ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

In Silico Structural and Functional Insight into the Binding Interactions of the Modeled Structure of Watermelon Urease with Urea

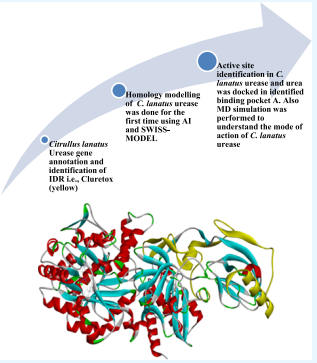
Prince Kumar, Arpan Kayastha, Vinay Kumar Singh, and Arvind M. Kayastha*

Cite This: *ACS Omega* 2024, 9, 2272−2285

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Urease (EC 3.5.1.5) is an amidohydrolase. This nickel-dependent metalloenzyme converts urea into NH_3 and CO_2. Despite their vital role in plants, the structure and function of watermelon (*Citrullus lanatus*) urease are unknown. We used third- and fourth-generation gene prediction algorithms to annotate the *C. lanatus* urease sequence in this investigation. The solved urease structure from *Canavalia ensiformis* (PDB ID: 4GY7) was utilized as a template model to identify the target 3-D model structure of the unknown *C. lanatus* urease for the first time. Cluretox, the *C. lanatus* urease intrinsic disordered area identical to Jaburetox, was also found. The *C. lanatus* urease structure was docked with urea to study atom interaction, amino acid interactions, and binding analyses in the urease−urea complex at 3.5 Å. This study found that amino acids His^{517}, Gly^{548}, Asp^{631}, Ala^{634}, Thr^{569}, His^{543}, Met^{635}, His^{407}, His^{490}, and Ala^{438} of *C. lanatus* urease bind urea. To study the molecular basis and mode of action of *C. lanatus* urease, molecular dynamics simulation was performed and RMSD, RMSF, Rg, SAS, and H-bond analyses were done. The calculated binding free energy (ΔG) for the urea−urease complex at 100 ns using the MM/PBSA method is −7.61 kJ/mol. Understanding its catalytic principles helps scientists construct more efficient enzymes, tailor fertilization to boost agricultural output, and create sustainable waste management solutions.



1. INTRODUCTION

Urea catabolized into NH_3 and CO_2 by urease (EC: 3.5.1.5), a nickel-dependent metalloenzyme.[1] Sumner crystallized jack bean urease, the first enzyme, and won the 1946 Nobel Prize in chemistry. He identified urease as a protein from its crystallized form.[2] In 1995, Hausinger and Karplus found the first bacterial crystal structure.[3] In 2010, Balasubramanian and Ponnuraj solved the first plant urease crystal structure from Jack bean,[4] a nickel-using hydrolase with a bi-nickel center per active site.[5] Basically, urease is a protein found in plants, bacteria, fungi, and invertebrates. It decomposes urea to provide nitrogen for growth.[6] Jack bean and pigeonpea ureases are the most studied plant ureases.[7] Pigeonpea urease was purified, characterized, immobilized, and a prototype urea biosensor was built.[8−13] Urease from soybeans provides the best plant genetic information.[14,15] Plant and fungal ureases are homo-oligomeric proteins with identical subunits, unlike bacterial ones.[16] Plant and bacterial urease share a catalytic mechanism, sequence similarity, and 3-D structure.[17] Comparing bacterial and plant urease catalytic areas showed significant amino acid residue conservation.[17] The catalytic site of Jack bean urease, like bacterial urease, has a bi-nickel core with nickel ions Ni1 and Ni2 separated by 3.7 Å.[17] The structural level of the urease process was studied by crystallizing JBU and PPU at a resolution of 2.05 Å.[4,18] Fabaceae and Cucurbitaceae have high levels of urease activity.[19]

In addition, seed-specific ureases from soybeans and other legumes have been claimed to be involved in plant defense rather than digestion.[20] In addition to the effects of urease, urease or urease-like proteins have been shown to have toxic effects on a variety of fungi and insects.[21,22] The investigation of urease neurotoxicity in rodents and insects led to the discovery of their nonenzymatic characteristics. Both proteins exhibit insecticidal and antifungal effects apart from their ureolytic activity because they have structural similarities with the main urease of the seed.[23−25] It has also been demonstrated that the urease from pigeonpea (*Cajanus cajan*) and soybean (*Glycine max*) exhibited insecticidal[18,22,26] and antifungal effects.[18,27,28]

The long-standing belief that biological function is a characteristic of a particular structure has been challenged by the discovery of intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs), which are unable to establish a stable tridimensional structure. These proteins are present in all species, especially eukaryotes, despite not having an organized structure and exhibiting essential biological functions.[29,30] IDPs and IDRs vary from structured proteins

Received: August 14, 2023
Revised: November 30, 2023
Accepted: December 1, 2023
Published: December 30, 2023

and domains, exhibiting substantial structural plasticity and conformational flexibility. Jaburetox-2Ec, a recombinant peptide containing 93 amino acids (~11 kDa) and corresponding to pepcanatox, was heterologously produced in *Escherichia coli* using the appropriate JBU isoform JBURE-II.[31] Later, a peptide known as Jaburetox was developed that lacked the V5 epitope of Jaburetox-2Ec but had the same urease-derived sequence and 6His tail.[32] The fact that both Jaburetox-2Ec and Jaburetox peptides have the same insecticidal activity indicates that V5 epitopes are not involved in their insect toxicity.[33]

According to the study on the interaction between urease and urea in the model plant *Arabidopsis thaliana*, urease residues Ala^{399}, Ile^{675}, Thr^{398}, and Thr^{679} play a vital role in the interaction with urea substrate.[33] A recent analysis of the molecular structure of two important legume species, soybean (*G. max*) and barrel medic (*Medicago truncatula*) of the Fabaceae family, showed that Glu, Thr, His, and Gly are commonly found as residue interactions in urease−urea-binding complexes while Glu was present in all docked structures.[34]

Despite the significance of urease, there is little structural information available, and the protein model database (PMDB) does not yet have the modeled structure of *C. lanatus* urease. In the present study, we predicted the *C. lanatus* urease sequence along with the Jaburetox-like intrinsically disordered regions in *C. lanatus* urease. We used the homology model to predict the 3D model structure of *C. lanatus* urease and molecular dynamics (MD) simulation to mimic the predicted structure. In addition, a ligand protein-based molecular binding technique was used to predict the interaction between urease and its substrate, urea. Ongoing work represents a breakthrough in our understanding of the activation mechanism of urease and the urea complex in the *Cucurbit* family. The significance of *C. lanatus* urease extends to its insecticidal and fungicidal properties. The present work has revealed significant breakthroughs in the identification of Cluretox in *C. lanatus* urease, which is homologous to Jaburetox and Soyuretox. It has been shown that Cluretox can serve as a promising toxin for the development of transgenic plants with enhanced resistance to insect herbivory and fungal pathogens. In silico study of watermelon urease provides valuable insights into its catalytic mechanism at a molecular level. By studying watermelon urease through computational simulations, scientists can gain a deeper understanding of its structure and function. Lastly, understanding these mechanisms can aid in the development of innovative waste management strategies, utilizing urease enzymes to break down urea-containing waste products and mitigate their harmful effects on the environment. By uncovering the underlying mechanisms and properties of watermelon urease, scientists can design novel strategies for crop improvement, such as developing genetically modified crops that can efficiently utilize nitrogen from fertilizers.

2. MATERIALS AND METHODS

2.1. In Silico Analysis and Gene Prediction. Gene prediction was based on comparative analysis and gene identification. In this case, the *Arabidopsis thaliana* urease gene (TAIR NO. ATIG67550, Gene Bank Acc.no.NM_105422, Uniprot Id-Q9SR52) was selected as the reference gene model. For gene prediction, the free online server Softberry, an HMM-based ab initio gene structure prediction server hosted by FGENESH,[35] and AUGUSTUS, an HMM-based eukaryotic gene prediction server (using the *A. thaliana* model data set).[36] The *C. lanatus* target urease was identified and obtained using the *C. lanatus* Whole Genome Shotgun (WGS) project of Project Accession AGCB00000000.[37]

The longest gene length prediction was accepted as the gene mode. The target was taken from the NCBI database (http://ncbi.nlm.nih.gov/protein/). Further genes of interest or urease sequences were used for full-length gene prediction based on FGENESH (Find Genes; www.softberry.com) and AUGUSTUS (http://augustus.gobics.de/submission) server. The complete CDS were translated into a protein sequence using the translate tool EXPASY[38] (https://web.expasy.org/translate/), and the protein sequences were saved in FASTA format. Gene structure display server 2.0 was used for gene feature visualization.[39]

2.2. Sequence Retrieval, Functional Domain Identification, Physicochemical Analysis, Secondary Structure Prediction, and Phylogenetic Classification. *A. thaliana* urease sequence (TAIR ID: AT1G67550.1, UNIPROT ID: Q9SR52, GENPEPT ID: NP_176922.1) was retrieved in FASTA format from the TAIR/UNIPROT/NCBI database. The sequence was selected and retrieved on the basis of a homology search done using the similarity search tool BLAST.[40] Further, InterPro (www.ebi.ac.uk/interpro/) server was used to examine the protein domain. Using Protparam (http://web.expasy.org/protparam/), physicochemical properties were analyzed. The secondary structure features were generated using SOPMA server (https://npsa-prabi.ibcp.fr/cgi-bin/npsaautomat.pl?page=/NPSA/npsa sopma.html). For multiple sequence alignment (MSA) of homologous sequences, ClustalW (https://www.genome.jp/tools-bin/clustalw) was utilized. CLC Sequence Viewer was used to see the aligned sequences (https://clc-sequence viewer.software.informer.-com/8.0/). The phylogenetic tree was subsequently constructed using MEGA (https://www.megasoftware.net/) software, and a tree was used to analyze the phylogenetic relationship and protein evolution based on conserved, variable, parsim informative, and singleton sites. In this tree construction, an unweighted pair group method with the arithmetic mean approach (UPGMA) was used to construct a phylogenetic tree from a distance matrix.

2.3. Structure Prediction and Evaluation. To access the best quality model, structure prediction was performed utilizing four servers, including SWISS-MODEL, AlphaFold2, I-TASSER, and RosettaFold. For Homology Modeling, the SWISS-MODEL server was used to predict the 3D model structure of *C. lanatus* urease using crystallographic structure analysis of urease from Jack bean (*Canavalia ensiformis*) Protein Data Bank file (PDB ID: 4GY7). For 3D model construction, the best template (PDB ID: 4GY7) was selected from the SWISS-MODEL template library (SMTL). Maximum query coverage, maximum identity, high score, and low e-value parameters were taken into consideration when choosing the template. SWISS-MODEL server automation was used to forecast the 3D model structure.[41]

Further, the model structure from SWISS-MODEL of the *C. lanatus* urease was used for evaluation of phi (ϕ) and psi (Ψ) torsion angles and the quality of covalent bonds using PROCHECK, SAVES server (http://nihserver.mbi.ucla.edu/SAVES).[42−44] The evaluated model was deposited in the PMDB server (http://srv00.recas.ba.infn.it/PMDB/). SWISS-MODEL was also used to solve the trimeric structure, which was achieved by using the *C. lanatus* urease sequence as a target and PDB ID: 4GY7 of JBU as a template.

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

>*Citrullus lanatus* Urease
MKLTPREVEKLGLHNAGYLAQKRLARGLRLNYAEAVALIASQILEFVRDGNCSVAELMELGRELLGRR
QVLPSVPHLLNYIQVEGTFPDGTKLITVHKPIESEVGNLERALQGSFLPVPSLDKFPLMEDSKVPGEI
ICDDKKIAINVGRKAVILSVKNMGDRPIQVGSHYHFIETNPSLVFDRSKAYGMRLNILAGSATRFEPG
CPKDVTLVAIGGNQVIRGGNCIADGPVDKSKLKEVMEAVHARGFGHLEEKDAREGVTGGEDEIFTTRV
LREEYANHYGPTTRDKVRLGDTDLYAEIEHDFCVYGDECMFGGGKVIRDGMGQACGHPPASSLDTVIT
NAVIIDYTGIFKADIGIKDGFIISLGKAGNPDIMDRVCPNLIIGANTEVISGEGLLVTSGAIDCHVHY
ICPQMVHEAISSGITTLVGGGTGPAEGSCATTCTPSPALMKMILQSTDNLPLNFGFTGKGNTSKPDEI
HEIIRAGAMGLKLHEDWGCTPAAIDSCLAVAEKYDIQVNIHTDTLNESGFVEDTIAAFKGRTIHTYHT
EGAGGGHAPDIIKVCSVKNVLPSSTNPTRPFTTNTIDEHLDMLMVCHHLDPNIEEDVAFAESRIRKET
IAAEDILHDMGAISMISSDSQAMGRIGEVISRTWQTAHKMKLQREPLDSTKPDDDNLRIKRYVSKYTI
NPAIANGISEYVGSVEVGKWADLVLWKPSFFGTKPEKVLKGGIIAWANMGDPNASIPTPEPVLMRPMY
GSMGKAASGSSIAFVSKAAFNAGVKDMYGLNKRVVAVSNTRKLTKLDMKWNDALPNIQVDPDHYDVKV
DGDVLTCLPATTLPLSRDYFLF

Figure 1. Full-length predicted *C. lanatus* urease sequence in FASTA format.

2.4. Intrinsic Disorder Propensity of Cluretox in *Citrullus lanatus* and Physicochemical Properties Analysis. To identify the disorder region in *C. lanatus* urease, the *C. ensiformis* urease Jaburetox sequence was used as a reference for comparison and identification of *C. lanatus* Cluretox region.[45] The disorder amino acid region Cluretox was examined using the tool PONDR (Predictor of Natural Disordered Regions; http://www.pondr.com/). PONDR is the best tool to understand the links between sequence anomalies and susceptibility to intrinsic disorder. The intrinsic disorder tendency of Cluretox was assessed using three different PONDR family disorder predictors: (a) PONDR VL-XT using multiple composition probabilities and hydrophobic amino acids measurements as input features of the artificial neural network for prediction;[46] although it is no longer the most accurate predictor, it is very sensitive to local constituents and can therefore be used to identify potential molecular interaction motifs; (b) PONDR VSL2 is used to accurately evaluate short and long disordered regions;[47] (c) PONDR VL3 is used to locate persistent regions of disorder.[48] The amino acid sequences were entered into PONDAR (http://www.pondr.com/) to produce the visuals. The Protparam server was used for physicochemical properties calculations. The secondary structure features were generated using SOPMA server.[49]

2.5. Active Site Identification. POCASA server was used for active binding sites by spotting pockets and cavities. POCASA (POcket-CAvity Search Application) implements the Roll method, which may predict binding sites with known three-dimensional (3D) structures.

2.6. Ligand Selection. The urea (ligand) was taken as the ligand for retrieval from the PubChem database (PubChem CID 1176). The PubChem database (https://pubchem.ncbi.nlm.nih.gov/) was searched for the chemical structure of urea (PubChem CID 1176). Using UCSF CHIMERA 1.10 (https://www.cgl.ucsf.edu/chimera/), the 3D structure of urea was then energy-minimized and exported to the Protein Data Bank (PDB) file format.

2.7. Molecular Docking. An in silico approach of molecular docking was used to investigate the molecular interaction between *C. lanatus* urease and urea. In this docking investigation, urease from *C. lanatus* was used. The grid-based docking of urea was performed with the *C. lanatus* urease using BIOVIA DISCOVERY STUDIO 2019 (https://discover.3ds.com/). The compared active sites were identified as POCASA site A, and BIOVIA DISCOVERY STUDIO and a sphere was generated. A receptor grid was generated for the molecular docking study, having boxed dimensions of X = 144.797, Y = 151.274, Z = 122.203, and radius = 15.398 Å.

2.8. Molecular Dynamics Simulation. The binding stability, conformation, and interaction modes between ligand as urea and receptor as urease were assessed using molecular docking and dynamic simulation. Using the GROMACS built into the WebGro server (https://simlab.uams.edu/), the chosen urease and urease−urea complex were subjected to molecular dynamics simulation (protein in water simulation). The steepest descent approach for 5000 steps was used in molecular dynamic simulation. The complex structure was solvated using a simple point charge (SPC) water model in a cubic periodic box of 0.5 nm. The complex system was then kept at a proper salt concentration of 0.15 M by appropriately introducing Na^+ and Cl^- counterions. For the simulation, the NPT ensemble isothermal−isobaric, constant-number-of-particles, constant-pressure, and constant-temperature equilibration was applied to the complex for a simulation time of 100 ns. The simulation result analysis was done and RMSD, RMSF, Rg (structural compactness), SASA (solvent accessible surface area), and H-bonds were calculated.

2.9. Molecular Mechanics Poisson−Boltzmann Surface Area (MMPBSA) Calculation. For finding the binding energy of the urea−urease complex, molecular dynamics simulations were performed using CHARMM36 force fields and water as a solvent for 100 ns in GROMACS. In this work, MD simulations have been carried out to evaluate the binding energy of the urease−urea complex. These simulations were investigated for the complex for 100 ns via GROMACS (Groningen Machine for Chemical Simulations) software. The topology of the protein, as well as the ligand, was created using CHARMM36 force fields. The complex was immersed in a dodecahedron box of simple point charge (SPC) water molecules. The solvated system was neutralized by adding counterions. Energy minimization of the solvated structures was done using a conjugate gradient algorithm until the maximum force reached below 100 kJ/mol/nm. To equilibrate, the system was then subjected to the position-restrained dynamics simulation (NVT and NPT) at 300 K for 100 ps. Finally, this system was subjected to an MD production run for 100 ns at 300 K temperature and 1 bar pressure. The binding free energy of the urea−urease complex was determined using the MM/PBSA method. The GROMACS software program g_mmpbsa was used to measure the binding free energy (ΔG).[50]

3. RESULTS AND DISCUSSION

3.1. In Silico Analysis and Gene Prediction. Based on homology search, it has been concluded that Chromosome 11 stretch of *C. lanatus* cultivar 97103, whole genome shotgun

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

Table 1. Physicochemical Properties and Secondary Structure Features of Selected Plant Ureases[a]

species name	primary sequence features				secondary structure features			
	sequence length	mol mass (kDa)	theoretical pI	amino acid richness (%)	α- helix (%)	β- turn (%)	extended strand (%)	random coil (%)
Citrullus lanatus	838	90.95	5.71	Gly 9.4	29.36	8.83	20.05	41.77
Arabidopsis thaliana (reference)	838	91.02	5.92	Gly 10	30.19	9.07	19.69	41.05
Cucurbita moschata	838	90.13	5.37	Gly 10.3	28.40	9.31	20.53	41.77
Cucurbita pepo	837	90.07	5.30	Gly 10.3	29.03	9.32	20.19	41.46
Momordica charantia	844	91.74	5.75	Gly 9.7	29.74	9.36	20.73	40.17
Cucumis sativus	833	89.88	5.69	Gly 9.8	29.65	9.00	20.29	41.06
Cucurbita maxima	833	89.97	5.53	Gly 9.4	28.45	9.48	20.77	41.30
Cucumis melo	833	89.90	5.84	Gly 10.0	28.69	9.60	20.29	41.42
Carica papaya	837	89.69	5.60	Gly 10.9	29.15	9.80	21.03	40.02
Mangifera indica	835	89.87	5.64	Gly 9.9	28.62	9.34	20.84	41.20
Gossypium hirsutum	837	90.10	5.24	Gly 10.4	29.39	8.96	20.07	41.58
Cannabis sativa	837	90.18	5.58	Gly 10.4	28.67	9.32	20.19	41.82
Citrus sinensis	837	89.74	6.15	Gly 10.0	27.96	9.32	20.67	42.05
Cajanus cajan	837	90.07	5.43	Gly 10.0	29.03	9.44	20.43	41.10
Glycine max	837	90.71	5.78	Gly 9.6	27.72	9.32	20.91	42.05
Cicer arietinum	838	90.15	5.36	Gly 9.9	29.95	9.31	20.29	40.45
Vigna unguiculata	837	90.66	5.69	Gly 9.7	28.67	9.32	20.55	41.46

[a]The 17 ureases used, including *C. lanatus* urease protein, had similar physicochemical properties, including 833−844 amino acid residues, 89.69−91.74 kDa molecular mass, having an acidic pI (5.24−6.15) and glycine-rich. Secondary structure features reveal a higher prevalence of α-helix conformation compared to β-turn conformation

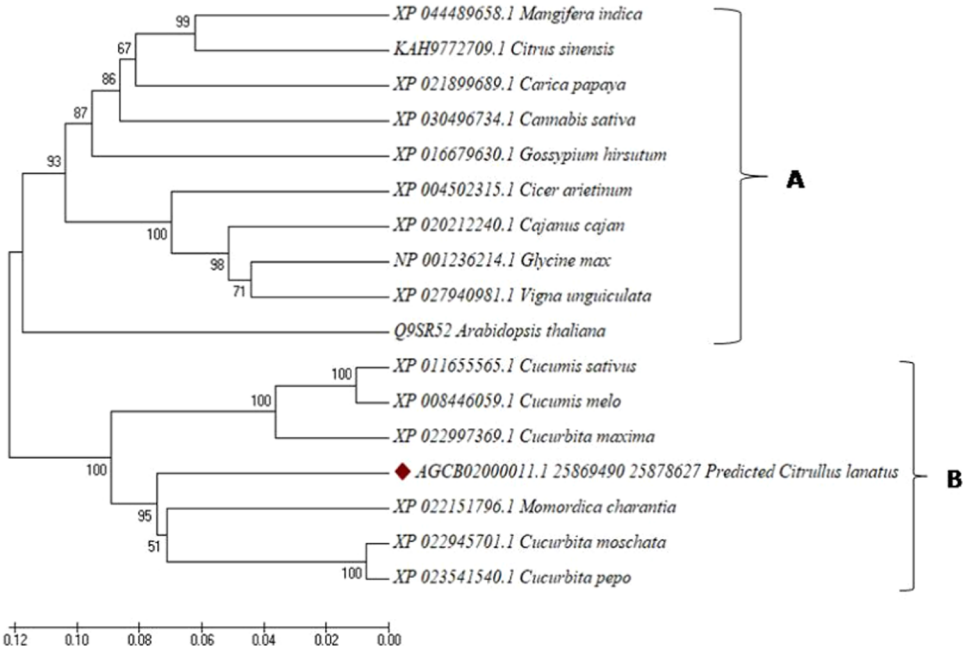


Figure 2. Phylogenetic tree prediction using MEGA based on ureases from *C. moschata* (Acc. No. XP_022945701.1), *C. pepo* (Acc. No. XP_023541540.1), *M. charantia* (Acc. No. XP_022151796.1), *C. sativus* (Acc. No. XP_011655565.1), *C. maxima* (Acc. No. XP_022997369.1), *C. melo* (Acc. No. XP_008446059.1), *C. papaya* (Acc. No. XP_021899689.1), *M. indica* (Acc. No. XP_044489658.1), *G. hirsutum* (Acc. No. XP_016679630.1), *C. sativa* (Acc. No. XP_030496734.1), *C. sinensis* (Acc. No. KAH9772709.1), *C. cajan* (Acc. No. XP_020212240.1), *G. max* (Acc. No. NP_001236214.1), *Cicer arietinum* (Acc. No. XP_004502315.1), *V. unguiculata* (Acc. No. XP_027940981.1), *A. thaliana* (Q9SR52), *C. lanatus* (Acc. No. AGCB02000011.1, from 25869490 to 25878627) organisms.

sequence (Accession No. AGCB02000011.1, from 25869490 to 25878627) having full genomic length of 9138nt from start codon to end codon was predicted and analyzed using ab initio gene prediction tool FGENESH and AUGUSTUS, and the results were exported in FASTA format (Figure S1). Based on retrieved genomic sequences, full-length CDS was predicted with 2517nt (Figure S2). The predicted *C. lanatus* urease FASTA protein sequence with 838 aa was given below (Figure 1). The exon-intron boundaries were predicted using Gene structure display server 2.0. Full-length genomic sequence composed of 18 exons and 17 introns (Figure S3). It can also be elucidated that *C. lanatus* urease is homologous to *A. thaliana* urease with the same number of amino acid residues and gene structure.[33]

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

3.2. Sequence Retrieval, Functional Domain Identification, Physicochemical Analysis, Secondary Structure Prediction, and Phylogenetic Classification. After the successful identification of urease from *C. lanatus* genome, the predicted complete protein sequences of urease were used for homology search using the similarity search tool BLAST. Homologous sequences of *Cucurbita moschata* (Acc. No. XP_022945701.1), *Cucurbita pepo* (Acc. No. XP_023541540.1), *Momordica charantia* (Acc. No. XP_022151796.1), *Cucumis sativus* (Acc. No. XP_011655565.1), *Cucurbita maxima* (Acc. No. XP_022997369.1), *Cucumis melo* (Acc. No. XP_008446059.1), *Carica papaya* (Acc. No. XP_021899689.1), *Mangifera indica* (Acc. No. XP_044489658.1), *Gossypium hirsutum* (Acc. No. XP_016679630.1), *Cannabis sativa* (Acc. No. XP_030496734.1), *Citrus sinensis* (Acc. No. KAH9772709.1), *Cajanus cajan* (Acc. No. XP_020212240.1), *Glycine max* (Acc. No. NP_001236214.1), *Cicer arietinum* (Acc. No. XP_004502315.1), and *Vigna unguiculata* (Acc.No.XP_027940981.1) organisms were retrieved in FASTA format, and multiple sequence alignment was performed (Figure S4).

A total of 17 ureases from various plant species including *C. lanatus* urease protein were used and all urease sequences shared similar physicochemical characteristics, such as 833−844 amino acid residues, 89.69−91.74 kDa molecular mass, primarily acidic (5.24−6.15) character and having glycine richness (Table 1). The four multidomains amidohydro_rel (IPR006680), urease_gamma (IPR002026), urease_beta (IPR002019), and urease_alpha_N_dom (IPR011612) were used to characterize all protein sequences. γ, β, α, and amidohydro_rel domains roughly mapped to amino acids ranging from 1−102, 131−232, 272−390, and 394−724, respectively, in the *C. lanatus* sequences (Figure S5). The secondary structural characteristics of *C. lanatus* ureases, such as extended strand, β-turn, random coil, active site residues, and metal cofactor binding site, were also used to demonstrate the conserved protein architecture in plants, which primarily comprises α-helix and random coil structures (Figure S6)

The phylogenetic tree (Figure 2) was inferred using the UPGMA method of selected 16 homologous protein sequences with *C. lanatus* urease. Based on the phylogeny result, two main functional groups (A and B) were identified by alignment of the phylogenetic tree. Plant urease from target genome *C. lanatus* was successfully classified with *C. sativus*, *C. melo*, *C. maxima*, *M. charantia*, *C. moschata*, and *C. pepo* in group B cluster while the remaining were in cluster A.

Based on sequence length, molecular mass (kDa), theoretical pI, amino acid richness (%), α-helix (%), β-turn (%), extended strand (%), and random coil (%), both the primary and secondary structures analysis of identified and selected 16 ureases from various plant species revealed conserved secondary protein architecture and similar physicochemical characteristics. This indicates that the discovered *C. lanatus* urease gene (Acc. No. AGCB02000011.1, from 25869490 to 25878627) has been successfully conserved throughout the course of the phylogenetic expansion of plants and the evolutionary process. All plant ureases, including *C. lanatus*, had the four domain structures: urease Υ, urease β, urease_alpha_N_dom, and amidohydro _re1.

3.3. Structure Prediction and Evaluation. Structure prediction was made using four servers, i.e., SWISS-MODEL, AlphaFold2, I-TASSER, and RosettaFold, to access the best quality model. The details of stereochemical qualities generated using the SAVES server are given in Table 2. Further comparison was made between the I-TASSER, RosettaFold, AlphaFold2 generated model, and SWISS-MODEL based predicted model. The model derived with the best PROCHECK result was used for further analysis, which was generated using the SWISS-MODEL server.

Table 2. Ramachandran Plot Based on Comparative Analysis of Several Modeled Structures Using Different Servers[a]

	Ramachandran plot analysis			
server	most favored regions (%)	additionally allowed regions (%)	generously allowed regions (%)	disallowed regions (%)
SWISS-MODEL	91.6	7.9	0.6	0.0
AlphaFold2	90.3	9.1	0.6	0.0
RosettaFold	89.3	10.5	0.0	0.1
I-TASSER	83.7	14.0	1.8	0.4

[a]The SWISS-MODEL algorithm produced the best model, as confirmed through analysis using the Ramachandran plot. The investigation revealed that 91.6% of the residues were located in the core regions, whereas AlphaFold2 has 90.3%, RosettaFold 89.3%, and I-TASSER 83.7% in most favored regions

The SWISS-MODEL server was used to build a 3D model of the *C. lanatus* urease using *C. ensiformis* (PDB ID: 4GY7) as the template (Figure 3). The other verification servers, PROCHECK, Verify3D, ERRAT, and PROVE, were used for verification to assess the quality of the predicted model.

In PROCHECK analysis of the predicted model (SWISS-MODEL), the Ramachandran plot showed that there were 91.6% residues in core regions (most favored regions), 7.9% in allowed (additionally allowed regions), 0.6% in generously allowed regions, and 0.0% in disallowed regions for *C. lanatus* urease model (Figure 4A). For a good model structure, obtained at high resolution, one would expect this percentage to be over 90% in the most favored region [A, B, L]. In the case of *C. lanatus* urease model, 91.6% of residues were observed in the most favored region and no residue was in the disallowed region. The result indicated that the stereochemical quality was good for the predicted model.

The predicted model quality was observed based on other verification servers. According to the Verify3D analysis, 85.08% of the residues had an averaged 3D-1D score ≥0.2 (Figure 4B); for a good quality model, there should be at least 80% of residues scoring ≥0.2 in the 3D-1D profile. This result also validated that the predicted model is good. The overall quality factor of the *C. lanatus* model was 86.3855 from the ERRAT server (Figure 4C). Thus, for further analysis, the quality and dependability of the model built were confirmed across all three servers. The PROVE analysis resulted in 0% buried outlier protein atoms.

The predicted 3D model of the *C. lanatus* plant urease using the SWISS-MODEL server with the best quality, the validation server also proved that the predicted model was in good quantity and quality. The correctness and dependability of the predicted model were examined using three different servers, which utilize different internal evaluation scores. All three selected servers could qualify the structure of the built-in model for additional examination. The modeled structure (SWISS-MODEL) of *C. lanatus* urease was deposited with PMDB ID no. PM0084348. Earlier reports suggested that JBU is composed of a trimer or

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

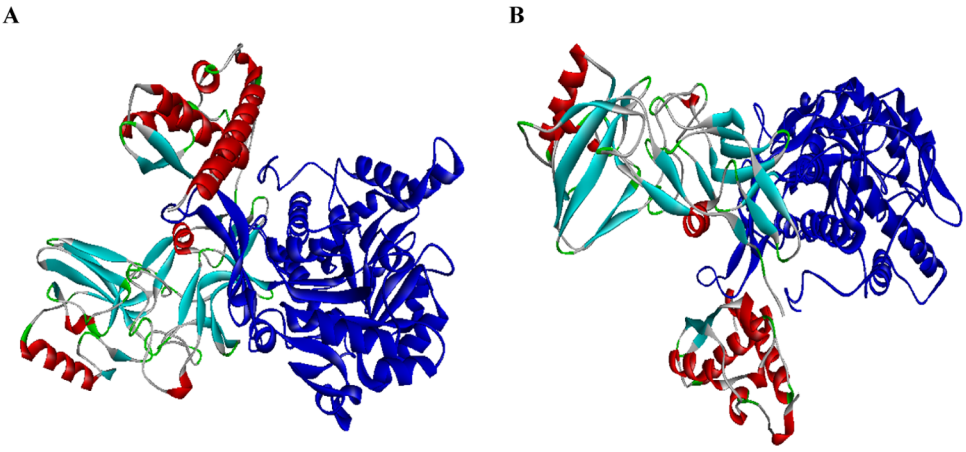


Figure 3. Template A. *C. ensiformis* (PDB ID: 4GY7) and Target B. *C. lanatus* urease structure in a ribbon form. The presence of a blue ribbon in both urease enzymes indicates the presence of a domain.

hexamer comprising six identical 90 kDa subunits,[5] each of which contains two nickel ions. As depicted in Figure 5, *C. lanatus* urease produces a trimer similar to JBU, each containing two nickel ions.[6] One notable similarity between the modeled structure of *C. lanatus* urease and the preexisting crystal structures of JBU is that in the case of *C. lanatus*, the catalytic site of urease has a striking resemblance to that of JBU since both enzymes include a bi-nickel center. This center consists of two nickel ions, Ni1 and Ni2, which are separated by a distance of 3.7 Å in JBU and 3.61 Å in *C. lanatus* urease.[4] It is important to highlight that the modeled watermelon urease exhibits the same amino acid residues as those discovered in the crystal structures of JBU.[4] The active site residues are mentioned in Section 3.5.

3.4. Intrinsic Disorder Propensity of Cluretox in *Citrullus lanatus* and Physicochemical Properties Analysis. The identified intrinsic disorder region of *C. lanatus* urease is termed Cluretox (Figure 6). The identified Cluretox from *C. lanatus* urease has physicochemical characteristics, with 91 amino acid residues, 10.19 kDa molecular mass, and primarily acidic (pI 4.88) character. The Cluretox sequence from *C. lanatus* urease sequence mapped to amino acids ranging from 230 to 321. The conserved protein architecture in plants, which mostly consists of α-helix (28.57%) and random coil (46.15%) structures, was also shown by the secondary structural properties of urease, including extended strand (18.68%), β-turn (6.59%). A family of PONDR predictors used to analyze the per-residue disorder propensity of Jaburetox, Soyuretox,[51] and Cluretox found that this protein contains extended disordered patches (PONDR scores over 0.5), particularly in its N-terminal region (Figures 7 and S7).

The physicochemical characterization and intrinsically disordered regions of Cluretox, a polypeptide synthesized from a watermelon ubiquitous urease and homologous to Jaburetox and Soyuretox (Figure S8), a similar recombinant polypeptide derived from a *C. ensiformis* and *G. max* urease, respectively were presented in this work.[31] Cluretox is fundamentally disordered, as shown by bioinformatics software. It can also be anticipated that Cluretox, which is homologous of Jaburetox and Soyuretox, could be used as a risk-free alternative for developing resistance in transgenic plants against insect herbivory and/or fungal infection. We can also characterize Cluretox structurally, and its impact on insects and fungi may be discussed in the future.

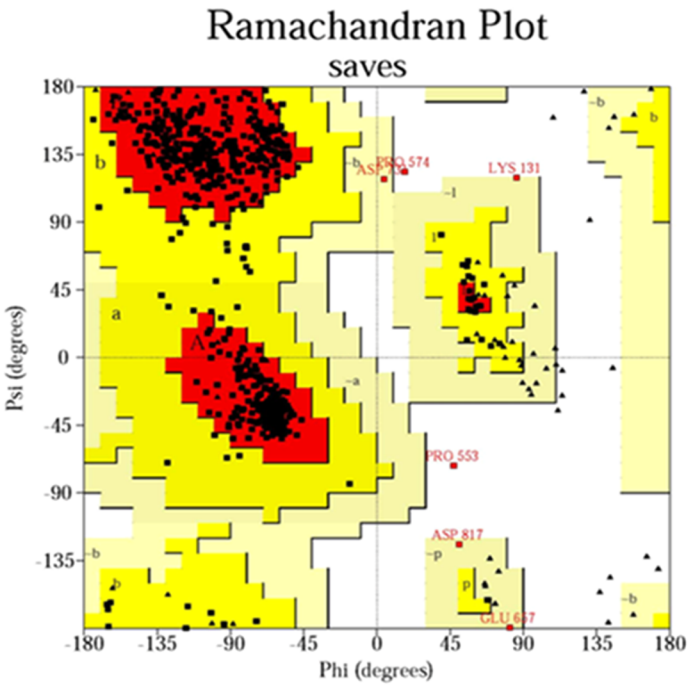
3.5. Active Site Identification. According to binding site analysis, using the predicted and verified structure of *C. lanatus* urease resulted in 5 binding pockets. Out of 5 binding pockets, pocket A was involved in interaction with urea and was already reported as the binding site for jack bean. For reference, Jack bean (*C. ensiformis*) 4GY7 and 3LA4 PDB Ids were taken and prominent active site residues were His^{407}, His^{409}, His^{492}, His^{519}, His^{545}, Ala^{440}, Ala^{636}, Gly^{550}, Asp^{633}, and Met^{637} (Figure S9). Similarly, *C. lanatus* urease prominent and essential residues Thr^{569}, His^{543}, Met^{635}, His^{407}, His^{490}, Ala^{438}, Gly^{548}, Ala^{634}, Asp^{631}, and His^{517} were identified as major active binding residues.[4] The comparative analysis using BLAST, the prominent active site residues in Jack bean as a reference with *C. lanatus* urease, was done successfully for sphere generation and docking calculation (Figure S10). Figure 8 illustrates the specific residues inside the protein that were shown to interact with the metal cofactor, i.e., Ni. In order to predict more accurate scoring of the selected urea pose, a receptor grid was generated for the chosen *C. lanatus* urease based on reported binding site residues.[52−56]

3.6. Molecular Docking. The probable binding sites and affinities in the urease−urea complexes of *C. lanatus* were examined using docking complexes. BIOVIA Discovery Studio 2019 was used to visualize the best-scored complex in order to display the atomic contact maps between urease and urea. The findings of the discovery studio demonstrated that the amino acids in urease His^{517}, Gly^{548}, Asp^{631}, and Ala^{634} bound (contacted) with urea through H-bonds and interacting atoms of urea and respective amino acids are shown in Table 3. Residues Thr^{569}, His^{543}, Met^{635}, His^{407}, His^{490}, and Ala^{438} were predicted on the surface of interaction sites within 3.5 Å with van der Waals energy (Figure 9).

To identify the possible binding sites in *C. lanatus* and examine the interacting amino acids and atom distances in the urea−urease docking complex, Jack bean (*C. ensiformis*) urease models (PDB ID: 4GY7) were used. The residues His^{517}, Gly^{548}, Asp^{631}, Ala^{634}, Thr^{569}, His^{543}, Met^{635}, His^{407}, His^{490}, and Ala^{438} were found to interact with urea. His^{519}, His^{545}, and Lys^{490} residues were reported as active sites of a complex of Nickel 1, while His^{407}, His^{409}, Asp^{633}, and Lys^{490} residues were liganded to Nickel 2 in *C. ensiformis* urease crystal structure.[4] In the case of *C. lanatus* urease and urea docking, His^{407}, His^{490}, His^{517}, and His^{543} were docked. The results suggested a potential function in nickel binding; however, further research in experiments is

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

A. PROCHECK



Plot statistics

Residues in most favoured regions [A,B,L]	652	91.6%
Residues in additional allowed regions [a,b,l,p]	56	7.9%
Residues in generously allowed regions [~a,~b,~l,~p]	4	0.6%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	712	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	79	
Number of proline residues	45	
Total number of residues	838	

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

B. VERIFY 3D

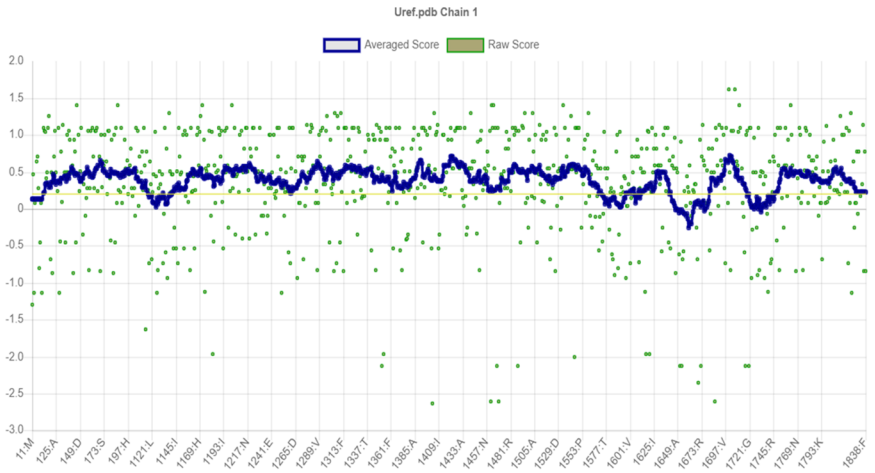


Figure 4. continued

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

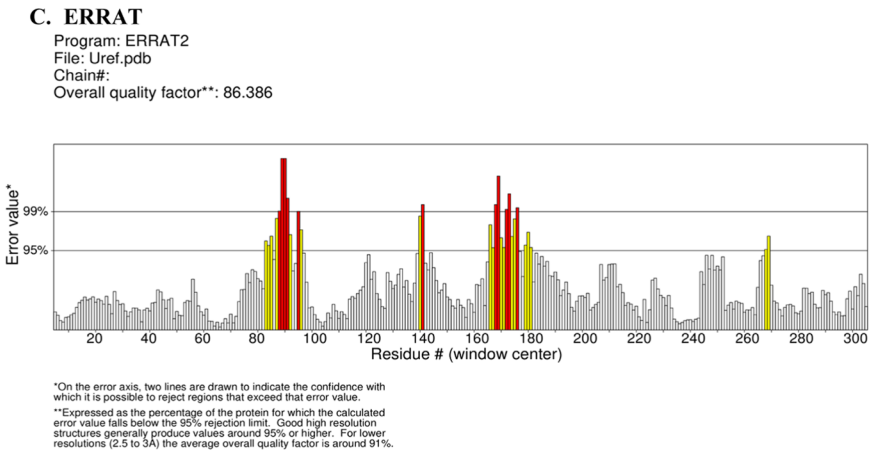


Figure 4. Structural verification details were generated for the SWISS-MODEL predicted model using SAVES SERVER (A) PROCHECK, (B) Verify3D, and (C) ERRAT.

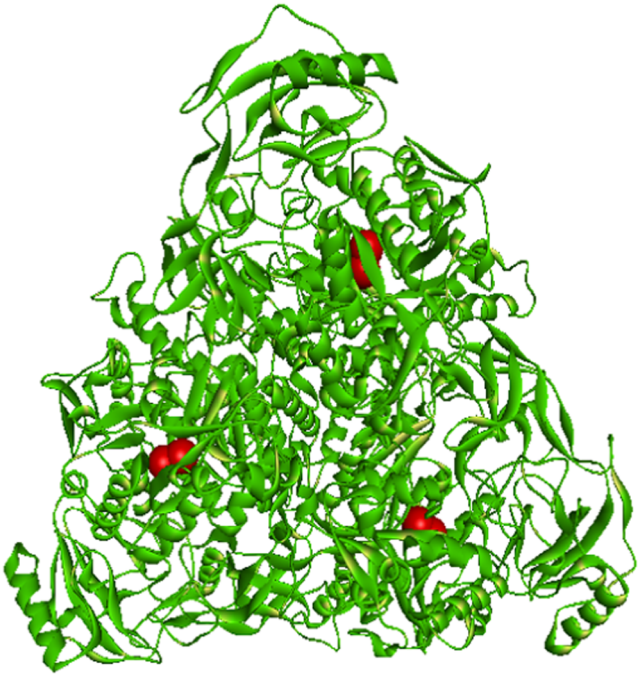


Figure 5. Ribbon scheme of the *C. lanatus* urease, three monomers are associated in a triangular fashion, generating a planar surface on the face of the triangle using SWISS-MODEL. The nickel atoms in the active site are shown as red spheres.

needed to validate this. Additionally, the *C. lanatus* urea−urease complex showed the presence of Ala^{438} and Ala^{634} residues that were in the docked structure. Balasubramanian and Ponnuraj[4] already observed the presence of Ala residue in plant urease active site architecture.

3.7. Molecular Dynamics Simulation. Although protein−ligand docking is extensively employed and possesses valuable uses, it primarily provides a static representation resembling an image of the ligand binding pose within the active area of the receptor. To model the temporal behavior of atoms within a system, it is essential to employ the method of molecular dynamics (MD), which involves the computational integration of Newton's equations of motion.[57] The receptor−ligand complex that scored top in the docking investigation underwent a 100 ns MD simulation, and the resulting data were then analyzed. The stability and fluctuation of the complex were

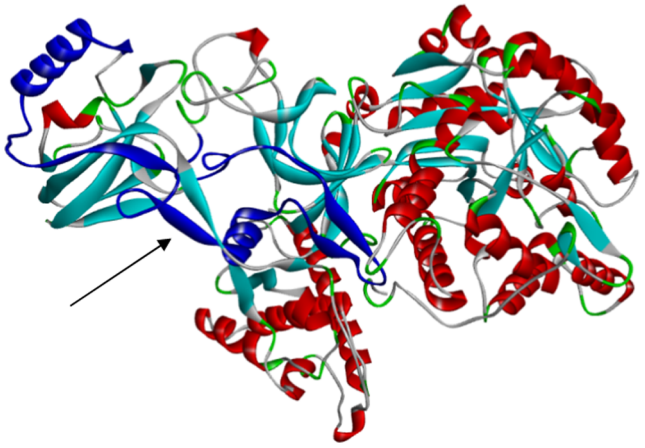


Figure 6. *C. lanatus* urease and entomotoxic peptide, i.e., Cluretox represented structurally. The arrow shows the location of the peptide (in blue) in the *C. lanatus* urease structure.

assessed by the utilization of MD trajectory analysis, using measures such as radius of gyration (RG), root mean square deviation (RMSD), root mean square fluctuation (RMSF), H-bonding, and solvent accessible surface area (SASA) of the receptor atoms. The RMSD is an essential measure for assessing the equilibration of MD trajectories and verifying the stability of complex systems during the simulation procedure. To examine alterations in structural conformation, a plot was generated to display the RMSD of the protein backbone atoms over time. The RMSD of the urease−urea complex exhibited early fluctuations within the first 70 ns, ranging from 0.2 to 0.55 nm. The stable configuration was achieved within the time interval of 75−100 ns, with no significant deviations observed in the collected data. The results of the simulation indicate that the protein went through minor structural alterations,[58] as depicted in Figure 10A.

RMSF is another significant characteristic to consider when evaluating the stability and adaptability of complex structures through simulation.[59] The behavior of amino acid residues of the target protein was analyzed using RMSF after their contact with a ligand.[60,61] The RMSF values of carbon atoms in the protein were computed and presented with respect to the individual residues. During the course of the simulation, minimal diversity was noted in the amino acid residues of the

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

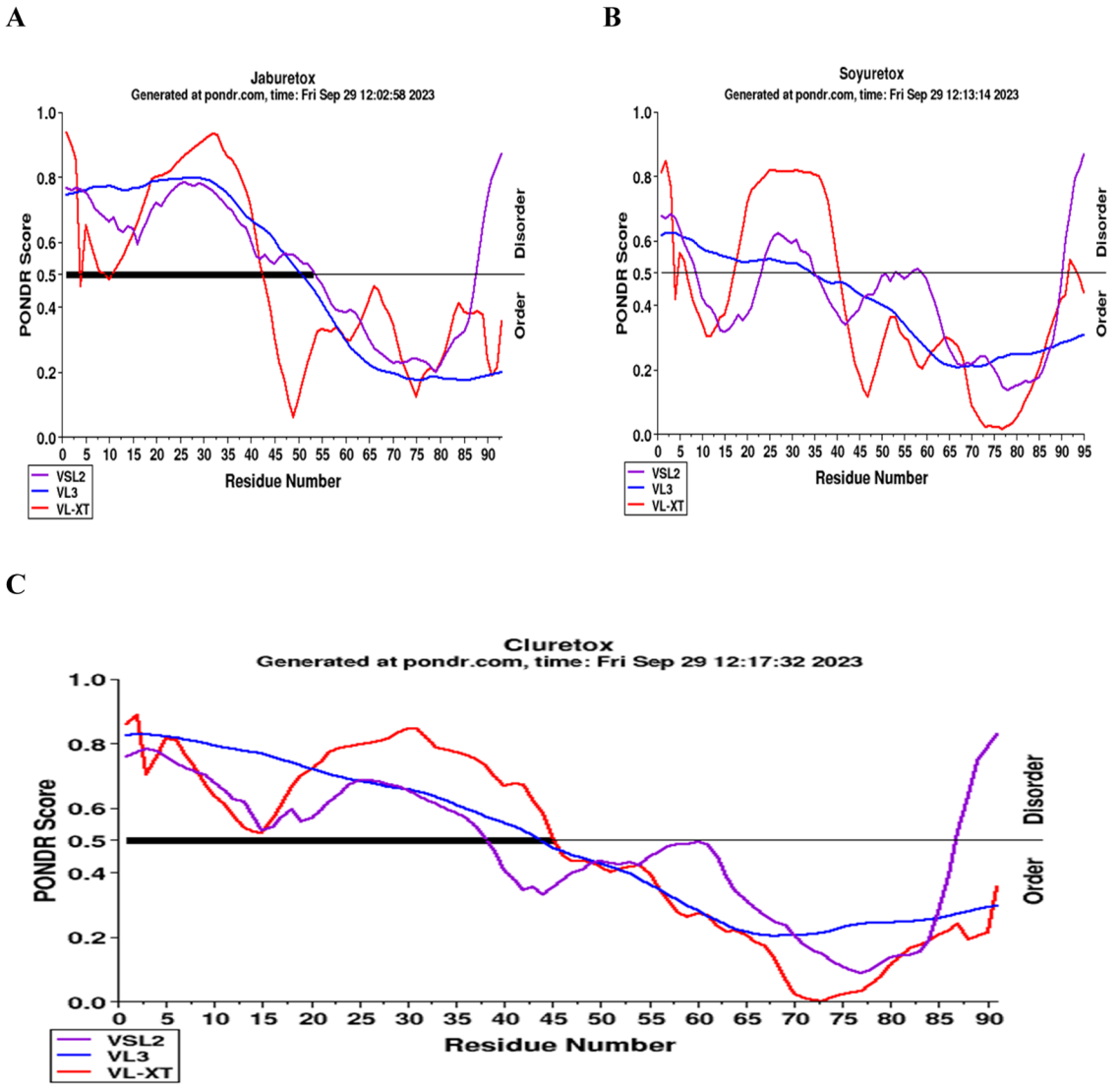


Figure 7. Evaluation disorder propensity of Jaburetox (A), Soyuretox (B), and Cluretox (C). An algorithm for predicting disorders the amino acid sequences of Jaburetox, Soyuretox, and Cluretox were compared using three PONDR family members PONDR VL-XT (red line), PONDR VSL2 (violet line), and PONDR VL3 (blue line) to assess the per-residue intrinsic disorder likelihood of Jaburetox, Soyuretox and Cluretox, sections with scores higher than 0.5 are considered to be disordered regions.

investigated complex.[58] The amino acids, namely, His^{517}, Gly^{548}, Asp^{631}, and Ala^{634}, which exhibited the lowest fluctuation values during the docking process and interacted with urea, were identified as the urease amino acids. The findings of this study indicate that the interaction between the ligand and the protein did not significantly affect the protein's flexibility, as depicted in Figure 10B.

Furthermore, an assessment was conducted on the radius of gyration (Rg) of the complex systems. The abbreviation "Rg" represents the root-mean-square distance of the protein atoms from the axis of rotation.[61] One of the most crucial attributes is the demonstration of how the simulation influences the protein structure's overall compactness and dimensions.[62] Proteins with a greater Rg value are characterized by reduced rigidity and increased flexibility, while proteins with a lower Rg value have enhanced stiffness and compactness.[59] In order to monitor the advancement of structural compactness, the Rg values of the atoms constituting the protein backbone were graphed during the course of time. The binding of urea resulted in a drop in the backbone Rg values to 20 ns. During the time interval spanning from 21 to 60 ns, there were no discernible fluctuations observed, and the magnitude of 3.0 nm remained rather stable. The Rg values were seen to exhibit constancy, with a range spanning from 2.90 to 2.95 nm. According to a comprehensive analysis, it was determined that the trajectory initially exhibited a peak value of 3.05 nm. The stability of the protein within the complex was subsequently demonstrated by the absence of any subsequent instances of this elevated value[58] (Figure 10C).

Also, an examination of the SASA was conducted on the complex. The SASA is often used in simulations to assess the extent to which solvent molecules are exposed to the receptor. This parameter provides valuable information about the molecular interactions within the system. Typically, the region of the receptor that comes into contact with the solvent may change due to the binding of a ligand.[61] The surface area variations were determined by plotting the measurements of protein SASA against time. The trajectory of the urease−urea complex exhibited a gradual decrease in values until reaching a point of stabilization at 15 ns. During the duration of the simulation, there were minor changes observed, except at certain

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

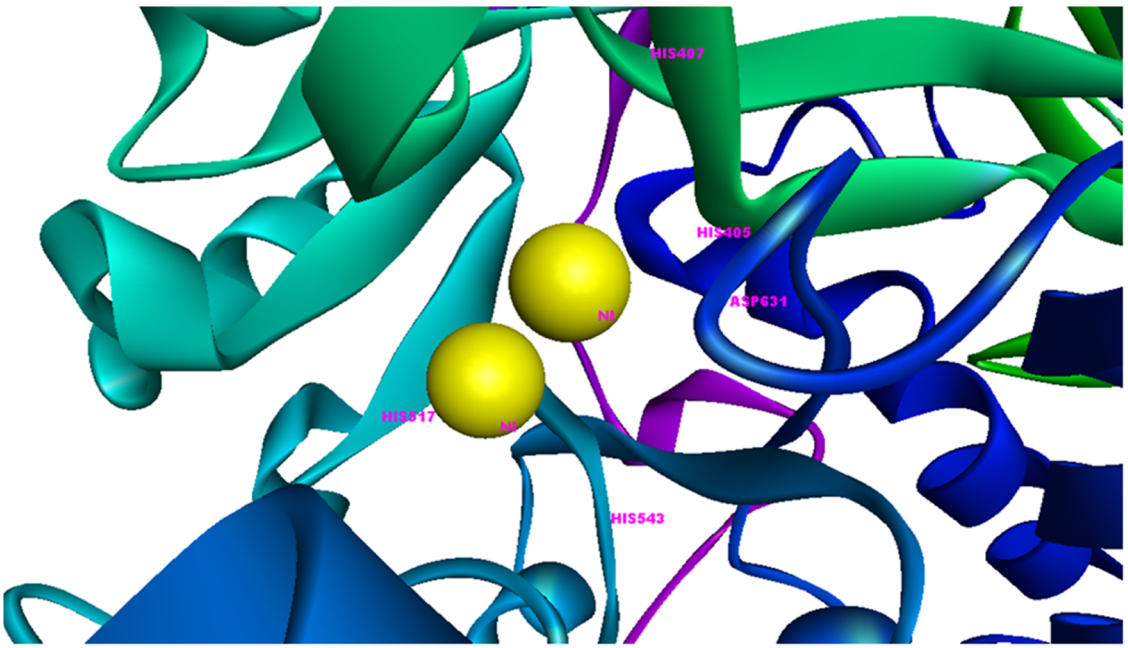


Figure 8. Specific residues His[405], His[407], His[517], His[453], and Asp[631] were present within *C. lanatus* urease that were shown to interact with the metal cofactor, i.e., Ni.

Table 3. Urea−Urease Docking Complexes, Binding Affinities, Interacting Amino Acids, and Atom Distances (H-Bond)[a]

		interacting amino acids/atoms (3-D)			
species	PDB ID	amino acids	amino acid atom	urea atom	distance (Å)
C. lanatus	predicted	His517	HD1	O1	2.81
		Gly548	HA2	O1	3.05
		Asp631	OD1	H7	2.31
		Ala634	H8	O	2.20

[a]The interacting amino acids of urease His517, Gly548, Asp631, and Ala634 interacted with urea through H-bonding within 3.05 Å.

time intervals (Figure 10D). The SASA exhibited a range of 350−300 nm^2, with a mean value of 323 nm^2. In general, the findings indicated that the simulation led to a reduction in the surface area of the protein complex.[63] To examine the affinity of the ligands for the target protein, the MD trajectories were analyzed to determine the extent of hydrogen bond formation during the entire simulation process. The SASA analysis revealed that there was a notable increase in the number of hydrogen bonds formed between the receptor protein and the target urea at different time intervals. This observation suggests a stronger binding affinity between the two molecules, as depicted in Figure 10E. Throughout the simulation, a consistent pattern of hydrogen bond production was observed, suggesting a high level of stability within the complex.[58] This finding provides convincing proof that the ligand urea exhibits the highest affinity for the target protein.

3.8. Binding Free Energy Analysis. The ΔG value, representing the binding free energy, was determined for the

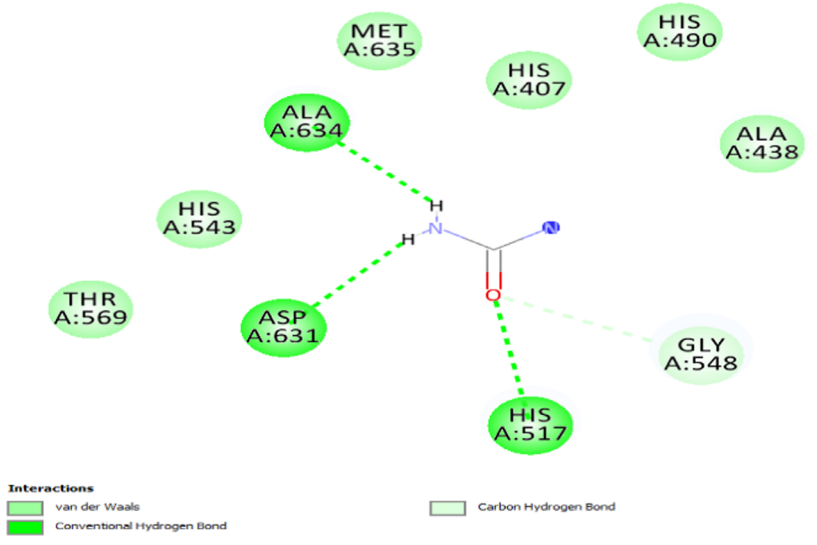


Figure 9. 2-D image docking structure of *C. lanatus* urease−urea and interacting amino acids within 3.5 Å.

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

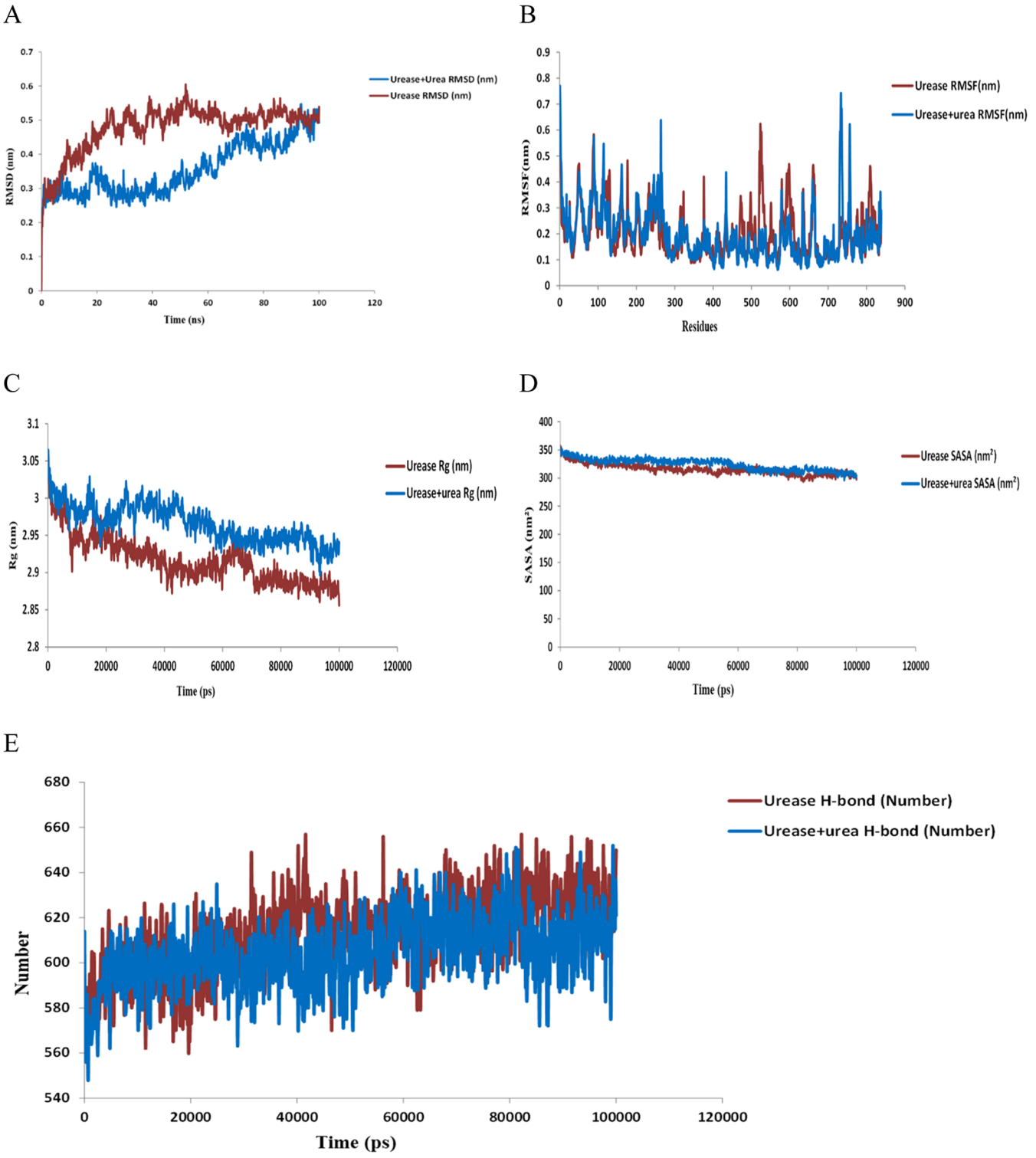


Figure 10. (A) RMSD study plot for 100 ns MD Simulation of urease and urea complex. (B) RMSF study plots for 100 ns MD Simulation. (C) Radius of gyration study plot for 100 ns MD Simulation of urease and urea complex. (D) Solvent accessible surface area study plot for 100 ns MD Simulation of urease and urea complex. (E) H-bond study plot for 100 ns MD Simulation of urease and urea complex.

urea−urease complex for a duration of 100 ns using the MM/PBSA method. The computed value was observed to be −7.61 kJ/mol. Table 4 includes specific information regarding energy details. The test ligand, i.e., urea exhibits a significant affinity for urease, as evidenced by the negative values of the free binding energies.[58]

4. CONCLUSIONS

In this study, after successful gene prediction, Protparam was used to analyze *C. lanatus* urease protein sequence, which showed a characteristic of 838 amino acid residues with 90.95 kDa mol. mass and an acidic theoretical pI of 5.71. Comparative studies of the physicochemical properties of 17 plant ureases revealed striking similarities in their primary, secondary, and

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285

Table 4. Binding Free Energy (ΔG) Calculated for the Urea−Urease Complex for 100 ns by Utilizing the MM/PBSA Method[a]

energy component	average (kJ/mol)
BOND	0
ANGLE	0
DIHED	0
UB	0
IMP	0
CMAP	0
VDWAALS	−7.68
EEL	−32.5
VDW	0
EEL	0
EGB	34.42
ESURF	−1.85
GGAS	−40.18
GSOLV	32.57
ΔG (binding energy)	−7.61

[a]The calculated value was −7.61 kJ/mol. Test ligand urea has a strong affinity for urease, as seen by its negative free binding energies.

domain structures, indicating significant conservation across the plant kingdom. The 3-D structure of *C. lanatus* urease was revealed in the current work using a homology modeling-based method. The modeled 3-D structure of *C. lanatus* urease was determined and the Ramachandran plot, ϕ, and Ψ calculations were done for quality assessment. PROCHECK analysis showed that 91.6% residues were present in most favored regions, and this result showed that the stereochemical quality was good for the predicted model. The findings of the molecular docking showed that the amino acids in urease His^{517}, Gly^{548}, Asp^{631}, and Ala^{634} were bound to urea through H-bonds. The results of urea−urease docking analysis suggested that different species might interact with one another in ways that are both common and rare depending on the species. To evaluate the stability of the modeled structure, a molecular dynamics simulation was performed. Molecular dynamic simulation results showed that during the simulation, the complex underwent minor structural modifications to the protein and the binding of the ligand had no noticeable impact on the protein flexibility. Overall, this study illustrates the salient structural feature as well as the urea and inhibitor binding interactions of the modeled structure of *C. lanatus*. *C. lanatus* urease has significant insecticidal and fungicidal properties, including the identification of Cluretox, a promising toxin for developing transgenic plants with enhanced resistance to insect herbivory and fungal pathogens. Understanding its catalytic mechanism can help design efficient enzyme-based drugs, improve crop yield through targeted fertilization, and develop sustainable waste management strategies. This knowledge can also aid in developing genetically modified crops that efficiently utilize nitrogen from fertilizers and targeted inhibitors to control urease activity in the soil.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.3c05993.

Full genomic length of *C. lanatus* urease (9138nt), full-length CDS of *C. lanatus* urease with 2517nt; Gene structure of *C. lantus* urease having exons and introns; ClustalW alignment of *C. moschata* (Acc. No. XP_022945701.1), *C. pepo* (Acc. No. XP_023541540.1), *M. charantia* (Acc. No. XP_022151796.1), *C. sativus* (Acc. No. XP_011655565.1), *C. maxima* (Acc. No. XP_022997369.1), *C. melo* (Acc. No. XP_008446059.1), *C. papaya* (Acc. No. XP_021899689.1), *M. indica* (Acc. No. XP_044489658.1), *G. hirsutum* (Acc. No. XP_016679630.1), *C. sativa* (Acc. No. XP_030496734.1), *C. sinensis* (Acc. No. KAH9772709.1), *C. cajan* (Acc. No. XP_020212240.1), *G. max* (Acc. No. NP_001236214.1), *C. arietinum* (Acc. No. XP_004502315.1), and *V. unguiculata* (Acc. No. XP_027940981.1); domain of *C. lanatus* urease. Key structural characteristics, such as the active site, cofactor binding sites, α helices, and β sheets of *C. lanatus* urease; PONDAR VL-TX, PONDAR VSL2, and PONDAR VL3 show intrinsic disorder propensity of Jaburetox in jack bean, Soyuretox in soybean, and Cluretox in *C. lanatus*; sequence alignment of Soyuretox from soybean and Cluretox from *C. lanatus* with the sequence of Jaburetox from jack bean; prominent active site residues in Jack bean (PDB id: 4GY7) for reference; and sequence alignment between Jack bean (*C. ensiformis*) and *C. lanatus* showing prominent active site residues (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Arvind M. Kayastha − *School of Biotechnology, Institute of Science, Banaras Hindu University, Varanasi 221005, India;* orcid.org/0000-0002-5090-7159; Phone: +91-9451000575; Email: kayasthabhu@gmail.com

Authors

Prince Kumar − *School of Biotechnology, Institute of Science, Banaras Hindu University, Varanasi 221005, India*

Arpan Kayastha − *Department of Biosciences and Bioengineering, IIT Roorkee, Roorkee 247667, India*

Vinay Kumar Singh − *School of Biotechnology, Institute of Science, Banaras Hindu University, Varanasi 221005, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c05993

Author Contributions

P.K., A.K., V.K.S., and A.M.K. contributed to writing the manuscript and editing it. All co-authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are thankful to the Centre of Bioinformatics, School of Biotechnology, B.H.U. for the computational facilities. P.K. thanks DBT, New Delhi, for Junior and Senior research fellowships. A.K. thanks UGC, New Delhi, for a Junior and Senior research fellowship.

■ REFERENCES

(1) Andrews, R. K.; Blakeley, R. L.; Zerner, B. Urea and urease. *Adv. Inorg. Biochem.* **1984**, *6*, 245−283.

(2) Sumner, J. B. Isolation and crystallization of the enzyme urease. *J. Biol. Chem.* **1926**, *69*, 435−441.

(3) Jabri, E.; Carr, M. B.; Hausinger, R. P.; Karplus, P. A. The crystal structure of urease from *Klebsiella aerogenes*. *Science* **1995**, *268*, 998−1004.
(4) Balasubramanian, A.; Ponnuraj, K. Crystal Structure of the first plant urease from Jack Bean: 83 Years of Journey from its first crystal to molecular structure. *J. Mol. Biol.* **2010**, *400*, 274−283.
(5) Dixon, N. E.; Gazzola, C.; Blakeley, R. L.; Zerner, B. Jack bean urease (EC 3.5.1.5), A metalloenzyme, A simple biological role for nickel. *J. Am. Chem. Soc.* **1975**, *97*, 4131−4133.
(6) Mobley, H. L. T.; Hausinger, R. P. Microbial ureases: significance, regulation and molecular characterization. *Microbiol. Rev.* **1989**, *53*, 85−103.
(7) Blakeley, R. L.; Zerner, B. Jack bean urease, the first nickel enzyme. *J. Mol. Catal.* **1984**, *23*, 263−292.
(8) Das, N.; Kayastha, A. M. Immobilization of urease from pigeonpea (*Cajanus cajan L.*) on flannel cloth using polyethyleneimine. *World J. Microbiol. Biotechnol.* **1998**, *14*, 927−929.
(9) Kayastha, A. M.; Das, N. Kinetics of thermal inactivation and molecular asymmetry of urease from dehusked pigeongea (*Cajanus cajan* L.) seeds. *J. Plant Biochem. Biotechnol.* **1998**, *7*, 121−124.
(10) Das, N.; Kayastha, A. M.; Malhotra, O. P. Immobilization of urease from pigeonpea (*Cajanus cajan L.*) in polyacrylamide gels and calcium alginate beads. *Biotechnol. Appl. Biochem.* **1998**, *27*, 25−29.
(11) Kayastha, A. M.; Srivastava, P. K. Pigeonpea (*Cajanus cajan L.*) urease immobilized on glutaraldehyde-activated chitosan beads and its analytical applications. *Appl. Biochem. Biotechnol.* **2001**, *96*, 41−53, DOI: 10.1385/abab:96:1-3:041.
(12) Reddy K, R. C.; Srivastava, P. K.; Dey, P. M.; Kayastha, A. M. Immobilization of pigeonpea (*Cajanus cajan*) urease on DEAE-cellulose paper strips for urea estimation. *Biotechnol. Appl. Biochem.* **2004**, *39*, 323−327, DOI: 10.1042/BA20030122.
(13) Dwevedi, A.; Routh, S. B.; Yadav, A. S.; Singh, A. K.; Srivastava, O. N.; Kayastha, A. M. Response surface analysis of nano-ureases from *Canavalia ensiformis* and *Cajanus cajan*. *Int. J. Biol. Macromol.* **2011**, *49*, 674−680.
(14) Polacco, J. C.; Holland, M. A. Roles of urease in plant cells. *Int. Rev. Cytol.* **1993**, *145*, 65−103.
(15) Polacco, J. C.; Holland, M. A. Genetic control of plant ureases. *Genet. Eng.* **1994**, 33−48.
(16) Mobley, H. L. T.; Island, M. D.; Hausinger, R. P. Molecular biology of microbial ureases. *Microbiol. Rev.* **1995**, *59*, 451−480.
(17) Follmer, C. Insights into the role and structure of plant ureases. *Phytochemistry* **2008**, *69*, 18−28.
(18) Balasubramanian, A.; Durairajpandian, V.; Elumalai, S.; Mathivanan, N.; Munirajan, A. K.; Ponnuraj, K. Structural and functional studies on urease from pigeon pea (*Cajanus cajan*). *Int. J. Biol. Macromol.* **2013**, *58*, 301−309.
(19) Sirko, A.; Brodzik, R. Plant ureases: roles and regulation. *Acta. Biochim. Polym.* **2000**, *47*, 1189−1195.
(20) Polacco, J. C.; Holland, M. A. Roles of urease in plant cells. *Int. Rev. Cytol.* **1993**, *145*, 65−103.
(21) Carlini, C. R.; Polacco, J. C. Toxic properties of urease. *Crop Sci.* **2008**, *48*, 1665−1672.
(22) Follmer, C.; Real-Guerra, R.; Wasserman, G. E.; Olivera-Severo, D.; Carlini, C. R. Jackbean, soybean and *Bacillus pasteurii* ureases: biological effects unrelated to ureolytic activity. *Eur. J. Biochem.* **2004**, *271*, 1357−1363.
(23) Carlini, C. R.; Ligabue-Braun, R. Ureases as multifunctional toxic proteins: A review. *Toxicon* **2016**, *110*, 90−109.
(24) Kappaun, K.; Piovesan, A. R.; Carlini, C. R.; Ligabue-Braun, R. Ureases: Historical aspects, catalytic, and non-catalytic properties-A review. *J. Adv. Res.* **2018**, *13*, 3−17.
(25) Ligabue-Braun, R.; Carlini, C. R. Moonlighting toxins: Ureases and beyond. In *Plant Toxins*; Springer, 2017; pp 199−219 DOI: 10.1007/978-94-007-6464-4_10.
(26) Real-Guerra, R.; Stanisçuaski, F.; Carlini, C. R. Soybean Urease: Over a Hundred Years of Knowledge. In *A Comprehensive Survey of International Soybean Research-Genetics, Physiology, Agronomy and Nitrogen Relationships*; IntechOpen, 2013.
(27) Becker-Ritt, A. B.; Martinelli, A. H. S.; Mitidieri, S.; Feder, V.; Wassermann, G. E.; Santi, L.; Carlini, C. R. Antifungal activity of plant and bacterial ureases. *Toxicon* **2007**, *50*, 971−983, DOI: 10.1016/j.toxicon.2007.07.008.
(28) Martinelli, A. H.; Lopes, F. C.; Broll, V.; Defferrari, M. S.; Ligabue-Braun, R.; Kappaun, K.; Carlini, C. R. Soybean ubiquitous urease with purification facilitator: An addition to the moonlighting studies toolbox. *Proc. Biochem.* **2017**, *53*, 245−258, DOI: 10.1016/j.procbio.2016.12.003.
(29) Oldfield, C. J.; Dunker, A. K. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **2014**, *83*, 553−584.
(30) Uversky, V. N. Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J.* **2009**, *28*, 305−325.
(31) Mulinari, F.; Staniscuaski, F.; Bertholdo-Vargas, L. R.; Postal, M.; Oliveira-Neto, O. B.; Rigden, D. J.; Carlini, C. R. Jaburetox-2Ec: an insecticidal peptide derived from an isoform of urease from the plant *Canavalia ensiformis*. *Peptides* **2007**, *28*, 2042−2050, DOI: 10.1016/j.peptides.2007.08.009.
(32) Martinelli, A. H.; Kappaun, K.; Ligabue-Braun, R.; Defferrari, M. S.; Piovesan, A. R.; Stanisçuaski, F.; Demartini, D. R.; Dal Belo, C. A.; Almeida, C. G.; Follmer, C.; Verli, H. Structure−function studies on jaburetox, a recombinant insecticidal peptide derived from jack bean (*Canavalia ensiformis*) urease. *Biochim. Biophy. Acta, Gen. Sub.* **2014**, *1840*, 935−944, DOI: 10.1016/j.bbagen.2013.11.010.
(33) Yata, V. K.; Thapa, A.; Mattaparthi, V. S. K. Structural insight into the binding interactions of modelled structure of *Arabidopsis thaliana* urease with urea: an *in-silico* study. *J. Biomol. Struct. Dyn.* **2015**, *33*, 845−851.
(34) Filiz, E.; Vatansever, R.; Ozyigit, I. I. Molecular docking of *Glycine max* and *Medicago truncatula* ureases with urea; bioinformatics approaches. *Mol. Biol. Rep.* **2016**, *43*, 129−140, DOI: 10.1007/s11033-016-3945-7.
(35) Solovyev, V.; Kosarev, P.; Seledsov, I.; Vorobyev, D. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* **2006**, *7*, No. S10, DOI: 10.1186/gb-2006-7-s1-s10.
(36) Stanke, M.; Steinkamp, R.; Waack, S.; Morgenstern, B. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res.* **2004**, *32*, W309−W312.
(37) Guo, S.; Zhao, S.; Sun, H.; Wang, X.; Wu, S.; Lin, T.; Ren, Y.; Gao, L.; Deng, Y.; Zhang, J.; Lu, X. Resequencing of 414 cultivated and wild watermelon accessions identifies selection for fruit quality traits. *Nat. Genet.* **2019**, *51*, 1616−1623, DOI: 10.1038/s41588-019-0518-4.
(38) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **2003**, *31*, 3784−3788.
(39) Hu, B.; Jin, J.; Guo, A. Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: an upgraded gene features visualization server. *Bioinformatics* **2015**, *31*, 1296−1297.
(40) Mullan, L. J.; Williams, G. W. BLAST and go. *Briefings Bioinf.* **2002**, *3*, 200−202.
(41) Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T. G.; Bertoni, M.; Bordoli, L.; Schwede, T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **2014**, *42*, W252−W258.
(42) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283−291.
(43) Colovos, C.; Yeates, T. O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* **1993**, *2*, 1511−1519.
(44) Bowie, J. U.; Lüthy, R.; Eisenberg, D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **1991**, *253*, 164−170.
(45) Lopes, F. C.; Dobrovolska, O.; Real-Guerra, R.; Broll, V.; Zambelli, B.; Musiani, F.; Uversky, V. N.; Carlini, C. R.; Ciurli, S. Pliable

natural biocide: Jaburetox is an intrinsically disordered insecticidal and fungicidal polypeptide derived from jack bean urease. *FEBS J.* **2015**, *282*, 1043−1064.
(46) Romero, P.; Obradovic, Z.; Li, X.; Garner, E. C.; Brown, C. J.; Dunker, A. K. Sequence complexity of disordered protein. *Proteins: Struct., Funct., Bioinf.* **2001**, *42*, 38−48.
(47) Peng, K.; Radivojac, P.; Vucetic, S.; Dunker, A. K.; Obradovic, Z. Length-dependent prediction of protein intrinsic disorder. *BMC Bioinform.* **2006**, *7*, No. 208, DOI: 10.1186/1471-2105-7-208.
(48) Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C. J.; Dunker, A. K.; Obradovic, Z. Optimizing long intrinsic disorder predictors with protein evolutionary information. *J. Bioinform. Comput. Biol.* **2005**, *3*, 35−60, DOI: 10.1142/s0219720005000886.
(49) Geourjon, C.; Deleage, G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput. Appl. Biosci.* **1995**, *11*, 681−684.
(50) Kumari, R.; Kumar, R.; Consortium, O. S. D. D.; Lynn, A. g_mmpbsa-A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **2014**, *54*, 1951−1962, DOI: 10.1021/ci500020m.
(51) Grahl, M. V. C.; Lopes, F. C.; Martinelli, A. H. S.; Carlini, C. R.; Fruttero, L. L. Structure-Function insights of Jaburetox and Soyuretox: Novel intrinsically disordered polypeptides derived from plant ureases. *Molecules* **2020**, *25*, 5338−5360.
(52) Saraswat, J.; Singh, P.; Patel, R. A computational approach for the screening of potential antiviral compounds against SARS-CoV-2 protease: Ionic liquid vs herbal and natural compounds. *J. Mol. Liq.* **2021**, *326*, No. 115298.
(53) Kumar, D.; Kumari, K.; Vishvakarma, V. K.; Jayaraj, A.; Kumar, D.; Ramappa, V. K.; Patel, R.; Kumar, V.; Dass, S. K.; Chandra, R.; Singh, P. Promising inhibitors of main protease of novel corona virus to prevent the spread of COVID-19 using docking and molecular dynamics simulation. *J. Biomol. Struct.* **2021**, *39*, 4671−4685.
(54) Saraswat, J.; Riaz, U.; Patel, R. In-silico study for the screening and preparation of ionic liquid-AVDs conjugate to combat COVID-19 surge. *J. Mol. Liq.* **2022**, *359*, No. 119277.
(55) Saraswat, J.; Kumar, S.; Alzahrani, K. A.; Malik, M. A.; Patel, R. Experimental and Computational Characterisation of the Molecular Interactions between 1-Butyl-1-methyl-pyrrolidin-1-ium bis (trifluoromethanesulphonyl) imide and Human Serum Albumin. *ChemistrySelect* **2023**, *8*, No. e202204159, DOI: 10.1002/slct.202204159.
(56) Azeem, K.; Ahmed, M.; Uddin, A.; Singh, S.; Patel, R.; Abid, M. Comparative investigation on interaction between potent antimalarials and human serum albumin using multispectroscopic and computational approaches. *Luminescence* **2023**, *38*, 2018−2033, DOI: 10.1002/bio.4590.
(57) Sharma, A.; Datta, D.; Balasubramaniam, R. Molecular dynamics simulation to investigate the orientation effects on nanoscale cutting of single crystal copper. *Comput. Mater. Sci.* **2018**, *153*, 241−250.
(58) Kalimuthu, A. K.; Panneerselvam, T.; Pavadai, P.; Pandian, S. R. K.; Sundar, K.; Murugesan, S.; Ammunje, D. N.; Kumar, S.; Arunachalam, S.; Kunjiappan, S. Pharmacoinformatics-based investigation of bioactive compounds of Rasam. *Sci. Rep.* **2021**, *11*, No. 21488, DOI: 10.1038/s41598-021-01008-9.
(59) Kumar, B.; Parasuraman, P.; Murthy, T. P. K.; Murahari, M.; Chandramohan, V. In silico screening of therapeutic potentials from *Strychnos nux-vomica* against the dimeric main protease (Mpro) structure of SARS-CoV-2. *J. Biomol. Struct. Dyn.* **2021**, *40*, 7796−7814, DOI: 10.1080/07391102.2021.1902394.
(60) Abdelrheem, D. A.; Ahmed, S. A.; Abd El-Mageed, H. R.; Mohamed, H. S.; Rahman, A. A.; Elsayed, K. N.; Ahmed, S. A. The inhibitory effect of some natural bioactive compounds against *SARS-CoV-2* main protease: insights from molecular docking analysis and molecular dynamic simulation. *J. Environ. Sci. Health, Part A* **2020**, *55*, 1373−1386.
(61) Rahman, M. M.; Saha, T.; Islam, K. J.; Suman, R. H.; Biswas, S.; Rahat, E. U.; Hossen, M. R.; Islam, R.; Hossain, M. N.; Mamun, A. A.; Khan, M. Virtual screening, molecular dynamics and structure-activity relationship studies to identify potent approved drugs for Covid-19 treatment. *J. Biomol. Struct. Dyn.* **2021**, *39*, 6231−6241, DOI: 10.1080/07391102.2020.1794974.
(62) Gupta, S.; Singh, A. K.; Kushwaha, P. P.; Prajapati, K. S.; Shuaib, M.; Senapati, S.; Kumar, S. Identification of potential natural inhibitors of *SARS-CoV2* main protease by molecular docking and simulation studies. *J. Biomol. Struct. Dyn.* **2021**, *39*, 4334−4345.
(63) Ghosh, S. K.; Saha, B.; Banerjee, R. Insight into the sequence-structure relationship of TLR cytoplasm's Toll/Interleukin-1 receptor domain towards understanding the conserved functionality of TLR 2 heterodimer in mammals. *J. Biomol. Struct. Dyn.* **2021**, *39*, 5348−5357.

https://doi.org/10.1021/acsomega.3c05993
ACS Omega 2024, 9, 2272−2285