.5e-31) 144..290(1.4e-49) <a href="#">Detail</a>	PF01014, Uricase
Mob_Pre	7..30(0.23) <a href="#">Detail</a>	PF01076, Plasmid recombination enzyme

**Fig. 6.** Motif finder tool result for the uricase from *Aspergillus niger* (ABI79319.1) displaying two functional motifs and their locations.



**Fig. 7.** The protein-protein interaction map depicts the predicted proteins that interact with the uricase protein (ABI79319.1).

positions 36, 104, and 291, which were found to be unbound by disulfide bonds. Cysteine residues play a critical role in conferring thermostability to proteins, whereas disulfide bonds are essential for protein folding. Furthermore, the functional analysis revealed the presence of two distinct functional motifs within the protein sequence under investigation, namely uricase and mob\_pre (Fig. 6). The STRING tool identified ten predicted functional partners in the protein-protein interaction network of *Aspergillus niger* uricase (Fig. 7). Proteins typically operate by engaging in interactions with other proteins, forming intricate protein complexes and networks. Unraveling these complex protein interactions provides valuable insights into the functions of novel proteins that regulate cellular behavior [37]. The uricase from *Aspergillus niger* is a member of a single protein superfamily which is a urate oxo superfamily with a single type of conserved domain found from the conserved domain analysis (Supplementary Fig. 7) Proteinase K, pepsin, chymotrypsin, and thermolysin were used to perform *in silico* hydrolysis of *Aspergillus niger* uricase. PeptideCutter found that proteinase K has an average of 148.18 cutting sites, pepsin (pH > 2) has 165.98, chymotrypsin-low specificity has 184.43, and thermolysin has 142.56 cutting sites (Supplementary Fig. 8). Various authors previously worked out similar functional studies of various proteins [37,43].

Nelapati *et al.* [17] conducted a study focusing on the structure and properties of uricase from various organisms using *in silico* analysis. They performed multiple sequence alignment of protein sequences from different sources, revealing differences in conserved amino acid residues among bacterial, fungal, plant, and animal uricase. Similarity

among selected sequences ranged from 51-388, with bacterial uricase showing homology between 16-337, fungal uricase between 14-339, plant uricase between 12-317, and animal uricase between 37-361. Phylogenetic analysis grouped uricase sequences into distinct clusters corresponding to their sources. Physicochemical analysis indicated amino acid residues within the range of 300-338, molecular weights from 33-39 kDa, and theoretical pI values from 4.95 to 8.88. Valine was the most prevalent amino acid, averaging 8.79% frequency across species. The same author analyzed uricase enzymes from various *Bacillus* species computationally, revealing activity in acidic to neutral conditions, thermal stability, and molecular weights ranging from 35.59 to 59.85 kDa. Secondary structure analysis showed alpha-helices and sheets predominance. *Bacillus simplex* (WP\_063232385.1) was chosen as a representative species, with a modeled protein quality score of 94.64. A strong correlation among *Bacillus* species in uricase enzymes and genes was observed through phylogenetic analysis [23]. In our current investigation, *in silico* analysis of amino acid sequences from diverse fungal species was conducted to establish correlations based on shared characteristics.

## CONCLUSION

In this study, attempts were undertaken to assess a comprehensive analysis of uricase protein sequences from various fungal species using bioinformatics tools. Through multiple sequence analysis and homology searches, it was noted that the selected sequences displayed similarities and found highly conserved amino acids. Phylogenetic analysis of the chosen sequences from various fungal sources revealed

distinct clusters, indicating sequence similarity correlated with the organism source. The cluster analysis conducted on all retrieved protein and cDNA sequences provided a clear comprehension of the evolutionary connections among various fungal species of uricase at the molecular level. The analysis of computational physicochemical features of all the chosen uricase enzymes provided a comprehensive understanding of their properties, including pI, EC, Ai, Ii, and GRAVY, which predominantly exhibited basic characteristics. These enzymes typically had molecular weights ranging from 33 kDa to 35 kDa. The amino acid valine has an average frequency of 8.40% across all fungal species, showing its pivotal role in uricase formation. The secondary structure study indicates that uricase exhibits a high abundance of alpha-helices and beta-sheets. Uricase serves as a beneficial treatment for hyperuricemia and gout. Hence, comprehending both the structural and functional characteristics of the protein will enhance our understanding of its enzyme action mechanism. Given the unavailability of the crystal structure, investigating the *in-silico* structure-function aspects of the protein seems to be a guiding light in the darkness. The biological functions may be better understood with the aid of the predicted three-dimensional structure. Further, the study on the molecular docking of uricase with its substrate uric acid to analyze binding interactions along with simulations indicates the future scope of the present research. Therefore, the current investigation has the potential to contribute to the field of computational proteomics by enhancing our comprehension of the uricase protein. Future researchers could use the results of this study to design and implement wet lab experiments to learn more about the structure, function, isolation, and characterization of the fungal uricase enzyme and its potential uses in industry.

## ACKNOWLEDGMENTS

The authors express thanks to VFSTR (Deemed to be University)

## REFERENCES

- [1] Piermarini, S.; Migliorelli, D.; Volpe, G.; Massoud, R.; Pierantozzi, A.; Cortese, C., Uricase biosensor based on a screen-printed electrode modified with Prussian blue for detection of uric acid in human blood serum. *Sens. Actuators B: Chem.*, **2013**, *179*, 170-4. DOI: 10.1016/j.snb.2012.10.090.
- [2] Kilimci, U.; Uygun, D. A., Preparation of PEGylated uricase attached magnetic nanowires and application for uric acid oxidation. *J. Biotechnol.*, **2023**, *373*, 12-9. DOI: 10.1016/j.jbiotec.2023.06.005.
- [4] Dudala, S. S.; Venkateswarulu, T. C.; Venkata Narayana, A.; John Babu, D., Modeling and optimization of uricase production from a novel *Pseudomonas mosselii* using response surface methodology and artificial neural network. *Biomass Conv Bioref.*, **2023**. DOI: 10.1007/s13399-023-04468-3.
- [5] Zhang, C.; Fan, K.; Zhang, W.; Zhu, R.; Zhang, L.; Wei, D., Structure-based characterization of canine-human chimeric uricases and its evolutionary implications. *Biochimie.*, **2012**, *94* (6), 1412-20. DOI: 10.1016/j.biochi.2012.03.016.
- [6] Gruia, F.; Parupudi, A.; Baca, M.; Ward, C.; Nyborg, A.; Remmele, R. L.; Bee, J. S., Impact of Mutations on the Higher Order Structure and Activity of a Recombinant Uricase. *J. Pharm. Sci.*, **2017**, *106*, 1018-1024. DOI: 10.1016/j.xphs.2016.12.028.
- [7] Noma, S. A. A., Investigation of Improved Uricase Release and Kinetic Parameters through Dual Affected Responsive Nanopolymers. *Process Biochem.*, **2023**, *131*, 52-58. DOI: 10.1016/j.procbio.2023.06.003.
- [8] Colloc'h, N.; el Hajji, M.; Bachet, B.; L'Hermite, G.; Schiltz, M.; Prangé, T.; Castro, B.; Mornon, J. P., Crystal Structure of the Protein Drug Urate Oxidase-Inhibitor Complex at 2.05 Å Resolution. *Nat. Struct. Biol.*, **1997**, *4*, 947-952. DOI: 10.1038/nsb1197-947.
- [9] Yainoy, S.; Phuadraksa, T.; Wichit, S.; Sompoppokakul, M.; Songtawee, N.; Prachayasittikul, V.; Isarankura-Na-Ayudhya, C., Production and Characterization of Recombinant Wild Type Uricase from Indonesian Coelacanth (*L. Menadoensis*) and Improvement of Its Thermostability by *In Silico* Rational Design and Disulphide Bridges Engineering. *Int. J. Mol. Sci.*, **2019**, *20* (6), 1269. DOI: 10.3390/ijms20061269.
- [10] Tripathi, S.; Parmar, J.; Kumar, A., Structure-Based Immunogenicity Prediction of Uricase from Fungal (*Aspergillus Flavus*), Bacterial (*Bacillus Subtilis*) and

- Mammalian Sources Using Immunoinformatic Approach. *Protein J.*, **2020**, *39* (2), 133-144. DOI: 10.1007/s10930-020-09886-0.
- [11] Zhang, C.; Fan, K.; Ma, X.; Wei, D., Impact of Large Aggregated Uricases and PEG Diol on Accelerated Blood Clearance of PEGylated Canine Uricase. *PLoS One.*, **2012**, *7* (6), e39659. DOI: 10.1371/journal.pone.0039659.
- [12] Nanda, P.; Jagadeesh Babu, P. E., Studies on the Site-Specific PEGylation Induced Interferences Instigated in Uricase Quantification Using the Bradford Method. *Int. J. Pept. Res Ther.*, **2016**, *22*, 399-406. DOI: 10.1007/s10989-016-9518-8.
- [13] Punnappuzha, A.; Ponnannettiappan, J.; Nishith, R. S.; Hadigal, S.; Pai, P. G., Synthesis and Characterization of Polysialic Acid-Uricase Conjugates for the Treatment of Hyperuricemia. *Int. J. Pept. Res. Ther.*, **2014**, *20*, 465-472. DOI: 10.1007/s10989-014-9411-2.
- [14] Nanda, P.; Babu, P. E. J., Isolation, Screening and Production Studies of Uricase Producing Bacteria from Poultry Sources. *Prep. Biochem. Biotechnol.*, **2014**, *44* (8), 811-821. DOI: 10.1080/10826068.2013.867875.
- [15] Khade, S.; Srivastava, S. K., Uricase and Its Clinical Applications., *Int. J. Biol. Med. Res.* **2015**, *6*, 5211-5215.
- [16] El-Naggar, N. E. -A.; Haroun, S. A.; El-Wesly, E. M.; Metwally, E. A.; Sherief, A. A., Mathematical Modeling for Bioprocess Optimization of a Protein Drug, Uricase, Production by *Aspergillus Welwitschiae* Strain 1-4. *Sci. Rep.*, **2019**, *9* (1), 12971. DOI: 10.1038/s41598-019-49201-1.
- [17] Nelapati, A. K.; Ponnannettiappan, J., Computational Analysis of Therapeutic Enzyme Uricase from Different Source Organisms. *Curr. Proteom.*, **2019**, *17* (1), 59-77. DOI: 10.2174/1570164616666190617165107.
- [18] Nanda, P.; Jagadeesh Babu, P. E. Solid Phase PEGylation of Uricase. *Materials Today: Proceedings*, **2017**, *4*, 10494-10497. DOI:10.1016/j.matpr.2017.06.407
- [19] Nelapati, A. K.; Das, B. K.; Ponnannettiappan, J. B.; Chakraborty, D., *In-Silico* Epitope Identification and Design of Uricase Mutein with Reduced Immunogenicity. *Process Biochem.*, **2020**, S1359511319311808. DOI: 10.1016/j.procbio.2020.01.022.
- [20] Nelapati, A. K.; Meena S, K., An Approach to Increase the Efficiency of Uricase by Computational Mutagenesis. *Phys. Chem. Res.*, **2023**, *11* (3), 481-491. DOI: 10.22036/PCR.2022.345329.2115.
- [21] Najjari, A.; Rahimi, H.; Nojoumi, S. A.; Omidinia, E., Computational Approach for Rational Design of Fusion Uricase with PAS Sequences. *Int. J. Mol. Cell. Med.*, **2020**, *9* (1), 90-103. DOI: 10.22088/IJMCM.BUMS.9.1.90.
- [22] Mohseni-Shahri, F. S.; Moeinpour, F.; Malaekhe-Nikouei, B.; Nassirli, H., Combined Multispectroscopic and Molecular Dynamics Simulation Investigation on the Interaction between Cyclosporine A and  $\beta$ -Lactoglobulin. *Int. J. Biol. Macromol.*, **2017**, *95*, 1-7. DOI: 10.1016/j.jbiomac.2016.10.107.
- [23] Nelapati, A. K.; Meena, S.; Singh, A. K.; Bhakta, N.; Ponnannettiappan, J., *In Silico* Structural and Functional Analysis of *Bacillus* Uricases. *Curr. Proteom.*, **2020**, *18* (2), 124-142. DOI: 10.2174/1570164617999200512081127.
- [24] Pramanik, K.; Soren, T.; Mitra, S.; Maiti, T. K., *In Silico* Structural and Functional Analysis of *Mesorhizobium* ACC Deaminase. *Comput. Biol. Chem.*, **2017**, *68*, 12-21. DOI: 10.1016/j.compbiolchem.2017.02.005.
- [25] Moeinpour, F.; Mohseni-Shahri, F. S.; Malaekhe-Nikouei, B.; Nassirli, H., Investigation into the Interaction of Losartan with Human Serum Albumin and Glycated Human Serum Albumin by Spectroscopic and Molecular Dynamics Simulation Techniques: A Comparison Study. *Chem. Biol. Interact.*, **2016**, *257*, 4-13. DOI: 10.1016/j.cbi.2016.07.025.
- [26] Hoda, A.; Tafaj, M.; Sallaku, E., *In Silico* Structural, Functional and Phylogenetic Analyses of Cellulase from *Ruminococcus Albus*. *J. Gene. Eng. Biotechnol.*, **2021**, *19* (1), 58. DOI: 10.1186/s43141-021-00162-x.
- [27] Pramanik, K.; Ghosh, P. K.; Ray, S.; Sarkar, A.; Mitra, S.; Maiti, T. K., An *in Silico* Structural, Functional and Phylogenetic Analysis with Three-Dimensional Protein Modeling of Alkaline Phosphatase Enzyme of *Pseudomonas Aeruginosa*. *J. Genet. Eng. Biotechnol.*, **2017**, *15* (2), 527-537. DOI: 10.1016/j.jgeb.2017.05.003.
- [28] Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G., Clustal W and Clustal X

- Version 2.0. *Bioinformatics.*, **2007**, 23 (21), 2947-2948. DOI: 10.1093/bioinformatics/btm404.
- [29] Tamboli, A. S.; Rane, N. R.; Patil, S. M.; Biradar, S. P.; Pawar, P. K.; Govindwar, S. P., Physicochemical Characterization, Structural Analysis and Homology Modeling of Bacterial and Fungal Laccases Using *in Silico* Methods. *Netw. Model. Anal. Health Inform. Bioinform.*, **2015**, 4, 17. DOI: 10.1007/s13721-015-0089-y.
- [30] Kumar, S.; Stecher, G.; Tamura, K., MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.*, **2016**, 33 (7), 1870-1874. DOI: 10.1093/molbev/msw054.
- [31] Dubey, A. K.; Yadav, S.; Kumar, M.; Singh, V. K.; Sarangi, B. K.; Yadav, D., *In Silico* Characterization of Pectate Lyase Protein Sequences from Different Source Organisms. *Enzyme Res.*, **2010**, 2010950230, 1-11. DOI: 10.4061/2010/950230.
- [32] Saleem, A.; Rajput, S., Insights from the *in Silico* Structural, Functional and Phylogenetic Characterization of Canine Lysyl Oxidase Protein. *J. Genet. Eng. Biotechnol.*, **2020**, 18, 20. DOI: 10.1186/s43141-020-00034-w.
- [33] Rahmatabadi, S. S.; Sadeghian, I.; Nezafat, N.; Negahdaripour, M.; Hajighahramani, N.; Hemmati, S.; Ghasemi, Y., *In Silico* Investigation of Pullulanase Enzymes from Various *Bacillus* Species. *Curr. Proteom.*, **2017**, 14 (3), 175-185. DOI: 10.2174/1570164614666170306164830.
- [34] Dwivedi, V. D.; Arora, S.; Kumar, A.; Mishra, S. K., Computational Analysis of Xanthine Dehydrogenase Enzyme from Different Source Organisms. *Netw. Model. Anal. Health Inform. Bioinform.* **2013**, 2, 185-189. DOI: 10.1007/s13721-013-0029-7.
- [35] Bailey, T. L.; Williams, N.; Misleh, C.; Li, W. W., MEME: Discovering and Analyzing DNA and Protein Sequence Motifs. *Nucleic Acids Res.*, **2006**, 34, W369-373. DOI: 10.1093/nar/gkl198.
- [36] Ramya, L. N.; Pulicherla, K. K., Molecular Insights into Cold Active Polygalacturonase Enzyme for Its Potential Application in Food Processing. *J. Food Sci. Technol.*, **2015**, 52 (9), 5484-5496. DOI: 10.1007/s13197-014-1654-6.
- [37] Pramanik, K.; Kundu, S.; Banerjee, S.; Ghosh, P. K.; Maiti, T. K., Computational-Based Structural, Functional and Phylogenetic Analysis of *Enterobacter* Phytases. *3 Biotech.* **2018**, 8 (6), 262. DOI: 10.1007/s13205-018-1287-y.
- [38] Pramanik, K.; Saren, S.; Mitra, S.; Ghosh, P. K.; Maiti, T. K., Computational Elucidation of Phylogenetic, Structural and Functional Characteristics of *Pseudomonas* Lipases. *Comput. Biol. Chem.*, **2018**, 74, 190-200. DOI: 10.1016/j.compbiolchem.2018.03.018.
- [39] Geourjon, C.; Deléage, G., SOPMA: Significant Improvements in Protein Secondary Structure Prediction by Consensus Prediction from Multiple Alignments. *Comput. Appl. Biosci.*, **1995**, 11 (6), 681-684. DOI: 10.1093/bioinformatics/11.6.681.
- [40] Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J., Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat Protoc.* **2015**, 10 (6), 845-858. DOI: 10.1038/nprot.2015.053.
- [41] Tamboli, A. S.; Waghmare, P. R.; Khandare, R. V.; Govindwar, S. P., Comparative Analyses of Enzymatic Activity, Structural Study and Docking of Fungal Cellulases. *Gene Rep.*, **2017**, 9, 54-60. DOI: 10.1016/j.genrep.2017.08.008.
- [42] Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.*, **2018**, 46, W296-W303. DOI: 10.1093/nar/gky427.
- [43] Verma, A.; Singh, V. K.; Gaur, S., Computational Based Functional Analysis of *Bacillus* Phytases. *Comput. Biol. Chem.*, **2016**, 60, 53-58. DOI: 10.1016/j.compbiolchem.2015.11.001.
- [44] Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K. P.; Kuhn, M.; Bork, P.; Jensen, L. J.; von Mering, C., STRING V10: Protein-Protein Interaction Networks, Integrated over the Tree of Life. *Nucleic Acids Res.*, **2015**, 43, D447-452. DOI: 10.1093/nar/gku1003.
- [45] Hirokawa, T.; Boon-Chieng, S.; Mitaku, S., SOSUI: Classification and Secondary Structure Prediction System for Membrane Proteins. *Bioinformatics*, **1998**, 14, 378-

379. DOI: 10.1093/bioinformatics/14.4.378.
- [46] Clark, W. T.; Radivojac, P., Analysis of Protein Function and Its Prediction from Amino Acid Sequence. *Proteins*, **2011**, *79* (7), 2086-2096. DOI: 10.1002/prot.23029.
- [47] Nei, M.; Kumar, S., *Molecular Evolution and Phylogenetics*; Oxford University Press New York, **2000**, pp. 1-329.
- [48] Hasegawa, M.; Kishino, H.; Yano, T., Dating of the Human-Ape Splitting by a Molecular Clock of Mitochondrial DNA. *J. Mol. Evol.*, **1985**, *22* (2), 160-174. DOI: 10.1007/BF02101694.
- [49] Rani, S.; Pooja, K., Elucidation of Structural and Functional Characteristics of Collagenase from *Pseudomonas Aeruginosa*. *Process Biochem.* **2018**, *64*, 116-123. DOI: 10.1016/j.procbio.2017.09.029.
- [50] Ramya, L. N.; Pulicherla, K. K., Studies on Deimmunization of Antileukaemic L-Asparaginase to Have Reduced Clinical Immunogenicity- An *in-Silico* Approach. *Pathol. Oncol. Res.*, **2015**, *21* (4), 909-920. DOI: 10.1007/s12253-015-9912-0.
- [51] Ikai, A. Thermostability and Aliphatic Index of Globular Proteins. *J. Biochem.*, **1980**, *88* (6), 1895-1898.
- [52] Ji, Y. -Y.; Li, Y. -Q., The Role of Secondary Structure in Protein Structure Selection. *Eur. Phys. J. E. Soft. Matter.* **2010**, *32* (1), 103-107. DOI: 10.1140/epje/i2010-10591-5.
- [53] Pramanik, K.; Pal, P.; Soren, T.; Mitra, S.; Ghosh, P. K.; Sarkar, A.; Maiti, T. K., *In Silico* Structural, Functional and Phylogenetic Analysis of *Klebsiella* Phytases. *J. Plant Biochem. Biotechnol.*, **2018**, *27*, 362-372. DOI: 10.1007/s13562-018-0445-y.
- [54] Beedkar, S. D.; Khobragade, C. N.; Bodade, R. G.; Vinchurkar, A. S., Comparative Structural Modeling and Docking Studies of Uricase: Possible Implication in Enzyme Supplementation Therapy for Hyperuricemic Disorders. *Comput. Biol.Med.* **2012**, *42* (6), 657-666. DOI: 10.1016/j.compbio.2012.03.001.