

### **Open Access** Original Article



# A molecular dynamic modeling and *in silico* analysis approach to three-dimensional structure of glutelin type-B 5-like proteins from proso millet: effects of temperature and electric field

Felix Akharume<sup>1</sup>, Ashutosh Singh<sup>2</sup>, Konstantin V. Korotkov<sup>3</sup>, Akinbode Adedeji<sup>1\*</sup><sup>10</sup>

<sup>1</sup>Department of Biosystems and Agricultural Engineering, University of Kentucky, Lexington, KY 40506, USA <sup>2</sup>School of Engineering, University of Guelph, Guelph, ON N1G2W1, Canada <sup>3</sup>Department of Molecular and Cellular Biochemistry, University of Kentucky, Lexington, KY 40506, USA

\*Correspondence: Akinbode Adedeji, Department of Biosystems and Agricultural Engineering, University of Kentucky, 128 Charles E. Barnhart Building, Lexington, KY 40506, USA. akinbode.adedeji@uky.edu Academic Editor: Zuhaib F Bhat, SKUAST-Jammu, India Received: January 18, 2025 Accepted: July 3, 2025 Published: July 17, 2025

**Cite this article:** Akharume F, Singh A, Korotkov KV, Adedeji A. A molecular dynamic modeling and *in silico* analysis approach to three-dimensional structure of glutelin type-B 5-like proteins from proso millet: effects of temperature and electric field. Explor Foods Foodomics. 2025;3:101088. https://doi.org/10.37349/eff.2025.101088

## Abstract

**Aim:** Protein structure-function relationship understanding is very important to the study of protein chemistry. In this study, we determined the 3D structure of glutelin type-B 5-like (GTB) protein from proso millet and investigated the effects of simulated temperature and static electric field on GTB protein's secondary structures, solvent accessibility surface area (SASA), the radius of gyration (Rg), root mean square deviation (RMSD), and total dipole moment with the view to elucidate its structural behavior under different simulated processing stresses.

**Methods:** This study leveraged homology modeling to determine the 3D structure of the GTB protein from proso millet. We further explored *in silico* modeling using molecular dynamic modeling tools to study the behavior of the GTB protein structure under simulated temperatures and static electric field processing stresses.

**Results:** The 3D protein structure of GTB is majorly a protomer that comprises three monomers with one jelly-like  $\beta$ -barrel and two extended helix domains, and the remaining 35% as coils. From the molecular dynamic simulation, the secondary structure of the protein was not disrupted at temperatures between 300–400 K and static electric fields between 0.1–3 V/nm, albeit with the loss of a few amino acid residues. The RMSD increased significantly with temperature, static electric field, and simulation time increase. However, the Rg and SASA decreased for the same conditions, suggesting that the protein structure is compacting during thermal and electrical treatment.

**Conclusions:** We concluded that low levels of simulated temperature and low levels of the electric field may cause a temporary reversible conformational change of GTB protein, not enough to cause a major disruption as may be expected during empirical processing of protein. To achieve greater disruption that would mimic protein denaturation, there is a need for *in silico* study of protein molecules under more

© The Author(s) 2025. This is an Open Access article licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



aggressive simulated stressors for a longer time, and consideration given to protein structure development from AlphaFold3.

## **Keywords**

Comparative homology modeling, molecular dynamic modeling, proso millet, glutelin proteins, structurefunction relationship, three-dimensional structure

## Introduction

The three-dimensional structures of food proteins dictate their biological and sometimes processing functionality. The structure-function relationship has remained a central dogma in elucidating the mechanism of protein functionality at the molecular level. This understanding provides information to identify processes and technologies to improve protein functionality and/or suppress undesirable functionality such as allergenicity and unwanted flavor [1]. For instance, Tandang-Silvas et al. [2] compared the three-dimensional structure of 11S proglobulins of amaranth, procruciferin, pumpkin pro-11S, soybean proA1aB1b, soybean proA3B4, and pea prolegumin and showed that the smaller cavity  $(4,167.1 \text{ Å}^3)$  of 11S proglobulin of amaranth seed was responsible for the higher midpoint denaturation temperature  $(T_m)$ values (98.85°-102.85°C) and the thermal stability, compared to other 11S proglobulins. Pantoja-Uceda et al. [3] revealed that the presence of a hydrophobic patch (five methionine residues that partially hide Trp76) at the surface of the three-dimensional structure protein surface may have been responsible for the high emulsifying properties of a 2S albumin protein from seeds of sunflower. The authors showed that the presence of a superhelix motif, especially a "hypervariable loop" in the sunflower albumin 8 (SFA-8) structure of sunflower protein, which is also common to other 2S albumin proteins, may have been responsible for the allergenicity of the sunflower proteins. Fukuda et al. [4] compared the threedimensional structures of 7S globulin proteins from dzuki beans (7S1, 7S3), soybeans ( $\beta$ -conglycinin  $\beta$ ,  $\beta$ conglycinin  $\alpha' c$ ), and mung beans (8S $\alpha$ ). Their work revealed that the differences in their thermal stability and solubilities are related to the size of the cavity found inside the molecules of each of these proteins. Furthermore, they showed that the difference in electrostatic surface potential distribution may cause differential solubility at pH 8. For instance, while the 7S1, 7S3, and 8S $\alpha$  were soluble at pH 8, the  $\beta$ conglycinin  $\beta$  and  $\beta$ -conglycinin  $\alpha$ 'c were not [4].

Molecular dynamic (MD) modeling is one of several tools that is gaining interest in studying the structure-function relationship of food proteins. Unlike traditional empirical investigations, MD provides an elucidation of structural changes at the atomic and molecular levels under dynamics-simulated processing stresses. Several authors have studied the structure-function relationship of food protein using MD modeling. Vagadia et al. [5] investigated the effects of temperature (300 K to 393 K) and oscillating electric field (0.5 V/nm and 2.45 GHz) of the soybean trypsin inhibitor (STI) *in silico*. The study showed that the effects of the external stresses disrupted the amino acid (AA) residues (31–41, 61–66) and a rearrangement of the secondary structure (coils and turns), but the core of the STI was still stable due to an antiparallel  $\beta$ -sheet structure [5].

In another earlier study, a myoglobin molecule in water exposed to pulsed and static electric fields ( $10^8$  V/m to  $10^9$  V/m) was examined using MD; the authors reported that a remarkable structural rearrangement in the myoglobin molecules can be achieved by the application of  $10^9$  V/m pulsed electric field for a few picoseconds or static electric field for a few nanoseconds [6]. It is confirmed by several authors that only a high electric field strength can cause a major disruption in the secondary structure of protein molecules [1, 6, 7].

Glutelin is the second-largest group of storage proteins in proso millet, with the protein concentration accounting for about 34–39% of the total protein content [8]. The glutelin group forms most of the proteins in proso millet protein concentrate (MPC) as produced by acid/alkaline extraction methods; the MPCs are used as a protein ingredient for food applications such as making cookies, bread, and other ready-to-eat

products. Like other cereal proteins, the wide applications of MPC in food applications are limited by their low functionality (solubility, emulsion, and foam properties) compared to other premium ingredients like similar proteins from soy, pea, and dairy [8, 9]. Knowledge of the structure-function relationship of protein is needed to understand how the functionality of MPC could be improved. The traditional method of structure-function relationships does not allow for multiple scenario analysis and could be costly to implement. There is a significant benefit in *in silico* modeling of stress environments that modify proteinstructure functions [8, 10, 11]. MD modeling has been used in biological sciences to study the protein structure of various food and non-food proteins under different environmental conditions, such as temperature, pH, electric field, ionic strength, and pressure. In particular, MD modeling allows researchers to simulate how these stressors induce conformational changes in protein secondary and tertiary structures, revealing mechanisms that affect solubility, aggregation, and binding interactions. For proso millet glutelin, MD modeling could be employed to identify specific AA residues or structural motifs that contribute to poor functional properties, thereby enabling the rational design of protein modifications or processing conditions to enhance emulsification, foaming, and solubility. This computational approach provides a powerful, cost-effective tool for screening multiple structural scenarios before experimental validation, accelerating the development of functional millet-based protein ingredients for diverse food systems [11, 12].

The first step to the structural prediction and structure-function study of such proteins begins with the identification of their AA sequence and, in some cases, their close homologs (usually AA similarity > 30%). Very recently, the genes that coded for the AA sequence glutelin storage proteins in proso millet were annotated and reported by Shi et al. [13]. They identified three glutelin protein subunits. One "glutelin type-B 5-like" protein with 256 AA residues and two "glutelin-2-like" proteins, possible isoforms, one having an AA residue of 156 and the other with AA residues of 161 AA residues [13]. In our previous study, we had identified that "glutelin type-B 5-like" protein has close homologs (above 40% structural identity) in the Protein Data Bank (PDB), which could facilitate easy determination of its three-dimensional structure, unlike the two "glutelin-2-like" proteins with zero homologs. We further cloned the gene that codes for the "glutelin type-B 5-like" protein and performed purification steps, but were unable to obtain a stable protein for X-ray crystallography [10]. Owing to the difficulties and time constraints associated with the empirical determination of the three-dimensional structure of the glutelin storage proteins from proso millet and lack access to the AlphaFold platforms (an AI program developed by DeepMind, a subsidiary of Alphabet, capable of performing prediction of protein structure at a higher level than the traditional approach) as at the time of the study, we, therefore, determined its three-dimensional structures using the traditional comparative homology modeling and predicted some of its physicochemical properties. Additionally, we investigated the effects of temperature with static electric fields on the "glutelin type-B 5-like" protein using MD modeling. We evaluated three levels of temperatures (300 K, 350 K, and 400 K) and three levels of static electric fields (0.1 V/nm, 1 V/nm, and 3 V/nm) on the root mean square deviation (RMSD), root mean square fluctuations (RMSF) per residue, the radius of gyration (Rg), total dipole moment, volume, density, secondary structure analysis, and solvent accessibility surface area (SASA). The study of "glutelin type-B 5like" in silico provides a model for understanding the 11–12S globulin proteins of most cereals, which are widely homologous to the "legumin" type globulin protein present in most legumes.

## **Materials and methods**

### Protein sequence identification and analysis

The AA sequence corresponding to the "glutelin type-B 5-like" proteins (256 AA) was retrieved from the National Centre for Biotechnology Information (NCBI, www.ncbi.nlm.nih.gov) in FASTA format with accession number RLM98351 [11, 12]. The AA sequence was searched by protein-protein BLAST against the PDB proteins to identify templates for model building. The templates with over 35% AA identity were selected for the "glutelin type-B 5-like" protein. A summary of the protein details used for homology modeling is presented in Table S1.

#### **Comparative homology modeling**

Template-based protein structure modeling of "glutelin type-B 5-like" protein was carried out using the Swiss-Model server (https://swissmodel.expasy.org/) [11–15] and the Robetta web server (http://robetta. bakerlab.org) [16] using the five templates with over 35% sequence identified from the PDB database (IDs: 2E9Q, 3QAC, 3KSC, 3C3V, and 1UD1). The procedure for modeling using the Swiss-Model and the Robetta web server has been documented by [14, 17]. Unlike the Swiss-Model server, the Robetta web server returned only the monomer of "glutelin type-B 5-like" protein, and since the active form of "glutelin type-B 5-like" protein is a trimer, the homo-trimers configuration of "glutelin type-B 5-like" protein was further modeled into a trimer with the Galaxy Homomer web server (https://galaxy.seoklab.org/cgi-bin/submit. cgi?type=PEPDOCK) [18, 19]. Structural comparison of all the models (a total of five) generated by the two structure-building servers was evaluated by the Partial Order Structure Alignment (POSA) server using the RMSD as an index of similarity [20]. Assessment of the model quality was performed by the Swiss-Model structural assessment tool to identify the best model quality from all models. The best model, determined based on Qualitative Model Energy Analysis (QMEAN) [15], was selected and used for geometry minimization.

#### Geometry minimization and model re-assessment

Geometry minimization of selected models for the "glutelin type-B 5-like" protein one from each modeling server was carried out by the geometry minimization tool of Python-based Hierarchical Environment for Integrated Xtallography (PHENIX) [21] with 1,000 maximum iterations and five macro-cycles. The evaluation of refined models was assessed by the Psi/Phi Ramachandran plot from PROCHECK [22]. The overall quality of the models was assured using Z-scores of the ProSA-web tool (https://prosa.services. came.sbg.ac.at/prosa.php/) [23, 24] and the QMEAN with Distance Constraints (QMEANDisCo) [25]. Furthermore, the final model was submitted to the Protein Model Database (PMDB, http://srv00.recas.ba. infn.it/PMDB/main.php) with PMDB identifier PM0083241.

#### **MD simulations**

The MD simulations of the "glutelin type-B 5-like" in water were carried out by the Groningen Machine for Chemical Structure (GROMACS) software package (version 2018.1, Stockholm Center for Biomembrane Research, Stockholm, Sweden) [26]. The "glutelin type-B 5-like" protein used in this study can be downloaded from the PMDB with the identity number PM0083241, where we had deposited it. The protein is a protomer of three non-covalently linked monomers, with each monomer chain comprising 5.0%  $\alpha$ -helix, 2.5% 3/10 helix, 23.8%  $\beta$ -sheet, and 69.17% coils/turns/bends/bridges (Figure 1). The protein-containing 11,574 atoms from 768 residues were kept in a cubic box of a dimension 11.683 nm × 11.683 nm × 11.683 nm as a periodic boundary condition, and the protein was solvated with 146,262 atoms of water molecules neutralized with 6 sodium ions.

The OPLS-AA/L all-atom force field [27] and SPC/E water model [28] were selected to provide potential energy function and water parameters to the system. Following neutralization of the systems to mimic the physiological state of the protein, the systems were energy minimized to converge at maximum force value < 1,000.0 kJ/(mol·nm) using a stepwise descent minimization algorithm for 100,000 steps, and the systems were further equilibrated to a constant number of particles, volume, and temperature (NVT) and a constant number of particles, pressure, and temperature (NPT) for 200 ps at 300 K and 1 bar. The MD simulations were run for 1 ns using a leap-frog integrator algorithm during which the temperature of the systems was maintained using a modified Berendsen thermostat, and the pressure was maintained using the Parrinello and Rahman barostat [29, 30]. A total of nine simulations were run for static electric field levels coupled with the temperature levels (Table 1). The results of simulations such as RMSD, Rg, and SASA were analyzed using GROMACS inbuilt tools. GROMACS inbuilt dictionary of the secondary structure of protein (DSSP) [31, 32] was used for the secondary structure analysis, and virtual MDs (VMD) [33] were used to visualize protein conformational change.



Figure 1. A snapshot of the three-dimensional structure of the "glutelin type-B 5-like" protein molecule

Temperature	Pressure	Static electric field	Simulation length
300 K	1 bar	0.1 V/nm	1 ns
		1 V/nm	
		3 V/nm	
350 K	1 bar	0.1 V/nm	1 ns
		1 V/nm	
		3 V/nm	
400 K	1 bar	0.1 V/nm	1 ns
		1 V/nm	
		3 V/nm	

Table 1. A summary of experimental process conditions

#### Experimental design and statistical analysis

A completely randomized  $2^3$ -factorial design was used in this experiment. Two treatment factors of temperature and static electric field, and three levels per treatment were considered. Six response variables (RMSD, total dipole moment, Rg, SASA, volume, and density) average replicates were analyzed by one-way-ANOVA using the GLM procedure of SAS (SAS Institute, 2004) to determine the differences in means at a significant level of *P* < 0.05. Tukey's test was used to assess the difference among specific treatment means at *P* < 0.05. The temperature and pressure levels were selected to represent typical industrial process conditions (room temperature, denaturation temperature, and super-critical temperature range).

### **Results**

### BLAST analysis of protein sequence alignment

Five templates were selected for building the "glutelin type-B 5-like" protein. The templates were obtained from the PDB by searching using the "glutelin type-B 5-like" protein AA as the query and for sequence alignment on the NCBI database. The search query returned about 16 hits (matches) with sequence similarity above 30%. Only five templates (ID: 2E9Q, 3QAC, 3KSC, 1UD1, and 1FXZ) were considered based on sequence similarity above 35%, the expected value (E-value), and conformational resemblance with the

empirical data of the structure of the protein (i.e., trimer). The E-value gives information on the likelihood of getting a hit or a protein by chance. Thus, the lower the score, the better the template fits our model. A summary of the results from the search query for the templates is shown in Table S2. Other crystal structures with similar identities and E-values were not selected due to non-conformation with the empirical data. For example, the crystal structure of coconut allergen cocosin (ID: 5WPW) with 38.35% similarity and an E-value of  $8 \times 10^{-25}$  was not selected because its natural state was a homo hexamer while our protein from empirical studies is a homo-trimer [11].

#### Model building and assessment

A total of ten models of the "glutelin type-B 5-like" protein were developed (five from each server) using two of the top-ranked and widely used structural building servers (Robetta and Swiss-Model servers) as ranked according to the community-wide critical assessment of predicted interactions (CAPRI) [12]. Unlike the Swiss-Model, which builds a model by transferring structural information of a conserved atom from the template onto the model, the Robetta web server builds up the model by turning the protein sequence into a parsed domain and then predicts the model from the domain [14, 17]. Protein residues that cannot be modeled were built by de novo structure prediction by both servers. Inputting both the target sequence ("glutelin type-B 5-like") and the templates into the servers generated five models, one for each template. Unlike the Robetta web server, the Swiss-Model server returned models with truncated residues: residues 1-76 and 245-256 were removed by moving the N- (AA residue) and C- (carboxyl group) terminals to glutamine (77) and glycine (244), respectively. This is expected since all the templates have five variable regions (I, II, III, IV, and V) that were removed from their original AA sequence because the residues in these regions caused disorder in the crystal growth during experimentation of their homologs [34]. These regions are mainly looped within all the AA sequences. Generally, the "glutelin type-B 5-like" protein model reveals a protomer composed of three monomers, which are linked non-covalently, as can be seen in Figure 2. The model tertiary structure contains a jelly-like  $\beta$ -barrel and two extended helix domains. These domains are highly conserved relative to the different 11S or pro-11S globulin proteins of pea, soybean, pumpkin, amaranth, rapeseed, and Ara h protein of peanuts, except that these other proteins have two jellylike  $\beta$ -barrels, not one as seen in their homologs [2, 34, 35]. The structure also contains two major variable regions at the terminals. This structure was not unexpected, as many of the "glutelin type-B 5-like" protein homologs displayed similar protomer structure. Empirically, from another study we conducted, we observed from cloning, expression, and purification of the "glutelin type-B 5-like" protein, that the protein is a trimer with a molecular weight of 96 kDa, as against its predicted molecular weight of 22.4 kDa [11].

To evaluate the ten models generated from the servers, the RMSD of the superimposed models from each of the servers showed a good fitting of the protein main chain with minimal deviations of 1.16 Å and 2.86 Å for models from the Swiss-Model and Robetta web servers, respectively. The higher RMSD value for Robetta may have been contributed by the disordered variable region. Furthermore, model quality assessment was done by QMEAN *Z*-scores [36] using the Swiss-Model structural assessment tool [15] as presented in Table S3. The closer this value is to zero, the more the predicted model fits the experimental structure of a similar size. Thus, the model generated from the 2E9Q template was further used for geometry minimization.

#### Model geometry minimization, re-assessment, and validation

To improve the structure of the "glutelin type-B 5-like" protein, the two models following the initial quality assessment were subjected to geometry minimization using PHENIX [21] with 1,000 maximum iterations and five macrocycles. Geometry minimization relaxes the backbone bond angle by allowing full flexibility of the bond angle and bond lengths, thereby improving the energy landscape of the protein [37]. The models following geometry minimization are presented in Figure 2 and were re-evaluated using the Ramachandran plot, *Z*-scores, and QMEANDisCo scores as presented in Table 2. Although the *Z*-scores seem not to have improved, the bond angles improved significantly, as can be seen in the Ramachandran plots. The bond angles increase from 94.01% to 99.8% and from 93.4% to 98% for the Swiss-Model and Robetta model, respectively (Figure S1). Comparing the minimized structure to the unminimized ones showed that



Figure 2. Visualization of the three-dimensional structure of the "glutelin type-B 5-like" protein model geometrically minimized using PHENIX. (A) Model generated from the Swiss-Model server; (B) model generated from the Robetta web server. The color code points to each chain of the monomer. PHENIX: Python-based Hierarchical Environment for Integrated Xtallography

there were zero bad angles and bad bonds for both minimized models and 6 bad bonds, 97 bad angles, 3 bad bonds, and 15 bad angles for both unminimized models, respectively.

Model type	Ramachandran plot (%)			Z-scores	QMEANDisCo scores
	Allowed	Generally allowed	Disallowed	_	
Swiss-Model server	99.8	0.7	0	-5.44	0.62 ± 0.05
Robetta web server	98	1.4	0.5	-5.11	0.48 ± 0.05

Table 2. Summary of the "glutelin type-B 5-like" protein geometry minimized model's quality assessment scores

QMEANDisCo: Qualitative Model Energy Analysis with Distance Constraints

#### RMSD

Following model optimization, the model from the Swiss-Model server was used in the MD simulation. The RMSD is a measure of the changes in the conformation of the protein due to the application of external processing stresses such as temperature, pressure, and electric field. It compares the protein's new conformation to its native or initial conformation as shown by the equation:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left| r_{final}(i) - r_{initial}(i) \right|^2}$$
(1)

Where  $r_{final}$  is the final coordinates of atom *i*,  $r_{initial}$  is the initial coordinate of atom *i*, and *N* is the number of atoms. The average of the RMSD for the "glutelin type-B 5-like" protein after 1 ns simulation from different combinations of temperature and static electric field is presented in Table 3 and Figure 3. The average values of RMSD for the "glutelin type-B 5-like" protein ranged from 0.28 to 0.61 across temperatures of 300–400 K, looking at both data for static and zero electric field treatments. It can be observed that the RMSD increases with temperature (Figure 3) at all electric field levels and decreases with a static electric field at all temperature levels, except at 300 K. Both the temperature and static field have significant effect (P < 0.0001) and interaction effects on the RMSD of the protein (P < 0.001). Additionally, the RMSD was observed to increase with simulation time for all treatment combinations (Figure 3). Static electric field treatments caused a higher RMSD compared to zero electric field were less pronounced compared to higher temperature thermal treatment. This phenomenon has been reported by Singh and coworkers [1] where it showed a static electric field of 0.002 V/nm and 0.004 V/nm had no effects on the stability of the structure of soybean hydrophobic protein at 300 K. Vanga and co-workers [38] also

reiterated this observation when they revealed that both static (0.05 V/nm) and oscillating electric field (0.05 V/nm, 2,450 MHz) at 300 K had no significant effects on the RMSD of Ara h 6 peanut protein allergen.

Table 3. Average root mean square deviation (RMSD), the radius of gyration (Rg), solvent accessibility surface area (SASA), volume, and density of the "glutelin type-B 5-like" protein after 1 ns simulation under different temperature and electric field conditions

Treatments	RMSD (nm)	Rg (nm)	SASA (nm <sup>2</sup> )	Vol (nm³)	Density (g/L)
300 K, 0 V/nm	$0.28 \pm 0.05^{a1}$	3.25 ± 0.02 <sup>a1</sup>	495.70 ± 5.38 <sup>a1</sup>	167.02 ± 1.28 <sup>a12</sup>	829.07 ± 6.35 <sup>a12</sup>
300 K, 0.1 V/nm	$0.29 \pm 0.06^{b2}$	3.22 ± 0.01 <sup>b1</sup>	490.94 ± 3.40 <sup>b1</sup>	166.69 ± 1.12 <sup>b2</sup>	830.72 ± 5.56 <sup>b1</sup>
300 K, 1 V/nm	$0.32 \pm 0.08^{c13}$	3.21 ± 0.02 <sup>c1</sup>	503.71 ± 6.50 <sup>c1</sup>	167.56 ± 1.19 <sup>c2</sup>	826.44 ± 5.87 <sup>c1</sup>
300 K, 3 V/nm	$0.30 \pm 0.06^{d_{23}}$	$3.23 \pm 0.01^{d_3}$	492.81 ± 3.82 <sup>d1</sup>	166.28 ± 1.33 <sup>d1</sup>	832.79 ± 6.64 <sup>d12</sup>
350 K, 0 V/nm	$0.40 \pm 0.09^{a1}$	$3.23 \pm 0.02^{a1}$	$495.55 \pm 6.96^{a1}$	167.09 ± 1.54 <sup>a12</sup>	828.75 ± 7.62 <sup>a12</sup>
350 K, 0.1 V/nm	$0.46 \pm 0.10^{b2}$	$3.24 \pm 0.04^{b1}$	484.91 ± 17.21 <sup>b1</sup>	166.24 ± 2.07 <sup>b2</sup>	833.07 ± 10.41 <sup>b1</sup>
350 K, 1 V/nm	$0.39 \pm 0.09^{c13}$	$3.18 \pm 0.04^{c^2}$	471.00 ± 13.43 <sup>c1</sup>	164.72 ± 1.64 <sup>c2</sup>	840.71 ± 8.37 <sup>c1</sup>
350 K, 3 V/nm	$0.45 \pm 0.12^{d^{23}}$	$3.20 \pm 0.03^{d_3}$	476.08 ± 15.57 <sup>d1</sup>	$164.98 \pm 2.05^{d1}$	839.41 ± 10.39 <sup>d2</sup>
400 K, 0 V/nm	$0.54 \pm 0.12^{a1}$	$3.23 \pm 0.06^{a1}$	462.85 ± 26.15 <sup>a1</sup>	$163.65 \pm 3.27^{a12}$	846.48 ± 16.88 <sup>a12</sup>
400 K, 0.1 V/nm	$0.61 \pm 0.16^{b2}$	$3.22 \pm 0.04^{b1}$	479.99 ± 15.84 <sup>b1</sup>	165.33 ± 2.23 <sup>b2</sup>	837.69 ± 11.24 <sup>b1</sup>
400 K, 1 V/nm	0.57 ± 0.13 <sup>c13</sup>	$3.23 \pm 0.02^{c^3}$	485.62 ± 15.03 <sup>c1</sup>	166.18 ± 2.21 <sup>c2</sup>	833.38 ± 11.11 <sup>c1</sup>
400 K, 3 V/nm	$0.59 \pm 0.14^{d_{23}}$	$3.30 \pm 0.06^{d_3}$	483.40 ± 20.29 <sup>d1</sup>	165.57 ± 2.61 <sup>d1</sup>	836.52 ± 13.11 <sup>d32</sup>

For a given attribute and corresponding mean, values with different letter or number superscript are significantly different from temperature and PEF standpoint, respectively, while values with common letter or number superscript are not significantly different

#### RMSF

RMSF reveals information on the flexible and rigid regions of the protein structure. The RMSF of the protein on a per-residue basis at the end of the 1 ns simulation is presented in Figure 4. From Figure 4, it can be observed that higher fluctuations were recorded at a high temperature of 400 K and the least at 300 K. The fluctuations only increased with an increase in temperature, but not with an increase in static electric field. The N- and C-terminals of the proteins showed a high degree of fluctuation consistently with all the treatments for all the chains (A, B, and C). The regions around the two distant  $\alpha$ -helices (R60–A65 and A233–K240) of the protein showed a higher RMSF compared to the  $\beta$ -sheet regions (D103–H106, S122–L128, V132–A133, Y138–T140, S145–V159, A165–L171, Q175–V179, T18–R200, and T20–R209). The average fluctuation for all residues recorded for the treatments are 0.05 ± 0.02, 0.05 ± 0.03, 0.06 ± 0.03, 0.

#### Secondary structure analysis

The secondary structure of the protein molecule was analyzed using the inbuilt DSSP tool of GROMACS. The changes in the number of residues count that contribute to the secondary structure profile of the protein after 1 ns of different treatment levels are presented in Figure 5, while the percentage distribution of the secondary structures is presented in Table S4. It is expected that following the application of processing stresses that cause unfolding and denaturation of the protein, the unstable structures, such as coil, turn, and bend, should increase while the more stable structures, such as helices and sheets, should decrease in count. It was observed that the increase in the number of residues for coils and bends was more obvious at 400 K, with a corresponding decrease in the number of residues for  $\beta$ -sheets and  $\alpha$ -helices. There was a progressive increase in the number of residues for bend with temperature, and the effects of the static electric field were observed. It is possible that short-time simulation may not be enough to cause an irreversible unfolding or permanent denaturation of the protein's secondary structure, which may explain the reason for the absence of a progressive or regressive trend in the behavior of the secondary structures of the protein. It is worthy of note that these changes were not statistically significant at P > 0.05. The expectation as reported by Li et al. [39] was that high temperature treatment should result in the unfolding of glutelin conformation, which should cause a decrease in the  $\alpha$ -helix content and increase the  $\beta$ -sheet content.



Figure 3. Root mean square deviation (RMSD) of the "glutelin type-B 5-like" protein under different simulation conditions (temperature and static electric field) for 1 ns simulation time



Figure 4. Root mean square fluctuations (RMSF) of the "glutelin type-B 5-like" protein per residue under different simulation conditions (temperature and static electric field) after 1 ns simulation time

The snapshot of the molecule after simulation (Figure S2) did not show an obvious disruption to the secondary structure when compared to all treatments. However, the changes in the percentage distribution profile of the secondary structure (Table S4) showed that some of the residues that make up the secondary



Figure 5. The number of residues present in the secondary structure of the "glutelin type-B 5-like" protein molecules under different simulation conditions (temperature and static electric field) after 1 ns simulation time

structure were lost without necessarily disrupting the secondary structure conformation. In addition, the interchange between turns structure and helices was observed with simulation time (data not shown). For instance, we analyzed the secondary structures of the protein at 300 K temperature level with STRIDE [40] (data not shown) and observed that the  $\alpha$ -helix in the region of P114–L119 on chain A lost one residue to become a turn within region I115–L119 after treatment with a static electric field of 0.1 V/nm. No loss of residue was observed with treatments of 1 V/nm and 3 V/nm. Also, an  $\alpha$ -helix formed on chain B in the region V222–Y229 when a static electric field of 0.1 V/nm was applied to the protein molecule. This was disrupted to  $3_{10}$ -helices when the static electric field changed to 1 V/nm and 3 V/nm (both at region V222–V224), losing 3 residues as turns.

#### **Total dipole moment**

The dipole moment is a good measure of the behavior of a protein to the external electric field [6]. It gives information on the polarization response of the protein to the external electric field, and any shift in the total dipole moment is an indication of protein denaturation or transition in the protein secondary structures [6, 7]. The total dipole moment of the "glutelin type-B 5-like" protein frames with simulation time is presented in Figure 6. It can be observed that as early as 100 ps, there were already deviations from the non-electric field treatment for temperatures 300 K, 350 K, and 400 K, indicating an early transition in the secondary structure of the protein, which became more pronounced between 400 ps to 600 ps and much more between 800 ps to 1,000 ps. The static electric field of 0.1 V/nm was observed to show the least deviation at all temperature levels. Both temperature and static electric field have significant effects (P < 0.05) on the differences observed in the total dipole moment.

#### Rg

The Rg provides information on the spread of the atoms of the protein molecule with respect to its center of mass. This information is useful for understanding if the proteins compact into a globular state or unfold. The higher the Rg, the less compact the protein and vice-versa [1, 41]. The Rg is calculated from the equation below:

$$Rg = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left| r(i) - r_{centre} \right|^2}$$
<sup>(2)</sup>

Where r(i) is the coordinate of atom *i*,  $r_{centre}$  is the coordinate of the center of mass of the protein, and *N* is the number of atoms. The Rg was observed to reduce with the simulation time for all the treatments, which indicates that the protein compacts over time with electrical and heat treatment. Proso millet protein has been reported to compact via hydrophobic aggregation during thermal treatment, which leads to reduced protein digestibility [42]. Similar phenomena of a decrease in Rg with simulation time were



Figure 6. The total dipole moment of the "glutelin type-B 5-like" protein under different simulation conditions (temperature and static electric field) after 1 ns simulation time

reported for peanut allergen (Ara h 6 protein) [38], and the Rg of STI protein was also reported to decrease because of heating. The heating of protein in solution at high temperature ( $80^\circ$ -100°C for pea protein) has been known to cause new aggregation through hydrophobic interaction following the revelation of hydrophobic residue during heating [43]. This behavior was observed for all the treatments where there was an initial increase in the Rg between 0–200 ps and then followed by a downtrend in Rg for the rest of the simulation time (Figure 7). This could be explained by the initial unfolding of the protein, leading to a high Rg, followed by the new aggregation for the rest of the simulation time. The average Rg for all treatments is presented in Table 3. For zero electric field, the Rg decreased significantly (P < 0.05) with an increase in temperature, and though there were significant differences in Rg with the static electric field and the interaction of the two factors, close observation of the data in Table 3 does not show a steady decrease or increase in Rg with the static electric field.

#### SASA

The SASA is one of the surface-active properties of the protein that impacts its functionality, such as surface hydrophobicity, which in turn affects the protein solubility, emulsifying, and foaming ability [8, 44]. The hydrophobicity of protein is inversely related to its SASA [45]. The values of SASA at the end of 1 ns simulation are presented in Table 3, while the trends with simulation time are presented in Figure 8. From Figure 8, the SASA trends downward as the simulation progresses, confirming that the protein compacts over time, as seen by the trend observed with the Rg. Similar to the Rg, there was an initial increase in SASA



Figure 7. The radius of gyration (Rg) of the "glutelin type-B 5-like" protein under different simulation conditions (temperature and static electric field) for 1 ns simulation time

of about 0–100 ps, which may indicate the initial unfolding due to applied treatment, causing an increase in SASA. Later, a decrease in SASA with simulation time was observed, which may have been due to the compaction of the protein molecule caused by hydrophobic aggregation. Vanga et al. [38] reported a similar downward trend in SASA with simulation time for Ara h 6 peanut allergy as a result of temperature and electric field treatment. Increasing the temperature from 300 K to 400 K causes a decrease in the average value of SASA at 0 V/nm and 0.1 V/nm static electric field, but at 1 V/nm and 3 V/nm static electric field, a decrease was observed from 300 K to 350 K and increased again at 400 K. The changes observed due to temperature were significant at P < 0.0001, but changes observed due to the static electric field were not at P < 0.05.

#### Volume and density

The volume and density of the "glutelin type-B 5-like" protein were stable with simulation time at 300 K for all static electric field treatment levels considered, with slight variations. However, at a temperature of 350 K and 400 K, a downward trend in the volume with simulation time was observed for all static treatment levels (Figure 9). At all levels of static electric field treatment, the average volume was observed to decrease from 300 K to 400 K, while the density, as expected, increased in reverse order. This further confirms that the protein molecules compacted as the temperature was ramped up from 300 K to 400 K. The changes observed in volume were significant (P < 0.05) for temperature, static electric field, and their interaction.



Figure 8. Solvent accessibility surface area (SASA) of the "glutelin type-B 5-like" protein under different simulation conditions (temperature and static electric field) for 1 ns simulation time

### Discussion

In summary, the three-dimensional structure of the "glutelin type-B 5-like" protein, a type of glutelin protein from proso millet, was successfully predicted using the Swiss-Model server and the Robetta web server. The two models obtained from the servers showed a good agreement in their main backbone and showed conserved domains with the globulin group of proteins from peas, amaranth, rapeseed, and peanuts. Ramachandra plot of the protein structure gave over 98% of residues in the favorable region. In modeling the effects of temperature and static electric field three levels of static electric field on the behavior of the "glutelin type-B 5-like" protein in an MD simulation environment, the simulated stresses caused a significant increase in the RMSD (0.28 nm to 0.61 nm across temperatures of 300–400 K) and a decrease in both the Rg and SASA, implying that the "glutelin type-B 5-like" protein is compacting under these simulated conditions. While the stresses led to the perturbation of the protein's secondary structure. Longer MD simulations may be needed to cause the disruptions of the protein secondary structure, which can provide more information on how the surface-active properties (such as SASA and Rg) may be affected. The temperature had significant effects on the changes observed for all parameters measured (RMSD, SASA, and Rg), while the effects of the static electric field were insignificant for SASA.



Figure 9. Volume changes were observed for the "glutelin type-B 5-like" protein under different simulation conditions (temperature and static electric field) for 1 ns simulation time

## Abbreviations

AA: amino acid

DSSP: dictionary of the secondary structure of protein

E-value: expected value

GROMACS: Groningen Machine for Chemical Structure

MD: molecular dynamic

MPC: millet protein concentrate

PDB: Protein Data Bank

PHENIX: Python-based Hierarchical Environment for Integrated Xtallography

PMDB: Protein Model Database

**QMEAN: Qualitative Model Energy Analysis** 

QMEANDisCo: Qualitative Model Energy Analysis with Distance Constraints

Rg: radius of gyration

RMSD: root mean square deviation

RMSF: root mean square fluctuations

SASA: solvent accessibility surface area

#### STI: soybean trypsin inhibitor

Explor Foods Foodomics. 2025;3:101088 | https://doi.org/10.37349/eff.2025.101088

## **Supplementary materials**

The supplementary tables for this article are available at: https://www.explorationpub.com/uploads/ Article/file/101088\_sup\_1.pdf. The supplementary figures for this article are available at: https://www. explorationpub.com/uploads/Article/file/101088\_sup\_2.pdf.

## **Declarations**

### Author contributions

FA: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing—original draft. AS and KVK: Methodology, Formal analysis, Investigation. AA: Conceptualization, Investigation, Methodology, Supervision, Writing—review & editing. All authors read and approved the submitted version.

### **Conflicts of interest**

The authors declare that they have no conflicts of interest.

Ethical approval

Not applicable.

**Consent to participate** 

Not applicable.

### **Consent to publication**

Not applicable.

### Availability of data and materials

The raw 3D structure of the glutelin type-B 5-like protein used can be obtained from the protein model database (PMDB; http://srv00.recas.ba.infn.it/PMDB/main.php) with PMDB identifier PM0083241. All other data will be made available by the authors to any qualified researcher.

### Funding

This work was supported by the Kentucky Agricultural Experiment Station (KAES), the National Institute of Food and Agriculture (NIFA), U.S. Department of Agriculture, Hatch-Multistate project # [1024529]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Copyright

© The Author(s) 2025.

## **Publisher's note**

Open Exploration maintains a neutral stance on jurisdictional claims in published institutional affiliations and maps. All opinions expressed in this article are the personal views of the author(s) and do not represent the stance of the editorial team or the publisher.

## References

- 1. Singh A, Orsat V, Raghavan V. Soybean Hydrophobic Protein Response to External Electric Field: A Molecular Modeling Approach. Biomolecules. 2013;3:168–79. [DOI] [PubMed] [PMC]
- Tandang-Silvas MR, Cabanos CS, Carrazco Peña LD, De la Rosa AP, Osuna-Castro JA, Utsumi S, et al. Crystal structure of a major seed storage protein, 11S proglobulin, from *Amaranthus hypochondriacus*: Insight into its physico-chemical properties. Food Chem. 2012;135:819–26. [DOI] [PubMed]

- 3. Pantoja-Uceda D, Shewry PR, Bruix M, Tatham AS, Santoro J, Rico M. Solution Structure of a Methionine-Rich 2S Albumin from Sunflower Seeds: Relationship to Its Allergenic and Emulsifying Properties. Biochemistry. 2004;43:6976–86. [DOI] [PubMed]
- 4. Fukuda T, Maruyama N, Salleh MRM, Mikami B, Utsumi S. Characterization and Crystallography of Recombinant 7S Globulins of Adzuki Bean and Structure–Function Relationships with 7S Globulins of Various Crops. J Agric Food Chem. 2008;56:4145–53. [DOI] [PubMed]
- 5. Vagadia BH, Vanga SK, Singh A, Raghavan V. Effects of thermal and electric fields on soybean trypsin inhibitor protein: A molecular modelling study. Innov Food Sci Emerg Technol. 2016;35:9–20. [DOI]
- 6. Marracino P, Apollonio F, Liberti M, d'Inzeo G, Amadei A. Effect of High Exogenous Electric Pulses on Protein Conformation: Myoglobin as a Case Study. J Phys Chem B. 2013;117:2273–9. [DOI] [PubMed]
- 7. Wang X, Li Y, He X, Chen S, Zhang JZH. Effect of Strong Electric Field on the Conformational Integrity of Insulin. J Phys Chem A. 2014;118:8942–52. [DOI] [PubMed]
- 8. Akharume F, Santra D, Adedeji A. Physicochemical and functional properties of proso millet storage protein fractions. Food Hydrocolloid. 2019;108:105497. [DOI]
- 9. Wouters AGB. Delcour JA. Cereal protein-based nanoparticles as agents stabilizing air-water and oilwater interfaces in food systems. Curr Opin Food Sci. 2019;25:19–27. [DOI]
- Akharume FU, Korotkov KK, Adedeji AA. Characterization of Recombinant Glutelin type-B 5-like Protein from Proso Millet. 2020 ASABE Annual International Virtual Meeting; 2020 July 13–15. Michigan: American Society of Agricultural and Biological Engineers; 2020. pp. 2000272. [DOI]
- Akharume F, Adedeji A. Molecular dynamic (*in silico*) modeling of structure-function of glutelin type-B
   5-like from proso millet storage protein: effects of temperature and pressure. J Food Sci Technol.
   2023;60:114–22. [DOI] [PubMed] [PMC]
- Kuroda Y, Suenaga A, Sato Y, Kosuda S, Taiji M. All-atom molecular dynamics analysis of multi-peptide systems reproduces peptide solubility in line with experimental observations. Sci Rep. 2016;6:19479.
   [DOI] [PubMed] [PMC]
- 13. Shi J, Ma X, Zhang J, Zhou Y, Liu M, Huang L, et al. Chromosome conformation capture resolved near complete genome assembly of broomcorn millet. Nat Commun. 2019;10:464. [DOI] [PubMed] [PMC]
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46:W296–303.
   [DOI] [PubMed] [PMC]
- Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis. 2009;30:S162–73. [DOI] [PubMed]
- 16. Song Y, DiMaio F, Wang RY, Kim D, Miles C, Brunette T, et al. High-Resolution Comparative Modeling with RosettaCM. Structure. 2013;21:1735–42. [DOI] [PubMed] [PMC]
- 17. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. 2004;32:W526–31. [DOI] [PubMed] [PMC]
- 18. Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res. 2012;40:W294–7. [DOI] [PubMed] [PMC]
- 19. Shin WH, Lee GR, Heo L, Lee H, Seok C. Prediction of Protein Structure and Interaction by GALAXY Protein Modeling Programs. Bio Design. 2014;2:1–11. Korean.
- 20. Li Z, Natarajan P, Ye Y, Hrabe T, Godzik A. POSA: a user-driven, interactive multiple protein structure alignment server. Nucleic Acids Res. 2014;42:W240–5. [DOI] [PubMed] [PMC]
- Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. Acta Crystallogr D Struct Biol. 2019;75:861–77. [DOI] [PubMed] [PMC]
- 22. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. *Procheck*: a program to check the stereochemical quality of protein structures. J Appl Crystallogry. 1993;26:283–91. [DOI]

- 23. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in threedimensional structures of proteins. Nucleic Acids Res. 2007;35:W407–10. [DOI] [PubMed] [PMC]
- 24. Sippl MJ. Recognition of errors in three-dimensional structures of proteins. Proteins. 1993;17:355–62. [DOI] [PubMed]
- 25. Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T. QMEANDisCo—distance constraints applied on model quality estimation. Bioinformatics. 2020;36:1765–71. [DOI] [PubMed] [PMC]
- 26. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. J Comput Chem. 2005;26:1701–18. [DOI] [PubMed]
- 27. Robertson MJ, Tirado-Rives J, Jorgensen WL. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. J Chem Theory Comput. 2015;11:3499–509. [DOI] [PubMed] [PMC]
- 28. Kusalik PG, Svishchev IM. The Spatial Structure in Liquid Water. Science. 1994;265:1219–21. [DOI] [PubMed]
- 29. Parrinello M, Rahman A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. Phys Rev Lett. 1980;45:1196. [DOI]
- 30. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys. 1984;81:3684–90. [DOI]
- 31. Touw WG, Baakman C, Black J, te Beek TA, Krieger E, Joosten RP, et al. A series of PDB-related databanks for everyday needs. Nucleic Acids Res. 2015;43:D364–8. [DOI] [PubMed] [PMC]
- 32. Kabsch W, Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogenbonded and geometrical features. Biopolymers. 1983;22:2577–637. [DOI] [PubMed]
- Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. 1996;14:33–8. [DOI] [PubMed]
- Tandang-Silvas MRG, Fukuda T, Fukuda C, Prak K, Cabanos C, Kimura A, et al. Conservation and divergence on plant seed 11S globulins based on crystal structures. Biochim Biophys Acta. 2010;1804: 1432–42. [DOI] [PubMed]
- 35. Adachi M, Takenaka Y, Gidamis AB, Mikami B, Utsumi S. Crystal Structure of Soybean Proglycinin A1aB1b Homotrimer. J Mol Biol. 2001;305:291–305. [DOI] [PubMed]
- 36. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics. 2011;27:343–50. [DOI] [PubMed] [PMC]
- 37. Conway P, Tyka MD, DiMaio F, Konerding DE, Baker D. Relaxation of backbone bond geometry improves protein energy landscape modeling. Protein Sci. 2014;23:47–55. [DOI] [PubMed] [PMC]
- 38. Vanga SK, Singh A, and Raghavan V. Effect of thermal and electric field treatment on the conformation of Ara h 6 peanut protein allergen. Innov Food Sci Emerg. 2015;30:79–88. [DOI]
- 39. Li Y, Zhang C, Liang Y, Wang L, Xiong W. Solubility and conformational characterization of rice glutelin after high temperature treatment. Int J Biol Macromol. 2022;223:1720–6. [DOI] [PubMed]
- 40. Heinig M, Frishman D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res. 2004;32:W500–2. [DOI] [PubMed] [PMC]
- 41. Fenwick S, Vanga SK, DiNardo A, Wang J, Raghavan V, Singh A. Computational evaluation of the effect of processing on the trypsin and alpha-amylase inhibitor from Ragi (*Eleusine coracana*) seed. Eng Rep. 2019;1:e12064. [DOI]
- 42. Gulati P, Li A, Holding D, Santra D, Zhang Y, Rose DJ. Heating Reduces Proso Millet Protein Digestibility via Formation of Hydrophobic Aggregates. J Agric Food Chem. 2017;65:1952–9. [DOI] [PubMed]
- 43. Chao D, Aluko RE. Modification of the structural, emulsifying, and foaming properties of an isolated pea protein by thermal pretreatment. J Food Sci. 2018;6:357–66.
- 44. Nakai S. Structure-function relationships of food proteins: with an emphasis on the importance of protein hydrophobicity. J Agric Food Chem. 1983;31:676–83. [DOI]
- 45. Gromiha MM, Nagarajan R, Selvaraj S. Protein Structural Bioinformatics: An Overview. Encyclopedia Bioinform Comput Biol. 2019;2:445–9. [DOI]