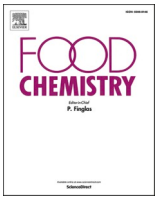
Food Chemistry 497 (2025) 147044

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem



Solvent-dependent structural and dynamic behavior of α-gliadin in aqueous ethanol systems: Insights from molecular dynamics simulations

Nandan Kumar, Yonghui Li *

Department of Grain Science and Industry, Kansas State University, Manhattan, KS 66506, USA

ARTICLE INFO

Keywords:
α-Gliadin
Gluten
Structure prediction
Molecular dynamics simulation
Solvent behavior
Protein conformation

ABSTRACT

Understanding the behavior of gliadin in mixed aqueous-organic solvents is vital for food material science. This study integrates advanced structure prediction with microsecond-scale molecular dynamics (MD) simulations to explore the conformational and solvation behavior of α-gliadin across ethanol-water mixtures. Our findings reveal that α-gliadin maintains structural stability across all solvent conditions, with a progressive increase in molecular expansion at higher ethanol concentrations. Ethanol promotes up to 30 % more salt bridges, enhances helical content, and reduces β-sheet formation. Residue-level flexibility analysis reveals that specific segments, such as Gln227-Ala261, Tyr121-Gln136, and the proline- and glutamine-rich region Pro64-Leu78, exhibit higher fluctuation, suggesting their sensitivity to solvent-induced conformational changes, whereas domains stabilized by electrostatic and covalent interactions remain rigid (show less fluctuation). Thermodynamic and PCA analyses highlight stronger protein-solvent interactions and greater conformational diversity at higher ethanol concentrations. This work provides a detailed molecular-level quantification of how ethanol modulates α-gliadin structure and solvation.

1. Introduction

Wheat gluten proteins are essential for the functional properties of dough, imparting the viscoelasticity and extensibility that define the quality of baked goods and other wheat-based products (Shewry et al., 2002). Gluten consists of two major fractions, gliadins and glutenins. Among these, α-gliadin, a monomeric protein soluble in aqueous ethanol and rich in glutamine and proline residues, is a key component influencing behavior of the gluten in various solvent environments, such as ethanol-water mixtures, commonly encountered in gluten extraction and processing (Wieser, 2007). The high solubility of gliadins, particularly α-gliadin, in ethanol-water mixtures is a distinguishing biochemical property that underpins their selective extraction from wheat flour and their functional role in food processing (Tatham & Shewry, 1985). This selective solubility not only facilitates the separation of gliadins from glutenins but also highlights the importance of ethanol-water environments in modulating gliadin structure and function. These solvent conditions significantly influence the structural properties of gluten, affecting its ability to form a cohesive network during dough development, which is vital for applications in baking and food formulation (Delcour et al., 2012; Urade et al., 2018). Beyond its role in food systems, α-gliadin has biomedical significance, as its glutamine-rich regions are primary antigens in celiac disease, a chronic autoimmune disorder affecting approximately 1 % of the global population (Fasano & Catassi, 2012; Sollid, 2002). Thus, understanding how ethanol-water mixtures modulate the structure and dynamics of α-gliadin is critical for both optimizing food processing and mitigating health risks associated with gluten consumption.

The effect of ethanol-water mixtures on gluten proteins has been studied extensively through experimental methods, which have demonstrated that ethanol disrupts intra-protein interactions, leading to unfolding, increased solvent exposure, and changes in secondary structure (Tatham & Shewry, 1985). For instance, ethanol has been shown to reduce hydrogen bonding within the protein while enhancing hydrophobic interactions, resulting in a more extended conformation that can alter solubility and aggregation properties of gluten (Shewry & Belton, 2024). However, experimental techniques such as circular dichroism and infrared spectroscopy offer limited atomic-level insight into the interplay between protein-solvent interactions, conformational dynamics, and thermodynamic stability (Delcour et al., 2012). Moreover, due to the lack of an experimentally determined full-length structure for food proteins, the atomistic understanding of their behavior in different

* Corresponding author.
E-mail address: yonghui@ksu.edu (Y. Li).

https://doi.org/10.1016/j.foodchem.2025.147044
Received 26 June 2025; Received in revised form 5 November 2025; Accepted 8 November 2025
Available online 9 November 2025

solvent environments remains incomplete.

To address these limitations, computational modeling, particularly MD simulations, has become an increasingly powerful tool for studying food proteins and peptides at atomic resolution (Kumar, Du, Amachawadi, et al., 2025; Singh et al., 2018; Yu, Li, et al., 2023). MD studies have been successfully applied to elucidate the solvent effects on protein folding and dynamics, revealing how co-solvents like ethanol modulate hydrogen bonding networks, solvation shells, and conformational ensembles (Gazi et al., 2023; Nassar et al., 2022). Recent advances in computational methods, including enhanced sampling techniques, machine learning, and integrative modeling, have enabled detailed investigations into the folding, stability, and functional properties of food proteins and peptides under various processing conditions (Kumar, Du, Amachawadi, et al., 2025; Kumar, Du, & Li, 2025; Kumar & Sastry, 2021; Yang et al., 2024; Yang, Kumar, Kuang, Li, & Song, 2025; Yang, Kumar, Kuang, Song, & Li, 2025). In this study, we employ MD simulations to investigate the structural, dynamic, and thermodynamic responses of α-gliadin in ethanol-water mixtures (0 %, 30 %, 70 %, and 100 % *v*/v ethanol). Using advanced computational techniques including Free Energy Landscape (FEL) analysis, Principal Component Analysis (PCA), Minimum-Distance Distribution Functions (MDDFs), and Kirkwood-Buff integrals, we characterize the conformational changes, solvent interactions, and stability of the protein across these conditions. Our comprehensive approach provides new molecular-level insights into how ethanol drives structural and conformational changes in α-gliadin. These findings have direct implications for food material science, informing how solvent conditions influence gluten functionality during processing, and also shed light on the exposure of glutamine-rich epitopes relevant to celiac disease immunogenicity. To our knowledge, this study provides the first integrated analysis combining advanced structure prediction with microsecond-scale MD simulations to investigate full-length α-gliadin across graded ethanol–water environments. By linking atomic-level solvation patterns, conformational energetics, and domain-specific flexibility, this work offers a deeper mechanistic understanding of the behavior of gliadin under processing-relevant solvent conditions that has not been reported previously.

2. Computational details

2.1. Structure prediction of α-gliadin

The study utilized the amino acid sequence of α-gliadin protein (accession number P18573) from wheat (*Triticum aestivum*), obtained from UniProt (https://www.uniprot.org/). The sequence and amino acid composition of α-gliadin are shown in Table S1. To predict the secondary structure of α-gliadin, established bioinformatics tools were employed. Specifically, the NPS@ server was used (Combet et al., 2000), which integrates algorithms like Garnier-Osguthorpe-Robson (GOR) III (Gibrat et al., 1987) and Ψ-Prediction of Secondary Structure (PHD) (Rost & Sander, 1993) to assign probabilities for each residue belonging to α-helices, β-strands, or random coils of α-gliadin. For tertiary structure prediction, a comprehensive suite of methods was employed, including deep learning-based methods (AlphaFold, DeepFold, ESMFold, OmegaFold, and Robetta), template-based methods (I-TASSER and C-I-TASSER) and a method combines both neural network-based methods and template-based approach (D-I-TASSER). Each method is accompanied by a link to access relevant information or resources, as provided in Table S2. AlphaFold, a deep learning-powered tool, was used for *ab initio* structure prediction of α-gliadin in the absence of homologous templates (Jumper et al., 2021). Concurrently, Robetta was also employed to generate alternative models for comparison, leveraging a different structure prediction approach (Baek et al., 2021). In addition, ESMFold (Lin et al., 2023) and OmegaFold (Wu et al., 2022) were also employed for rapid protein structure predictions leveraging transformer-based architectures. These comprehensive approaches ensured robust and reliable structural predictions for the full-length α-gliadin protein. For template-based modeling, I-TASSER was used to identify templates from the PDB and construct full-length atomic models through iterative template-based fragment assembly simulations (Yang & Zhang, 2015). Furthermore, C-I-TASSER (Zheng et al., 2021) and D-I-TASSER (Li et al., 2021) tools from the I-TASSER suite, along with DeepFold (Pearce et al., 2022) were used for potential model refinement. These tools incorporate additional information like inter-residue contacts and optimize the initial models for enhanced accuracy. This comprehensive strategy ensured the generation of reliable and structurally sound 3D models for the full-length α-gliadin protein.

2.2. Initial MD simulation for structural refinement

To refine the modeled structures into more compact and energetically favorable conformations, each of the α-gliadin models underwent a 100 ns MD simulation using GROMACS 2023.3 (Abraham et al., 2015). These simulations were conducted in a water solvent environment, employing the SPC/E water model and the Amber99sb-ILDN force field (Lindorff-Larsen et al., 2010) for the protein. Each system was placed in a cubic box with a minimum distance of 1.0 nm between the protein and the box edges. The system was then neutralized with counterions and energy-minimized using the steepest descent algorithm. Equilibration was carried out in two stages: first in the NVT ensemble for 100 ps at 300 Kelvin using the V-rescale thermostat, followed by 100 ps in the NPT ensemble at 1 bar using the Parrinello-Rahman barostat. The production simulations were run for 100 ns with a 2 fs time step, saving compressed coordinates every 10 ps. Long-range electrostatic interactions were handled using the Particle Mesh Ewald (PME) method with a 1.0 nm cutoff. Van der Waals interactions were cut off at 1.0 nm, with dispersion correction applied for energy and pressure. Bonds involving hydrogen atoms were constrained using the LINCS algorithm.

2.3. Model evaluation and selection

After 100 ns MD simulations, the final structure of each model were evaluated using multiple metrics to select the best model for further analysis. The Root Mean Square Deviation (RMSD) and radius of gyration (Rg) were computed to assess structural stability and compactness. Structural quality was further evaluated using ERRAT (Colovos & Yeates, 1993) on the SAVES server. ERRAT analyzes the statistics of nonbonded interactions between atoms within the protein structure and compares them to a database of reliable, high-resolution structures. This analysis helps identify regions of the model with potential errors. Furthermore, PROCHECK was used for Ramachandran plot analysis to assess the phi (ϕ) and psi (ψ) torsion angles of the protein backbone (Laskowski et al., 1993). These angles define the allowed conformations for amino acid residues within a protein structure. Deviations from the expected Ramachandran plot regions can indicate structural errors or strained conformations. Additionally, ProSA-web (Wiederstein & Sippl, 2007) was used for structural evaluations. ProSA-web employs a potential energy function to assess overall quality and stability of predicted structures. Furthermore, we evaluated overall model quality using an array of structural validation metrics, including MolProbity (global structural quality), Clash (steric hindrance), Ramachandran plot (backbone geometry), and QMEANDisCo Global score (consensus quality). These metrics provide a comprehensive assessment of the accuracy and reliability of predicted models.

2.4. Extended MD simulation in ethanol-water mixtures

The selected model was subjected to extended MD simulations in four ethanol-water solvent mixtures with ethanol concentrations of 0 %, 30 %, 70 %, and 100 % (*v*/v) to investigate solvent effects on the structure of protein. The number of water and ethanol molecules, as well as the total number of atoms and simulation box volume for each system, are provided in Table S3. Here, "XEtOH" denotes the mole fraction of

ethanol in the simulation box, with the corresponding percent volume/volume (v/v) ethanol concentration provided in parentheses for clarity. For each solvent condition, two independent replicas were run to ensure statistical reliability, resulting in a total of eight simulations. Simulations were performed using GROMACS 2023.3 for 1000 ns each, with a 2-fs time step, under the NPT ensemble at 300 K and 1 bar pressure. The Amber99sb-ILDN force field was used for the protein, the SPC/E model for water, and ethanol molecules were parameterized using the CHARMM General Force Field (CGenFF) via the acpype tool (v2022.7.21). The simulation setup, including system preparation, energy minimization, equilibration (100 ps NVT followed by 100 ps NPT), thermostat (V-rescale), barostat (Parrinello-Rahman), PME electrostatics (1.0 nm cutoff), van der Waals interactions (1.0 nm cutoff with dispersion correction), and bond constraints (LINCS, all bonds), was identical to the initial 100 ns simulations unless otherwise specified. Coordinates and energies were saved every 2 ps for analysis, providing higher temporal resolution for the extended simulations.

2.5. Analysis of structural dynamics and solvent interactions

Structural dynamics and solvent interactions of α-gliadin were analyzed using GROMACS, with custom Python scripts (MDAnalysis, Seaborn, Matplotlib) for data processing and visualization. Conformational stability was assessed by calculating the RMSD for two replicas per solvent condition. Protein compactness was evaluated by computing Rg over time across replicas. Secondary structure evolution (α-helices, β-sheets) was tracked using the DSSP algorithm. SASA was calculated, and mean values across replicas were analyzed for each ethanol concentration to determine overall solvent effects. Hydrogen bonding with water and ethanol was analyzed using gmx hbonds. RMSD distributions were visualized with histograms and kernel density estimation (KDE) to capture conformational heterogeneity within individual replicas (y-axis: probability). SASA distributions were visualized using KDE only to highlight the averaged solvent effect across replicas (y-axis: probability density). To compute the minimum distance distribution function (MDFF) and Kirkwood-Buff Integrals, we utilized the ComplexMixtures.jl package (Martínez, 2022). PCA was performed on C-α atoms of the MD trajectory using GROMACS. The covariance matrix of atomic positional fluctuations was calculated with gmx covar, and eigenvalue decomposition was carried out using gmx anaeig to obtain the principal components. The first two principal components (PC1 and PC2) were selected as reaction coordinates for further analysis. The MD trajectory was projected onto these components, and the resulting distributions were used to construct a two-dimensional free energy landscape with the gmx sham utility, estimating the Gibbs free energy as a function of PC1 and PC2. The minima on the FEL, corresponding to the most stable conformational states, were identified by locating the lowest free energy values, and the time frames closest to these minima were mapped back to the trajectory for structural characterization.

3. Results and discussion

3.1. Structure prediction and initial refinement of α-gliadin models

Due to the lack of an experimentally determined full-length structure for α-gliadin, its structure was predicted using various deep learning-based, template-based, or hybrid methods as described in Section 2.1. The templates selected by I-TASSER and C-I-TASSER are provided in Table S4. The predicted structures (Fig. S1) showed that deep learning-based methods (AlphaFold, DeepFold, ESMFold, OmegaFold) produced more extended conformations compared to template-based methods (I-TASSER, C-I-TASSER), while the hybrid D-I-TASSER method displayed intermediate characteristics. All predicted structures exhibited a combination of helical and coiled regions, which is consistent with prior experimental studies reporting 30–35 % α-helix content in α-gliadin (Tatham & Shewry, 1985; Urade et al., 2018) as well as in the secondary structure prediction from sequence (Fig. S2). However, none of the models captured the experimentally reported β-strand content (Tatham & Shewry, 1985; Urade et al., 2018), likely due to limitations in deep learning training datasets or biases in template selection. To refine these models, 100 ns MD simulations were performed, resulting in more compact structures (Fig. S3). Comparisons of pre- and post-simulation conformations revealed that template-based models retained greater structural integrity than deep learning-based models, suggesting that evolutionary constraints from template selection enhance stability during dynamics.

3.1.1. Conformational stability and compactness during initial refinement

The compactness and stability of the refined models were assessed using Rg and RMSD over the 100 ns simulations conducted in pure water, which serves as a standard reference environment for evaluating protein structural properties *in silico* (Lindorff-Larsen et al., 2010). Rg measures the compactness of protein, while RMSD indicates structural stability over time. Fig. 1A shows Rg for each model, with error bars representing variability across simulation time. C-I-TASSER exhibited the lowest Rg at 2.02 ± 0.02 nm, indicating a highly compact structure with minimal fluctuations. D-I-TASSER followed closely with an Rg of 2.14 ± 0.03 nm. I-TASSER and ESMFold showed moderate compactness with Rg values of 2.54 ± 0.05 nm and 2.63 ± 0.24 nm, respectively. OmegaFold, AlphaFold, and DeepFold had higher Rg values of 3.01 ± 0.29 nm, 3.16 ± 0.32 nm, and 3.19 ± 0.40 nm, respectively, while Robetta exhibited the highest Rg at 3.95 ± 0.53 nm, reflecting less compact structure with greater variability. Fig. 1B shows the RMSD, with C-I-TASSER and D-I-TASSER maintaining the lowest RMSD at 0.62 ± 0.08 nm and 0.81 ± 0.07 nm, indicating high stability. I-TASSER and ESMFold had RMSD values of 0.93 ± 0.12 nm and 1.67 ± 0.21 nm, respectively, suggesting moderate stability. OmegaFold, DeepFold, and AlphaFold showed higher RMSD values of 2.30 ± 0.34 nm, 2.09 ± 0.49 nm, and 2.98 ± 0.46 nm, respectively, indicating greater conformational flexibility. Robetta had the highest RMSD at 3.30 ± 0.67 nm, reflecting significant structural rearrangement. These trends suggest that template-based and hybrid methods (C-I-TASSER, D-I-TASSER, I-TASSER) produce more stable and compact models, likely due to their reliance on experimentally validated templates, while deep learning methods (AlphaFold, DeepFold, OmegaFold, Robetta) predict more flexible structures that undergo larger conformational changes during MD refinement. The variability in Rg and RMSD across methods highlights the challenge of modeling food proteins like α-gliadin, where flexibility is a functional trait but can lead to overprediction of extended conformations in deep learning models. These trends align with experimental observations of α-gliadins in 0.01 M NaCl (pH 4), which reported a low intrinsic viscosity (4.0 mL g^{-1}) and a hydrodynamic radius of ~2.0 nm, consistent with compact conformations (Cole et al., 1984). C-I-TASSER, D-I-TASSER, ESMFold, and I-TASSER models best captured this compactness, with Rg values closest to the experimentally inferred hydrodynamic radius (Cole et al., 1984). In contrast, AlphaFold, DeepFold, OmegaFold, and Robetta overestimated the size of the protein, potentially due to their tendency to predict extended conformations that relax more significantly during MD simulations.

3.1.2. Structural quality assessment of predicted models

The structural quality of the refined models was thoroughly assessed using multiple metrics, as summarized in Table 1. Among AI-based methods, ESMFold demonstrated highest overall quality, achieving a MolProbity score of 1.26, a Clash Score of 0.00, an ERRAT score of 66.31, a ProSA *Z*-score of −7.17, a QMEANDisCo Global score of 0.54 ± 0.05, and 91.61 % of residues in favored Ramachandran regions (Table 1 and Fig. S4). AlphaFold followed closely, with a MolProbity score of 1.28, a Clash Score of 0.82, an ERRAT score of 63.19, a ProSA Z-score of −5.83, and 91.97 % of residues in favored Ramachandran regions. DeepFold, while exhibiting minimal steric clashes (Clash Score: 0.62), showed lower reliability with an ERRAT score of 49.10. Among

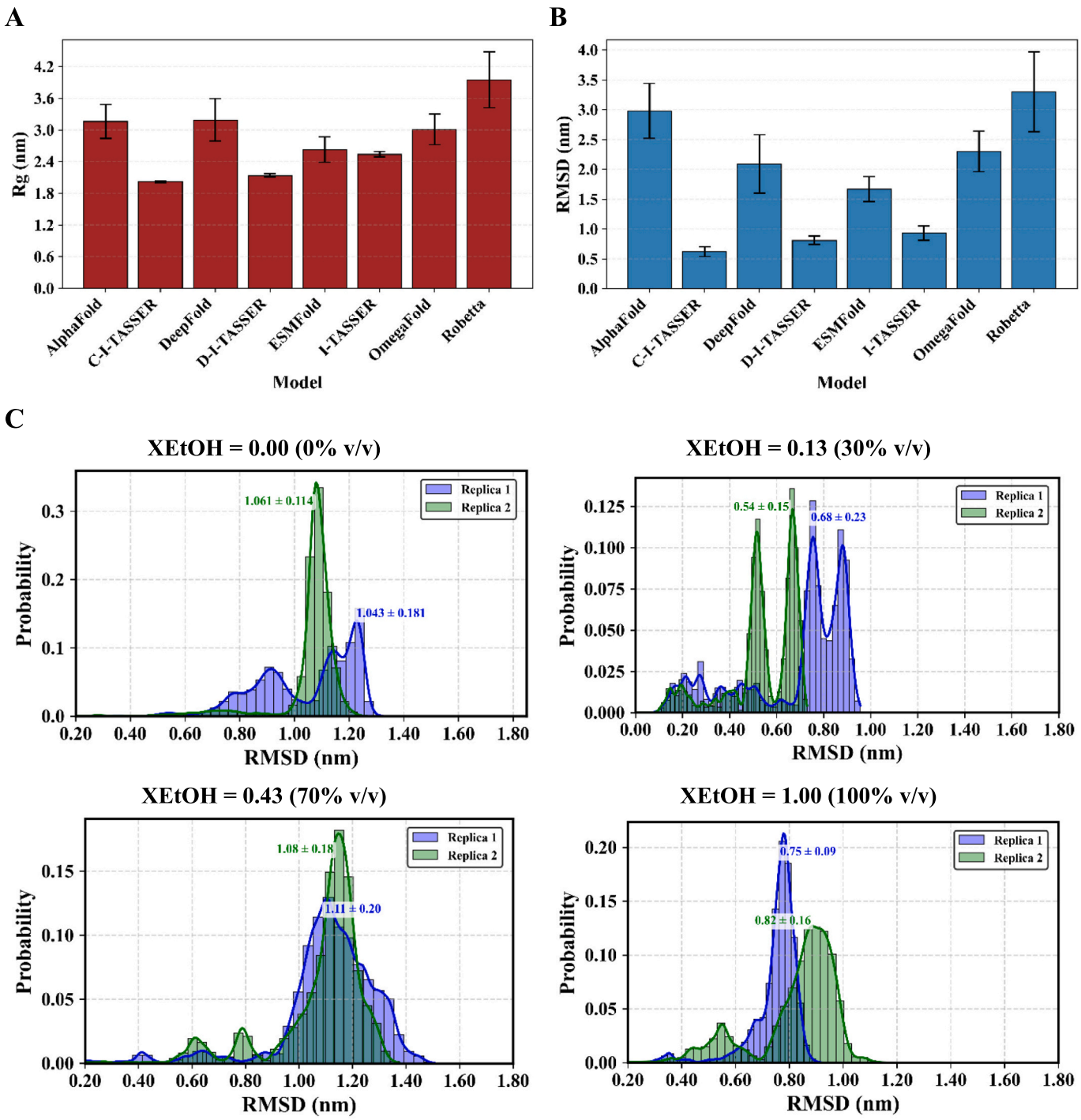


Fig. 1. Structural assessment of α-gliadin using multiple prediction models and molecular dynamics simulations. (A) Rg and (B) RMSD of α-gliadin predicted by eight structure prediction tools. Bars represent mean values (in nm), with error bars indicating standard deviations over 100 ns MD simulations. (C) Probability distributions of RMSD for α-gliadin in ethanol–water mixtures (0 %, 30 %, 70 %, and 100 % *v*/v EtOH), based on 1 μs MD simulations in two independent replicas. Annotated mean RMSD values ± standard deviations highlight the conformational stability and solvent-dependent structural variability of α-gliadin. "XEtOH" indicates the ethanol mole fraction; values in parentheses represent ethanol concentration as % v/v.

template-based approaches, C-I-TASSER performed best, with a MolProbity score of 1.81, a Clash Score of 1.03, an ERRAT score of 70.40, a ProSA Z-score of −5.41, and 85.40 % of residues in favored Ramachandran regions. D-I-TASSER, despite a high ERRAT score (74.39), had notable steric clashes (Clash Score: 1.44). Robetta achieved highest ERRAT score (76.39) among all methods but also exhibited moderate clashes (Clash Score: 1.24).

Although the C-I-TASSER, D-I-TASSER, and Robetta models exhibited a higher ERRAT score than ESMFold, and other models indicating favorable non-bonded interactions, the overall structural integrity of the ESMFold model was superior when evaluated across multiple complementary metrics. Specifically, ESMFold achieved the lowest MolProbity score, zero steric clashes, and the highest QMEANDisCo score, reflecting excellent stereochemical quality and global consistency. In contrast, the slightly higher ERRAT score of C-I-TASSER, D-I-TASSER, and Robetta models was balanced by its moderate clash score

Table 1
Assessment of structural quality in α-gliadin protein models through molprobity, clash, ERRAT, ProSA, Ramachandran favored, Ramachandran outliers, and QMEANDisCo global scores.

Prediction Methods	MolProbity Score	Clash Score	ERRAT	ProSA	Ramachandran		QMEANDisCo Global
					Favored	Outliers	
AlphaFold	1.28	0.82	63.19	−5.83	91.97	0.73	0.32 ± 0.05
C-I-TASSER	1.81	1.03	70.40	−5.41	85.40	2.92	0.35 ± 0.05
DeepFold	1.58	0.62	49.10	−5.56	87.23	0.73	0.34 ± 0.05
D-I-TASSER	2.04	1.44	74.39	−5.70	82.85	4.01	0.33 ± 0.05
ESMFold	1.26	0.00	66.31	−7.17	91.61	0.73	0.54 ± 0.05
I-TASSER	2.09	1.65	69.58	−4.27	83.21	4.01	0.29 ± 0.05
OmegaFold	1.54	0.21	67.90	−5.90	89.42	0.73	0.32 ± 0.05
Robetta	1.72	1.24	76.39	−6.69	89.42	1.09	0.36 ± 0.05

and lower QMEANDisCo value, suggesting local inconsistencies despite reliable non-bonded interactions. These results emphasize that no single quality metric fully determines model reliability; rather, combined evaluation across metrics provides a more accurate assessment.

Considering Rg, RMSD, and overall structural quality metrics, ESMFold (AI-based) and C-I-TASSER (template-based) emerged as top performers, balancing compactness, stability, and structural reliability. Zero Clash Score and high QMEANDisCo score of ESMFold model indicate a highly refined structure, while high ERRAT score and reasonable stereochemistry of the C-I-TASSER model make it the most reliable among template-based methods. In contrast, low ERRAT score of DeepFold model suggests potential structural inaccuracies despite minimal clashes, and high ERRAT score of the Robetta model is offset by the presence of steric clashes. Therefore, the ESMFold model was selected for extended simulations due to its lowest MolProbity score, absence of steric clashes, and highest QMEANDisCo score, demonstrating an optimal balance of compactness, stability, and accuracy during initial refinement.

3.2. Conformational stability and compactness in ethanol-water mixtures

To ensure comprehensive exploration of the conformational space of the refined α-gliadin model, we conducted extensive MD simulations, each lasting 1 μs and performed in two replicas for every solvent condition. This approach allowed the protein structure ample time to relax from its initial predicted conformation and to effectively sample the equilibrium conformational ensemble under varying ethanol concentrations (0 %, 30 %, 70 %, and 100 % *v*/v EtOH). The primary goal was not only to assess the stability of the model but also to characterize the range of conformations accessible to α-gliadin in different solvent environments. Conformational stability was evaluated using RMSD analysis across the two replicas for each condition. The resulting RMSD distributions, shown in Fig. 1C, reveal distinct trends as ethanol concentration increases. In pure water (0 % EtOH), the mean RMSD across replicas is 1.05 ± 0.18 nm, with individual replica values of 1.06 nm and 1.04 nm, indicating high stability and minimal variability (standard deviation of 0.01 nm). This suggests that the dominant conformational states are well-sampled within the 1 μs timescale. At 30 % EtOH, the mean RMSD decreases to 0.61 ± 0.19 nm (replica values: 0.68 nm and 0.54 nm), reflecting consistent structural behavior and low variability. At 70 % EtOH, the mean RMSD rises slightly to 1.10 ± 0.20 nm (replicas: 1.08 nm and 1.11 nm), yet variability between replicas remains minimal (standard deviation of 0.02 nm), indicating that the system maintains a stable conformational state even at this higher ethanol concentration. In pure ethanol (100 % EtOH), the mean RMSD is 0.79 ± 0.16 nm, with replica-specific values of 0.75 nm and 0.82 nm, showing moderate variability (standard deviation of 0.04 nm). These RMSD results suggest that the α-gliadin structure remains relatively stable across a wide range of ethanol concentrations, with RMSD values ranging from 0.61 to 1.10 nm. This indicates that α-gliadin can maintain its conformational stability despite significant changes in solvent composition.

To gain deeper insights into the compactness of the α-gliadin structure across varying ethanol concentrations (0 %, 30 %, 70 %, and 100 % *v*/v EtOH), we analyzed the Rg values and its corresponding free energy profile, as shown in Fig. 2A. In pure water (0 % EtOH), the average Rg is 2.397 ± 0.137 nm, with replica-specific values of 2.424 nm and 2.369 nm, indicating that α-gliadin maintains a relatively compact structure with moderate fluctuations. At 30 % EtOH, the Rg increases to 2.533 ± 0.044 nm (replica values: 2.511 nm and 2.555 nm), suggesting minimal variability and a slight expansion. At 70 % EtOH, the Rg rises further to 2.639 ± 0.113 nm, with replica-specific values of 2.616 nm and 2.662 nm, reflecting increased structural variability. In pure ethanol (100 % EtOH), the Rg reaches 2.645 ± 0.105 nm (replicas: 2.628 nm and 2.661 nm), indicating that α-gliadin adopts a notably expanded conformation. Overall, the trend of increasing Rg with higher ethanol concentrations (Fig. 2B) demonstrates that α-gliadin progressively adopts a less compact structure in ethanol-rich environments. This observation aligns well with previous experimental and computational studies, which report that α-gliadin, like α-zein (another prolamin in maize), loses compactness and assumes an extended conformation in 70 % aqueous ethanol (Li et al., 2012; Shewry & Tatham, 1997; Tatham & Shewry, 1985; Yu, Xu, et al., 2023). It also suggests that our simulations capture realistic solvent-induced conformational changes. The observed expansion may be attributed to the interaction of ethanol with hydrophobic residues, which disrupts the tertiary structure of protein by exposing nonpolar side chains and weakening hydrophobic packing (Feng et al., 2021, 2022). Additionally, ethanol competes with water for hydrogen bonding sites, thereby reducing the stabilizing intramolecular hydrogen bonds and destabilizing the native fold (Dixit et al., 2002; Tolmachev et al., 2023). Moreover, the reliability of our simulation results is strongly supported by their agreement with experimental observations from circular dichroism and infrared spectroscopy studies that report ethanol-driven increases in α-helical content and molecular expansion of gliadins (Li et al., 2012; Shewry & Tatham, 1997; Tatham & Shewry, 1985; Yu, Xu, et al., 2023). This concordance with experimental data provides key validation that our MD simulation approach accurately captures the solvent-dependent conformational behavior of α-gliadin, despite the complexity of ethanol-water mixtures.

3.3. Residual fluctuations in ethanol environments

To further investigate the residue-level flexibility of α-gliadin, we analyzed the Root Mean Square Fluctuation (RMSF) across all solvent conditions (Fig. S5). A closer examination of RMSF data reveals an interesting pattern: at all ethanol concentrations except for 100 %, the segment of α-gliadin spanning residues Phe238 to Glu257 consistently display elevated RMSF values. The increased flexibility observed in this segment, which is rich in specific residues such as glutamine, likely arises from reduced conformational constraints and the disruptive effect of ethanol on hydrophobic interactions. These interactions typically contribute to the stability and rigidity of protein structures. Interestingly, when the ethanol concentration reaches 100 %, the fluctuations in this region decrease markedly. Specifically, the residues Glu240 to Glu244 exhibit the highest RMSF values within this segment, but these

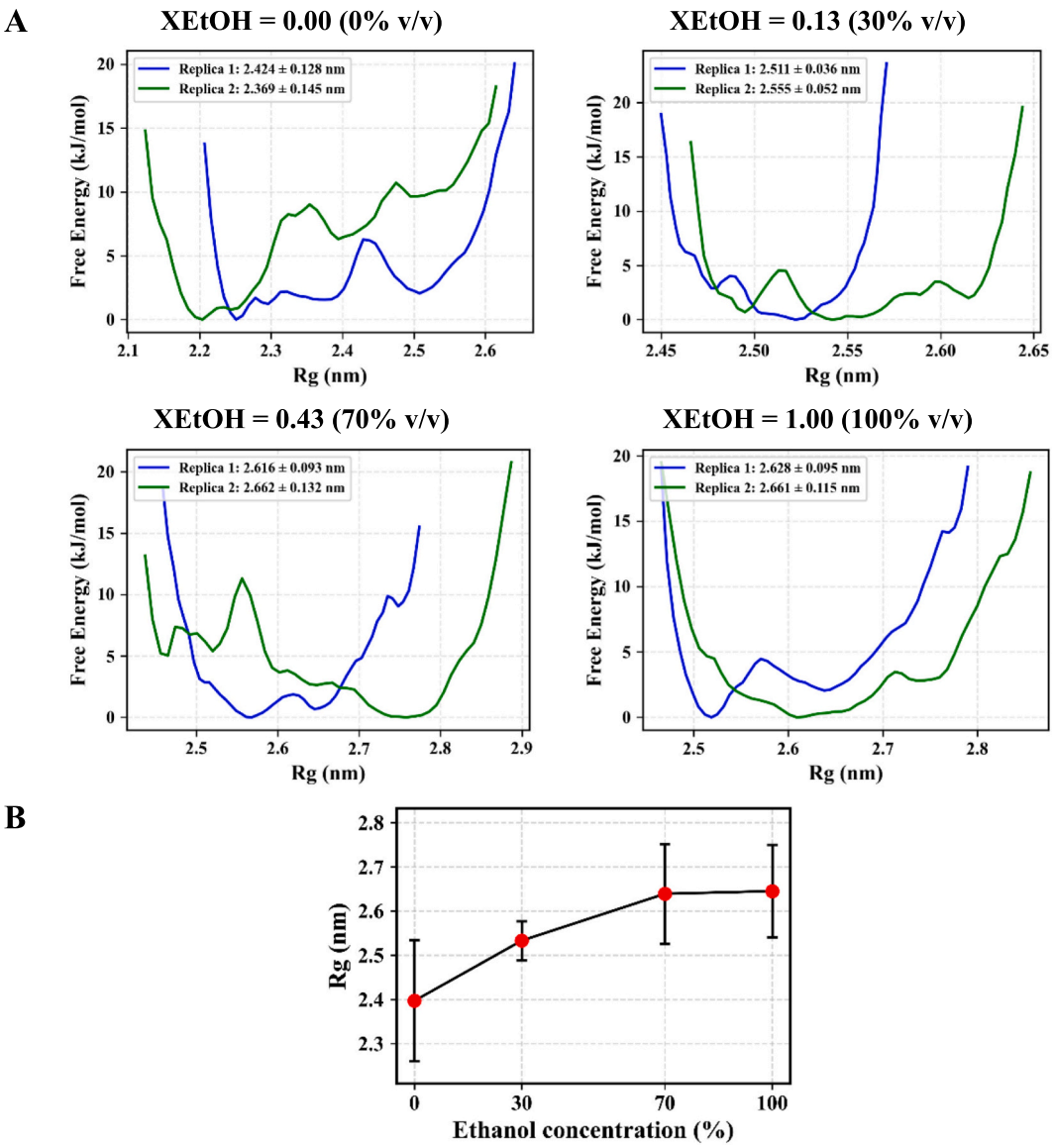


Fig. 2. (A) Free energy landscapes of α-Gliadin as a function of Rg in ethanol-water mixtures (0 %, 30 %, 70 %, and 100 % v/v EtOH) from 1 μs MD simulations, showing up to two replicas with mean ± SD. (B) Mean Rg values with standard deviations plotted against ethanol concentration, summarizing conformational trends across solvent conditions. "XEtOH" indicates the ethanol mole fraction; values in parentheses represent ethanol concentration as % v/v.

values are now reduced to 0.5–0.7 nm, compared to higher fluctuations at lower ethanol concentrations. This suggests that in pure ethanol, the protein backbone in this region becomes more stabilized, possibly due to reduced competition with water for hydrogen bonding or a collapse of the protein into a more compact conformation under these solvent conditions. In addition, the regions spanning residues Gln49 to Tyr75 and Pro65 to Pro95 exhibit pronounced fluctuations at higher ethanol concentrations, particularly at 70 % and 100 % ethanol. These segments are notably rich in proline and glutamine residues. Proline is well known for introducing kinks and disrupting regular secondary structures, while polar side chains of glutamine are highly solvent-accessible and prone to forming transient hydrogen bonds with the surrounding ethanol and water molecules. The abundance of these residues likely contributes to the increased flexibility observed in these regions, as the disruption of hydrophobic interactions and hydrogen bonding by ethanol further amplifies their dynamic behavior.

In contrast, the region from residues Gln135 to Gln225 remains relatively stable across all solvent conditions. The RMSF values in this segment range from 0.18 nm to 0.44 nm, with an average fluctuation of 0.20–0.30 nm, indicating a consistently rigid structure. This stability is likely attributed to the presence of charged residues such as lysine and aspartic acid, as well as cysteine, which can form stabilizing electrostatic interactions and disulfide bonds, respectively. These interactions enhance local rigidity and help maintain the structural integrity of this segment even in the presence of high ethanol concentrations. Taking together, these findings highlight the heterogeneous nature of α-gliadin's response to solvent environment: while proline- and glutamine-rich regions become highly flexible in ethanol-rich conditions, domains stabilized by electrostatic and covalent interactions retain their rigidity, contributing to the overall resilience of the core structure of protein.

These findings have important implications for the functional role of α-gliadin within gluten. The observed decrease in compactness of α-gliadin in ethanol-water mixtures likely enhances its solubility, a critical property leveraged during gluten extraction and dough preparation, where such solvents are routinely employed to selectively isolate gliadin fractions from other gluten components. Increased solubility in these mixtures facilitates the separation and purification of gliadins, which are essential for imparting extensibility and viscosity to dough. However, this structural expansion may have a downside, particularly in

the context of celiac disease. The more extended conformation of α-gliadin in ethanol-rich environments could increase the accessibility of immunogenic epitopes. The regions spanning residues Gln49-Tyr75 and Pro65-Pro95 are of particular interest due to their potential immunogenicity. Notably, the segment from residues Pro65 to Pro95 substantially overlaps with the well-characterized 33-mer peptide (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF), which contains multiple overlapping T-cell stimulatory epitopes recognized by HLA-DQ2/DQ8-restricted T cells in individuals with celiac disease (de Groot et al., 2020; Kim et al., 2004; Shan et al., 2002). Both regions are rich in proline and glutamine residues, a hallmark of immunogenic gliadin sequences, and contain motifs such as PQPQLPY and QLQPFPQPQ that are known to trigger immune responses.

3.4. Solvent accessibility and protein-solvent interactions

The exposure of α-gliadin to solvents across varying ethanol concentrations (0 %, 30 %, 70 %, and 100 % *v*/v EtOH) was evaluated by calculating the SASA, offering insights into protein-solvent interactions and their influence on structural dynamics. Fig. 3A-3B illustrates the SASA distributions for the entire protein and its hydrophobic residues, averaged over two replicas per condition. In pure water (0 % EtOH), the average SASA is 192.68 ± 8.39 nm^2, indicating a relatively compact structure with limited solvent exposure. This aligns with the low Rg of 2.397 nm. At 30 % EtOH, the mean SASA increases to 219.19 ± 2.39 nm^2, suggesting increased solvent accessibility as the protein begins to expand. This corresponds with an increased Rg of 2.533 nm and higher flexibility in regions such as residues Glu240 to Glu244. The enhanced solvent exposure and flexibility observed in segments like Gln227-Ala261 are largely due to disruption of intramolecular hydrogen bonding and increased ethanol interaction; this occurs with only a minor change in the overall salt bridge number (from 3.45 ± 1.01 at 0 % to 3.18 ± 1.05 at 30 % ethanol), which is discussed further in Section 3.4.2. The SASA peaks at 70 % EtOH, with a mean of 238.31 ± 5.56 nm^2. This

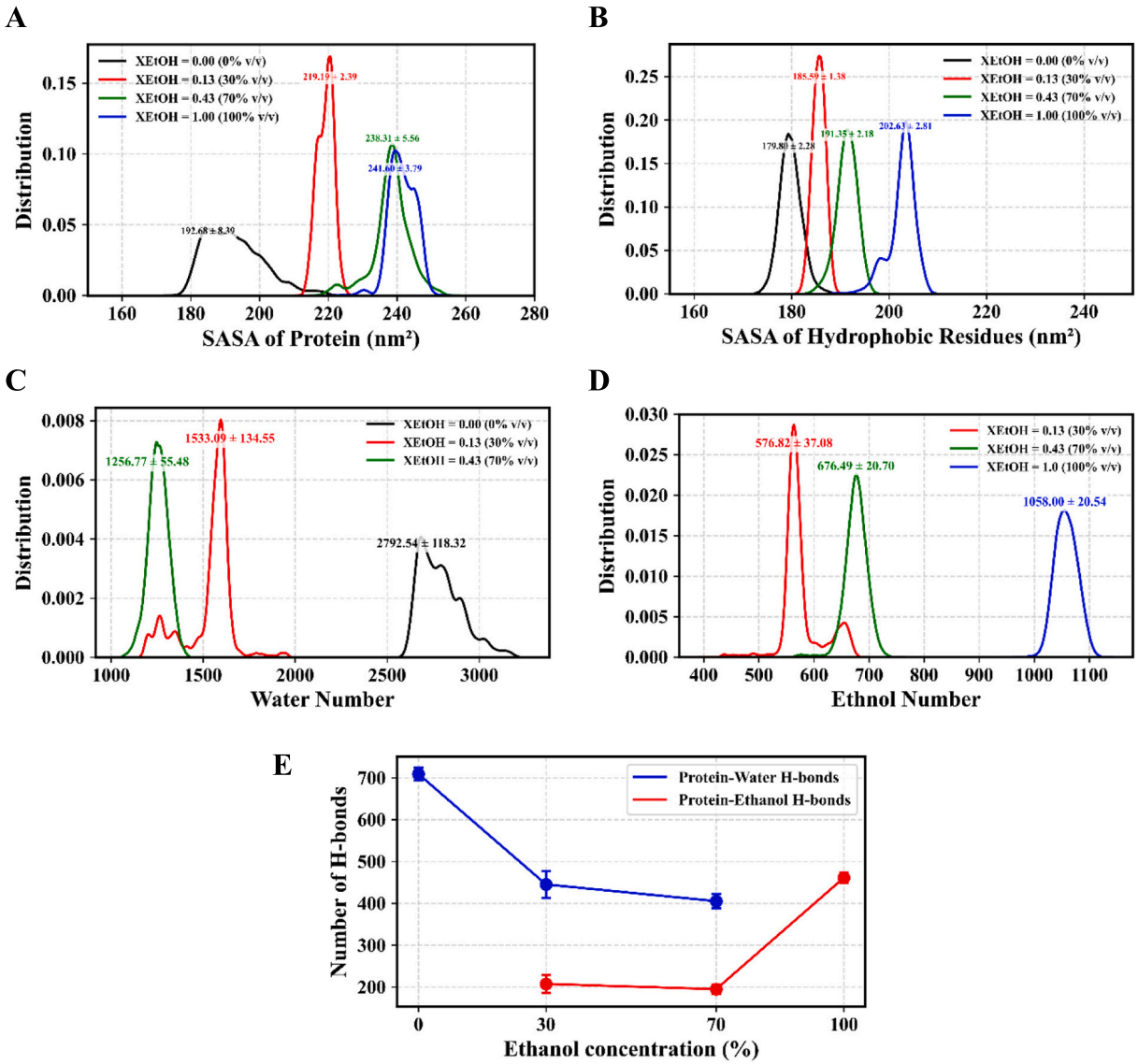


Fig. 3. Kernel density estimations of (A) SASA of the entire protein and (B) SASA of hydrophobic residues, as well as kernel density estimations of number of (C) water molecules and (D) ethanol molecule distributed within 0.6 nm of the protein surface. Panel (E) shows the number of hydrogen bonds (H-bonds) formed between α-gliadin and water (blue) or ethanol (red) in ethanol-water mixtures (0 %, 30 %, 70 %, and 100 % v/v EtOH) over 1 μs MD simulations. "XEtOH" indicates the ethanol mole fraction; values in parentheses represent ethanol concentration as % v/v. Each curve represents the average value across two replicas per solvent condition, with mean ± standard deviation annotated above the distribution peaks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

corresponds to an increased Rg of 2.639 nm and significant flexibility in residues Gln49 to Tyr75, as well as Glu240 to Glu244 (though with slightly decreased flexibility compared to 30 % EtOH). In pure ethanol (100 %), the mean SASA increases further to 241.60 ± 3.79 nm^2, indicating maximal solvent exposure and the lowest variability among replicates. This is consistent with an expanded protein conformation (higher Rg) and heightened flexibility in the Pro65 to Pro95 region. However, the fluctuation in residues Glu240 to Glu244 decreases, suggesting a stabilized yet expanded conformation. These findings indicate that at higher ethanol concentrations (70 % and 100 %), hydrophobic interactions within α-gliadin are maximally disrupted, resulting in large exposure of hydrophobic residues to the solvent.

For the hydrophobic residues specifically, the mean SASA in 0 % EtOH is 179.80 ± 2.28 nm^2, showing limited exposure in a water-dominated environment. At 30 % EtOH, it increases to 185.59 ± 1.38 nm^2, and at 70 % EtOH, it reaches 191.35 ± 2.18 nm^2, reflecting a progressive exposure of hydrophobic regions (e.g., tyrosine, isoleucine) as ethanol disrupts their burial within the protein core, corroborating the RMSF increase in flexible regions. In 100 % ethanol, mean SASA for hydrophobic residues climbs to 202.63 ± 2.81 nm^2, representing maximum exposure. This is likely due to ethanol's superior ability to solvate hydrophobic surfaces, which aligns with overall SASA trend, and the expanded protein structure indicated by Rg and RMSF data. Collectively, these SASA trends reinforce the observation that ethanol-rich solvents drive structural expansion in α-gliadin by competing with water for hydrogen bonding and interacting with hydrophobic residues, thereby disrupting the tertiary structure of protein. The increased solvent accessibility observed at 70 % and 100 % ethanol for both total protein and its hydrophobic residues likely enhance the solubility of α-gliadin, thereby facilitating its extraction during gluten processing. The progressive exposure of hydrophobic residues with rising ethanol concentrations indicates a transition from a compact, water-stabilized state to a more solvent-exposed, ethanol-stabilized conformation. This structural shift may significantly influence the functional properties of α-gliadin in high-ethanol environments, such as those encountered during food processing.

3.4.1. Hydrogen bonding analysis

To further elucidate protein-solvent interactions, we calculated both the number of hydrogen bonds (H-bonds) formed between α-gliadin and water or ethanol, as well as the number of water and ethanol molecules within 0.6 nm of the protein surface. These analyses provide insight into the solvent distribution around α-gliadin and are summarized in Fig. 3C-3E. In pure water (0 % EtOH), the number of protein-water H-bonds is 709.21 ± 15.81, while protein-ethanol H-bonds are 0 because ethanol is not present, suggesting a water-dominated H-bonding network with polar residues, with 2792.54 ± 118.32 water molecules within 0.6 nm. At 30 % EtOH, the number of protein-water H-bonds decreases to 445.15 ± 31.66, and protein-ethanol H-bonds is 207.38 ± 21.96, indicating a shift in solvent interactions, with 1533.09 ± 134.55 water and 576.82 ± 37.08 ethanol molecules. This corresponds to a 44 % reduction in water molecules relative to 0 % EtOH, as ethanol begins to solvate the protein, consistent with the SASA increase. At 70 % EtOH, protein-water H-bonds further decrease to 405.31 ± 10.69, while protein-ethanol H-bonds slightly decreased yet remain stable at 194.84 ± 10.69, with 1256.77 ± 55.48 water and 676.49 ± 20.70 ethanol molecules. This stability arises might be due to a dynamic competition between ethanol and residual water molecules near the protein surface, creating a transitional solvation environment. Such competition moderates the availability of solvent molecules for hydrogen bonding, resulting in fluctuating but overall maintained protein-ethanol hydrogen bonds. This reflects a 54.8 % reduction in water molecules compared to 0 % EtOH (a 10.8 % additional reduction from 30 % EtOH), with ethanol playing a larger role in solvation, aligning with the peak SASA (238.31 nm^2) and maximum structural expansion (higher Rg). The broader fluctuation range in protein-water hydrogen bonds at 70 % EtOH reflects the transitional solvent composition where water and ethanol dynamically compete to solvate the protein. This competition, combined with increased conformational flexibility at this specific concentration, produces greater variability in hydrogen bonding compared to more static solvent environments at lower or higher ethanol concentrations. In pure ethanol (100 % EtOH), protein-water H-bonds drop to 0 because water is absent, while protein-ethanol H-bonds increase sharply to 461.10 ± 12.27, with 1058.00 ± 20.54 ethanol molecules, indicating that ethanol fully solvates the protein, stabilizing the expanded conformation such as region Phe238 to Glu257. This shift in H-bonding and distribution of water and ethanol within 0.6 nm of the protein correlates with the increased exposure of hydrophobic residues (SASA 202.63 nm^2) and the flexibility observed in regions like residues Gln49 to Tyr75 and Pro65 to Pro95, suggesting that ethanol H-bonds with polar groups compensate for the loss of water H-bonds, maintaining the solvent-exposed state of protein.

Due to the higher H-bonding capability of water (up to four H-bonds per molecule compared to ethanol, which can form two), the number of protein-water H-bonds is significantly higher compared to protein-ethanol H-bonds at 30 % and 70 % EtOH. At 30 % EtOH, water molecules still dominate (70 % by volume, molar ratio ~ 87:13 water: ethanol), with 1558.08 ± 142.22 water molecules within 0.6 nm preferentially forming H-bonds with polar residues due to their stronger H-bonding potential. Even at 70 % EtOH (30 % by volume, molar ratio ~ 57:43 water: ethanol), water remains a significant component, maintaining 405.31 protein-water H-bonds and 1256.77 ± 55.48 water molecules within 0.6 nm. Ethanol, while capable of H-bonding, also solvates hydrophobic residues, as seen in the SASA increase for hydrophobic residues (from 179.80 nm^2 at 0 % to 191.35 nm^2 at 70 % EtOH) and the presence of 676.49 ± 20.70 ethanol molecules within 0.6 nm at 70 % EtOH, reducing its competition for H-bonding sites. Additionally, the semi-compact structure of the protein at these concentrations may limit the exposure of additional H-bonding sites, preserving the dominance of water until 100 % EtOH, where ethanol fully replaces water, forming 461.10 H-bonds with 1058.00 ± 20.54 ethanol molecules. These H-bonding dynamics support the observed structural changes, enhancing solubility in ethanol-rich mixtures.

3.4.2. Salt bridge analysis

To evaluate how ethanol concentration affects the structural stability of α-gliadin, we analyzed salt bridge interactions between charged residues within 4 Å under various solvent conditions, averaging results over two replicas for reliability (Fig. 4 and Table S5). At 0 % ethanol, the average number of salt bridges was 3.45 ± 1.01, with Arg164-Asp165 being the most prominent (47.46 %), followed by Arg277-Glu274 (40.77 %) and Arg205-Glu202 (11.70 %). The average distance between these interacting residues was 3.16 ± 0.41 Å. When the ethanol concentration increased to 30 %, the number of salt bridges decreased slightly to 3.18 ± 1.05, confirming only a minor reduction consistent with the modest structural expansion described in the SASA analysis (section 3.4.1). The Arg164-Asp165 pair become even more dominant (62.01 %), while Arg277-Glu274 and Arg205-Glu202 accounted for 25.33 % and 12.67 % of interactions, respectively. The average inter-residues distance also slightly decreased to 3.12 ± 0.35 Å, indicating that although the protein becomes more solvent-accessible, its major electrostatic contacts were retained. At 70 % ethanol, the number of salt bridges declined marginally to 3.12 ± 0.87, with Arg164-Asp165 remaining the most prevalent (52.51 %), followed by Arg277-Glu274 (32.25 %) and Arg205-Glu202 (15.24 %), and the average distance increased slightly to 3.19 ± 0.39 Å. This gradual changes in salt bridge count and distribution up to 70 % ethanol supports the notion that the early conformational expansion of α-gliadin occurs primarily through hydrogen bond disruption and increased ethanol-mediated solvation, rather than through loss of key electrostatic interactions. Notably, at 100 % ethanol, the number of salt bridges rose significantly to 6.99 ± 0.99, accompanied by a marked shift in dominant interactions: Arg205-

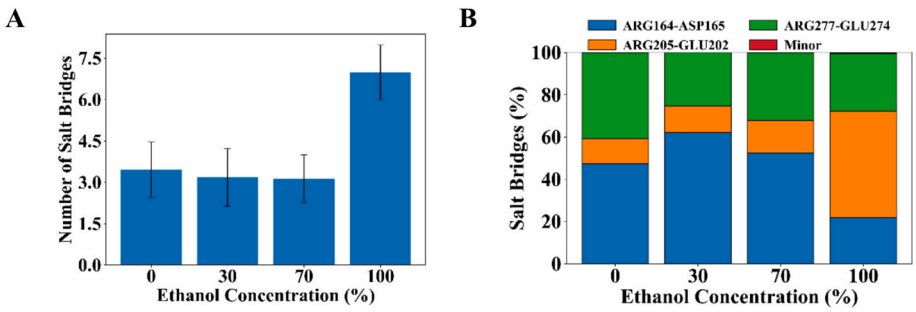


Fig. 4. Effects of ethanol concentration on salt bridge formation and residue pair contributions. (A) Average number of salt bridges, with error bars indicating standard deviations, and (B) relative contributions of key residue pairs to the total number of salt bridges at 0 %, 30 %, 70 %, and 100 % v/v ethanol.

Glu202 became the most prevalent (50.06 %), surpassing Arg277-Glu274 (27.30 %) and Arg164-Asp165 (22.00 %). The average distance between residues also tightened to 3.09 ± 0.40 Å. The increasing dominance of Arg205-Glu202-from 11.70 % at 0 % ethanol to 50.06 % at 100 % alongside the relative decrease or shift in Arg164-Asp165 and Arg277-Glu274, suggests an ethanol-driven reorganization of electrostatic interactions. This trend, coupled with the overall increase in the number of salt bridges from 3.45 to 6.99, reflects a transition from water-mediated stability to ethanol-stabilized contacts. The shift in dominant salt bridge pairs and the increased number of interactions at 100 % ethanol indicate that ethanol promotes the formation of new electrostatic contacts, thereby stabilizing the extended structure of α-gliadin.

3.5. Secondary structure evolution

The evolution of secondary structure of α-gliadin was examined across ethanol concentrations (0 %, 30 %, 70 %, and 100 % *v*/v EtOH) to elucidate solvent-induced structural changes. The secondary structure analysis presented is from one representative replicate, selected for clarity in visualization (Fig. S6). This replicate reflects the overall behavior observed across all simulation replicates, as evidenced by consistent structural stability and other metrics. In pure water (0 % ethanol), α-gliadin exhibits a secondary structure composition of approximately 37.0 % α-helix, 2.0 % β-sheet, 11.0 % β-turn, and 50.0 % random coil. This indicates a partially ordered structure with significant helical and disordered regions, and minimal β-sheet content. These results closely align with experimental circular dichroism studies, which report that gliadins in aqueous solution typically have 30–40 % α-helix, 45–55 % coil, and low β-sheet content (Shewry & Tatham, 1997; Tatham & Shewry, 1985; Urade et al., 2018).

At 30 % EtOH, secondary structure of α-gliadin shows subtle but meaningful changes: the α-helix fraction increases slightly to 37.5 %, β-sheet content decreases to 1.5 %, β-turns rise to 11.5 %, and random coil content drops to 49.5 %. These shifts suggest the onset of structural rearrangement as the protein begins to respond to the presence of ethanol. At 70 % EtOH, α-gliadin exhibits a further increase in α-helix content to 38.0 %, while β-sheet content remains low at 1.5 %. The fraction of β-turns decreases slightly to 11.0 %, and the coil content remains steady at 49.5 %. These results suggest stabilization of helical structure at intermediate ethanol concentrations, even as the protein undergoes significant expansion, as indicated by increased Rg and SASA. These results are consistent with experimental CD studies, which show that gliadins maintain or slightly increase α-helix content and remain largely disordered in 50–70 % ethanol (Shewry & Tatham, 1997; Tatham & Shewry, 1985).

In pure ethanol (100 % EtOH), the helix fraction reaches 38.5 %, β-sheet remains at 1.5 %, β-turns rise to 11.5 %, and coil content decreases to 48.5 %, indicating a slightly more ordered and stabilized helical structure in a fully ethanol environment. The gradual increase in helical content (from 37.0 % to 38.5 %) with rising ethanol concentration suggests that ethanol promotes helical stability, likely by disrupting water-mediated hydrogen bonds and enhancing intramolecular hydrogen bonding within the protein. The minimal change in β-sheet content (2.0 % to 1.5 %) indicates that β-sheet formation is not favored in ethanol, consistent with gliadin's intrinsic structural tendencies. Meanwhile, the slight fluctuation in β-turns (from 11.0 % to 11.5 %) reflects local structural rearrangements in flexible regions as the protein adapts to the solvent. The decrease in coil content (from 50.0 % to 48.5 %) further supports the notion of an expanded yet more stabilized protein conformation at 100 % ethanol.

3.6. Protein solvation structure and thermodynamics

To understand how the solvent environment affects the behavior of α-gliadin, we examined its local solvation structure and thermodynamics for one of the replicas of 1 μs MD simulation trajectories. Given the irregular shape of protein, we utilized Minimum-Distance Distribution Functions rather than standard Radial Distribution Functions to more accurately characterize how water and ethanol molecules arrange themselves around the protein (Figs. 5 and S7). The MDDF results revealed distinct solvation patterns. At 0 % EtOH, water forms two pronounced layers around the protein, with peaks at approximately 1.9 Å (height: 1.49) and 2.6 Å (height: 1.51). This structured hydration shell contributes to the stability and compactness of α-gliadin (Fig. 5A-5B). Upon increasing to 30 % EtOH, the water peak at 1.9 Å diminishes to a height of 1.18, indicating that ethanol begins to displace water from the protein surface. Ethanol itself displays a sharp peak at 2.3 Å (height: 4.51), suggesting preferential association with less polar or hydrophobic regions, which may initiate partial unfolding of the protein. At 70 % EtOH, the water peak at 1.9 Å becomes more pronounced (height: 2.05), likely reflecting water clustering around remaining polar or charged residues as bulk water is depleted. A broader water peak at 4.5 Å (height: 1.19) further supports this selective retention. The peak height of ethanol decreases, indicating competition between the two solvents as the protein expands. In 100 % EtOH, the ethanol peak shifts to 1.9 Å (height: 3.2) but appears broader and less intense, reflecting weaker and less specific interactions with the protein surface compared to water.

To visualize the residue-level specificity of solvent interactions, we analyzed MDDF densities for individual α-gliadin residues (Fig. S7). In pure water, water density was lowest around residues 151–225, a region enriched in hydrophobic and aliphatic residues (e.g., isoleucine, leucine, valine), suggesting these segments are relatively buried and adopt stable secondary structure with limited solvent exposure. At 30 % EtOH, the largest reductions in water density occurred in residues 1–45 and 151–225, consistent with the distribution observed under aqueous conditions. Both regions are dominated by hydrophobic and aliphatic side chains, while additional changes in 258–268 and 300–307 reflect increased solvent accessibility as the protein expands. Ethanol density increased notably in residues 76–150, a proline- and glutamine-rich segment, indicating preferential solvation of flexible, disordered regions. At 70 % EtOH, this pattern persisted, with water remaining near

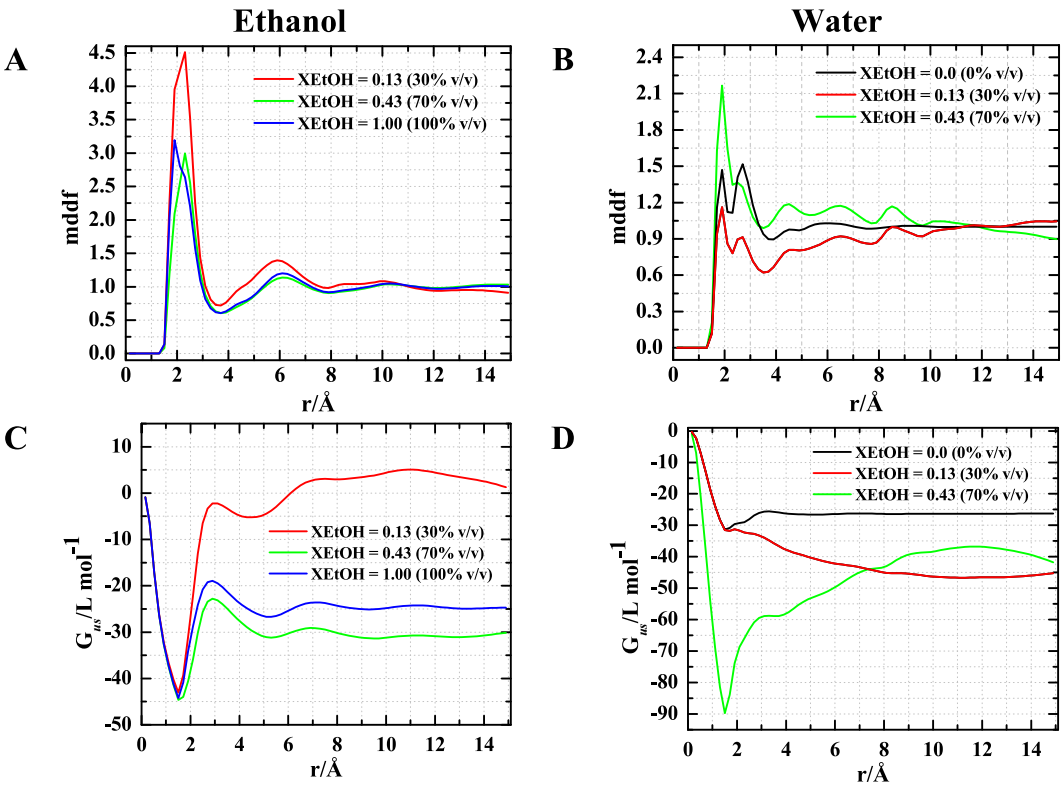


Fig. 5. Solvent organization around α-gliadin across ethanol–water mixtures. (A, B) Minimum Distance Distribution Functions (MDDF) of ethanol and water, respectively, relative to the protein at ethanol concentrations of 0 %, 30 %, 70 %, and 100 % (*v*/v). "XEtOH" indicates the ethanol mole fraction; values in parentheses represent ethanol concentration as % v/v. (C, D) Corresponding Kirkwood–Buff integrals (KBIs) for ethanol and water, quantifying preferential accumulation or exclusion of solvent species near the protein surface.

polar and charged residues and ethanol accumulating around proline- and glutamine-rich stretches, consistent with solvent competition for flexible and hydrophobic regions. In 100 % EtOH, high ethanol density was observed in the majority of N-terminal residues 1–45), proline- and glutamine-rich regions (57–67, 76–97, 118–157, 178–188, 219–276), and the C-terminal (298–307), corresponding to exposed hydrophobic and flexible polar residues. This residue-dependent redistribution of solvent reflects ability of ethanol to displace water, and preferentially solvate nonpolar and disordered surfaces, in line with established models of protein-solvent interactions (Prabhu & Sharp, 2006; Timasheff, 2002).

To quantify solvent preference, we calculated integrals (Fig. 5C-5D). In pure water, the KB integral for water starts at around −30 mol·L^{-1} and levels off at −26 mol·L^{-1} at 15 Å, indicating strong water enrichment and stabilization of the protein. At 30 % EtOH, the KB integral for water becomes more negative (−45 mol·L^{-1}), reflecting increased exclusion of water from the protein's vicinity as ethanol displaces it. Simultaneously, the KB integral for ethanol shifts from −43 mol·L^{-1} to about +5 mol·L^{-1} at longer distances, indicating ethanol's increasing association with the protein, likely due to greater exposure of hydrophobic regions. At 70 % EtOH, the KB integral for water drops sharply to −90 mol·L^{-1} initially and then settles at −41 mol·L^{-1}, while ethanol's integral moves from −45 to −30 mol·L^{-1}, suggesting ongoing competition between the solvents, with water preferentially stabilizing polar regions. In pure ethanol, the KB integral for ethanol levels out at −24 mol·L^{-1}, indicating a weaker, less specific interaction with the protein, as ethanol is less effective at stabilizing charged or polar residues compared to water. These findings demonstrate that the solvation shell around α-gliadin is dynamically remodeled as ethanol concentration increases. Water forms a structured, stabilizing layer that maintains protein compactness, while ethanol progressively disrupts this shell, leading to a looser and less stable environment.

3.7. Conformational dynamics

PCA was employed to elucidate the dominant modes of conformational variability in α-gliadin across ethanol-water mixtures. The analysis focused on the first two principal components (PC1 and PC2), which together captured a substantial proportion of the total variance (Fig. 6). In pure water (0 % EtOH), the conformational landscape was notably rugged in replica 1, displaying four distinct minima (PC1: −3.26 to −7.91, PC2: −4.40 to 2.81), each corresponding to well-sampled regions of the trajectory. In contrast, replica 2 revealed a single dominant minimum, indicating some variability in sampling but overall convergence toward major conformational states at low ethanol concentrations. The variance explained by PC1 and PC2 in water (49–47 % and 13–15 %, respectively) demonstrates that these axes effectively capture most of the large-scale motions.

At 30 % EtOH, the conformational space sampled by α-gliadin became more structured: replica 1 exhibited two closely spaced minima alongside a marked increase in variance explained by PC1 (76.7 %), while replica 2 revealed four closely spaced minima with PC1 and PC2 explaining 58.9 % and 16.2 % of the variance, respectively. This suggests a solvent-driven shift toward a more restricted conformational ensemble dominated by specific collective motions. At 70 % ethanol, the protein exhibited a more complex and heterogeneous landscape, with three distinct minima in both replicas and a lower proportion of variance captured by PC1 (36–39 %), reflecting increased conformational diversity and the presence of multiple accessible states. In 100 % EtOH, the results diverged between replicas: replica 1 exhibited a single deep minimum, while replica 2 sampled four well-separated minima, with PC1 and PC2 together accounting for over 54 % of the variance. These findings suggest that high ethanol content allows α-gliadin to access a broader range of conformations, likely due to the disruption of stabilizing intramolecular interactions and the formation of new electrostatic

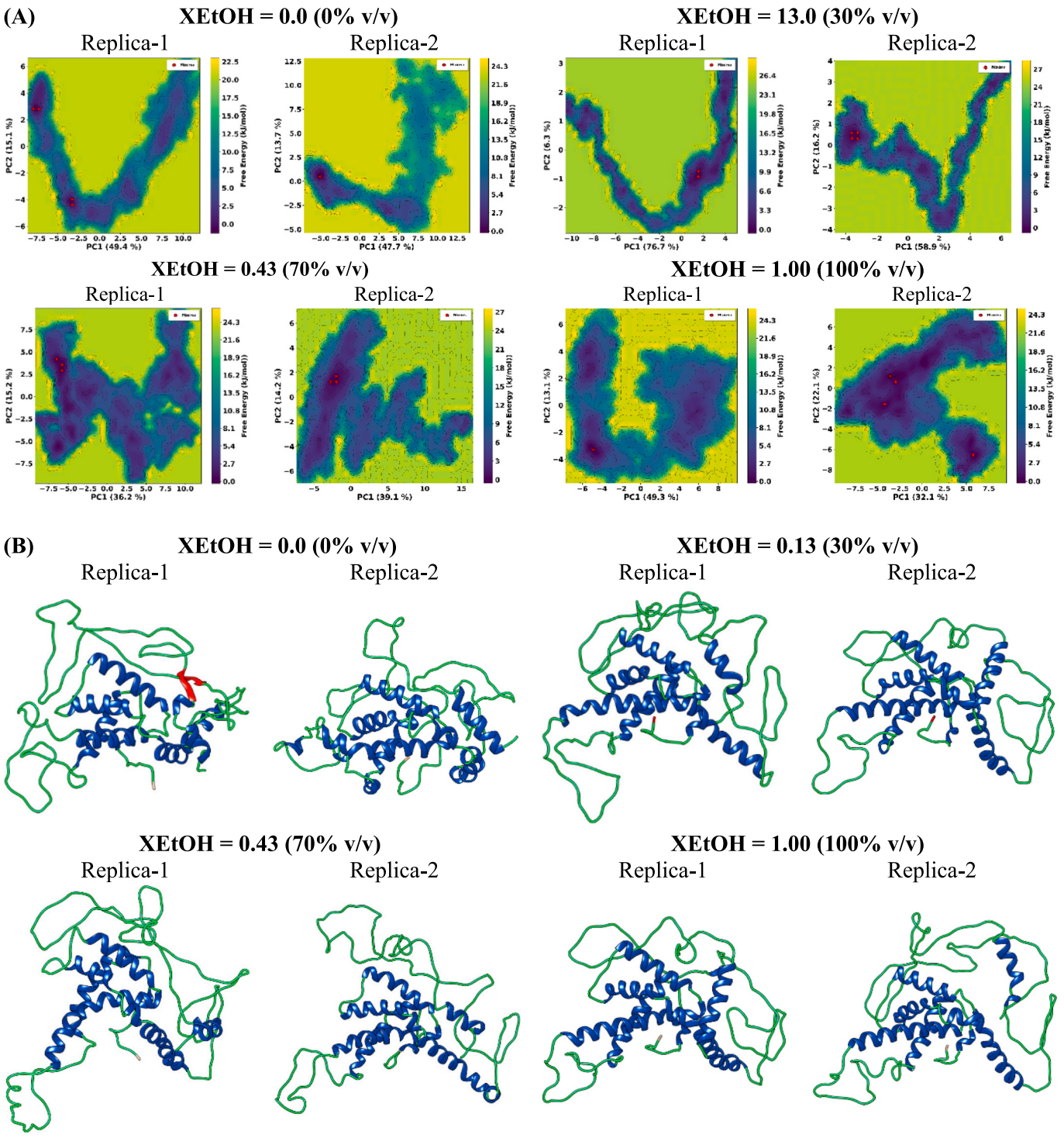


Fig. 6. (A) Free energy landscapes of α-gliadin derived from the first two principal components (PC1 and PC2) from both replicas of 1 μs MD simulations in ethanol-water mixtures (0 %, 30 %, 70 %, and 100 % v/v). "XEtOH" indicates the ethanol mole fraction; values in parentheses represent ethanol concentration as % v/v. Colour gradients indicate free energy (kcal/mol), and red dots mark the energy minima representing predominant conformational states. (B) Representative α-gliadin structures at the global minima, extracted from the free energy landscapes for each solvent condition. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

contacts, as evidenced by the emergence of additional salt bridges. This shift ultimately stabilizes the extended structure of α-gliadin in ethanol-rich environments.

To further interpret these global motions, the global minimum structures identified for each condition from one of the replicas were superimposed, using the global minimum structure obtained at 0 % EtOH as the reference. Per-residue RMSD values were then mapped onto the B-factor field and visualized with a blue-white-red gradient, where red highlights regions of highest structural divergence (Fig. 7). The overall RMSDs between global minima were 11.56 Å (30 % EtOH), 13.28 Å (70 % EtOH), and 10.72 Å (100 % EtOH), indicating significant conformational changes, with the largest deviation observed at 70 % ethanol. Residue-specific RMSD analysis highlighted pronounced flexibility in several regions (Fig. S8), most notably the C-terminal segment

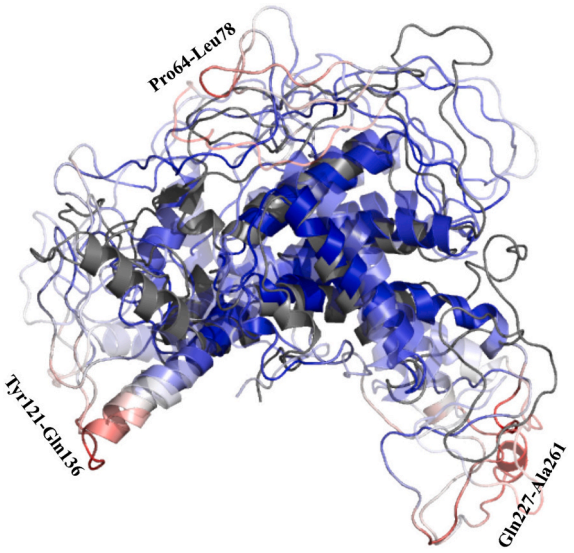


Fig. 7. Superimposed global minima structures of α-gliadin across ethanol concentrations. Structures from 0 % (solid gray), 30 % (solid cartoon with flat sheets), 70 % (slightly transparent with smooth loops), and 100 % (more transparent with simple helices) ethanol are shown superimposed. Structural differences are visualized by mapping per-residue RMSD values onto the B-factor field using a blue-white-red colour gradient, where red indicates regions of highest divergence. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Gln227-Ala261; 8–33 Å), where Ser234, Leu233, and Gln235 exhibited extreme deviations, suggestive of a disordered or solvent-exposed loop. The proline- and glutamine-rich segment Pro64-Leu78 (10–28 Å) also displayed high RMSDs, with Gln72 and Ser71 among the most flexible residues. Additionally, the region spanning Tyr121-Gln136 (12–27 Å) showed substantial deviations, particularly at Pro127, Ile128, Ser129, and Gln130, indicating another flexible segment sensitive to ethanol concentration, which displayed marked increases in both RMSD and RMSF. The N-terminal region (Met1-Pro30; 5–16 Å) was comparatively less flexible. Interestingly, the lower overall RMSD in 100 % ethanol compared to 70 % ethanol may reflect a stabilization effect, possibly due to dehydration or altered hydrophobic interactions. These results demonstrate that ethanol concentration profoundly modulates the conformational landscape of α-gliadin, with specific regions exhibiting pronounced structural variability likely driven by disrupted hydrogen bonding and altered solvent interactions. These findings are consistent with RMSD, Rg, and RMSF analyses, which indicated increased flexibility, expansion, and solvent exposure in ethanol-rich environments. Importantly, the ability of α-gliadin to sample diverse conformational states in ethanol may have implications for its solubility, its functional properties in food matrices, and the accessibility of immunogenic epitopes.

4. Conclusion

This study presents a comprehensive, multi-scale analysis of structural dynamics of α-gliadin in ethanol-water mixtures (0 %, 30 %, 70 %, and 100 % *v*/v), integrating advanced structure prediction, MD simulations, and both residue-level and global conformational analyses. Notably, this work is the first to quantitatively combine state-of-the-art structure prediction with microsecond-scale MD simulations to systematically probe the behavior of α-gliadin across a full range of ethanol concentrations. Comparative model assessment revealed that ESMFold (an AI-based method) and C-I-TASSER (a template-based method) produced more compact and stable α-gliadin structures, while several other AI-driven models tended to favor more extended, flexible conformations. This conclusion is supported by Rg and RMSD, as well as structural quality metrics, and aligns with experimental measurements of gliadin compactness in aqueous solution. After initial refinement and evaluation, the ESMFold model was selected for extended simulations due to its optimal balance of structural compactness and overall quality.

Our extensive MD simulations reveal that α-gliadin maintains remarkable conformational stability across all the solvent conditions. However, there is a clear trend toward expansion and increased solvent accessibility as ethanol concentration rises. Residue-level analyses reveal that proline- and glutamine-rich regions (Pro64-Leu78, Tyr121-Gln136, Gln227-Ala261) are the most dynamic, whereas the central region remains consistently rigid. Residues like Ser234, Leu233, and Gln235 exhibit extreme fluctuations, suggesting these segments act as disordered or solvent-exposed loops. In contrast, the central region (Gln135-Gln225) remains consistently rigid, stabilized by charged residues (lysine, aspartic acid) and cysteine-mediated disulfide bonds, which confer local structural integrity even in ethanol-rich environments. Secondary structure analysis reveals modest stabilization of helical content and a reduction in β-sheet content, especially at intermediate and high ethanol concentrations. Residue-level analyses underscore the heterogeneous flexibility of α-gliadin: proline- and glutamine-rich regions become highly dynamic in ethanol-rich environments, while segments stabilized by electrostatic or covalent interactions retain their rigidity. Solvation analyses using MDDF and KD integrals demonstrate a clear, residue-dependent remodeling of the protein's hydration shell: water forms a compact, stabilizing layer in pure aqueous conditions, while ethanol progressively displaces water and preferentially solvates hydrophobic and flexible polar regions. PCA provides a global perspective, showing that solvent composition critically shapes the conformational landscape of α-gliadin. Lower ethanol concentrations favor compact, convergent conformational basins, whereas higher ethanol levels promote greater conformational diversity and the sampling of multiple distinct minima. These global motions are driven by the same flexible regions identified in RMSF and SASA analyses, highlighting a direct link between local flexibility and large-scale structural transitions.

These findings provide deeper insights into the behavior of α-gliadin in food processing and potential immunogenicity, offering valuable insights for improved gluten extraction and immune responses mitigation strategies. Collectively, our results provide new molecular-level insights into the solvent-dependent dynamics of α-gliadin, with implications for its solubility, functional properties in food processing, and the accessibility of immunogenic epitopes relevant to celiac disease. This work not only advances our understanding of the structural plasticity of intrinsically disordered food proteins but also establishes a computational framework for future studies aimed at modulating gluten protein properties for enhanced food functionality and reduced immunogenicity. However, the absence of an experimentally determined α-gliadin structure and the complexity of real gluten environments beyond ethanol-water mixtures introduce certain limitations. Future research should focus on integrating experimental data for further refinement, exploring other food-relevant solvent conditions, and extending analyses to other gliadin isoforms and glutenin proteins. Such efforts will improve our fundamental understanding of gluten behavior, ultimately benefiting both food technology and health.

CRediT authorship contribution statement

Nandan Kumar: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yonghui Li:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This is contribution no. 25-243-J from the Kansas Agricultural Experimental Station.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2025.147044.

Data availability

Data will be made available on request.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., … Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science, 373*(6557), 871–876. https://doi.org/10.1126/science.abj8754

Cole, E. W., Kasarda, D. D., & Lafiandra, D. (1984). The conformational structure of A-gliadin: Intrinsic viscosities under conditions approaching the native state and under denaturing conditions. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 787*(3), 244–251. https://doi.org/10.1016/0167-4838(84)90315-7

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science, 2*(9), 1511–1519. https://doi.org/10.1002/pro.5560020916

Combet, C., Blanchet, C., Geourjon, C., & Deléage, G. (2000). NPS@: Network protein sequence analysis. *Trends in Biochemical Sciences, 25*(3), 147–150. https://doi.org/10.1016/S0968-0004(99)01540-6

Delcour, J. A., Joye, I. J., Pareyt, B., Wilderjans, E., Brijs, K., & Lagrain, B. (2012). Wheat gluten functionality as a quality determinant in cereal-based food products. *Annual Review of Food Science and Technology, 3*, 469–492. https://doi.org/10.1146/annurev-food-022811-101303

Dixit, S., Crain, J., Poon, W. C. K., Finney, J. L., & Soper, A. K. (2002). Molecular segregation observed in a concentrated alcohol–water solution. *Nature, 416*(6883), 829–832. https://doi.org/10.1038/416829a

Fasano, A., & Catassi, C. (2012). Clinical practice. Celiac disease. *The New England Journal of Medicine, 367*(25), 2419–2426. https://doi.org/10.1056/NEJMcp1113994

Feng, Y., Ma, X., Kong, B., Chen, Q., & Liu, Q. (2021). Ethanol induced changes in structural, morphological, and functional properties of whey proteins isolates: Influence of ethanol concentration. *Food Hydrocolloids, 111*, Article 106379. https://doi.org/10.1016/j.foodhyd.2020.106379

Feng, Y., Yuan, D., Kong, B., Sun, F., Wang, M., Wang, H., & Liu, Q. (2022). Structural changes and exposed amino acids of ethanol-modified whey proteins isolates promote its antioxidant potential. *Current Research in Food Science, 5*, 1386–1394. https://doi.org/10.1016/j.crfs.2022.08.012

Gazi, R., Maity, S., & Jana, M. (2023). Conformational features and hydration dynamics of proteins in Cosolvents: A perspective from computational approaches. *ACS Omega, 8*(3), 2832–2843. https://doi.org/10.1021/acsomega.2c08009

Gibrat, J.-F., Garnier, J., & Robson, B. (1987). Further developments of protein secondary structure prediction using information theory: New parameters and consideration of residue pairs. *Journal of Molecular Biology, 198*(3), 425–443. https://doi.org/10.1016/0022-2836(87)90292-0

de Groot, N. G., Otting, N., Maccari, G., Robinson, J., Hammond, J. A., Blancher, A., … Bontrop, R. E. (2020). Correction to: Nomenclature report 2019: Major histocompatibility complex genes and alleles of great and small ape and old and new world monkey species. *Immunogenetics, 72*(1), 131–132. https://doi.org/10.1007/s00251-019-01146-5

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kim, C.-Y., Quarsten, H., Bergseng, E., Khosla, C., & Sollid, L. M. (2004). Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease. *Proceedings of the National Academy of Sciences, 101*(12), 4175–4179. https://doi.org/10.1073/pnas.0306885101

Kumar, N., Du, Z., Amachawadi, R. G., Guo, X., Zhao, J., & Li, Y. (2025). Membrane selectivity mechanisms of the antimicrobial peptide Snakin-Z against prokaryotic and eukaryotic membrane models. *The Journal of Physical Chemistry B, 129*(18), 4392–4409. https://doi.org/10.1021/acs.jpcb.5c01013

Kumar, N., Du, Z., & Li, Y. (2025). pLM4CPPs: Protein language model-based predictor for cell penetrating peptides. *Journal of Chemical Information and Modeling, 65*(3), 1128–1139. https://doi.org/10.1021/acs.jcim.4c01338

Kumar, N., & Sastry, G. N. (2021). Study of lipid heterogeneity on bilayer membranes using molecular dynamics simulations. *Journal of Molecular Graphics and Modelling, 108*, Article 108000. https://doi.org/10.1016/j.jmgm.2021.108000

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography, 26*(2), 283–291. https://doi.org/10.1107/S0021889892009944

Li, Y., Li, J., Xia, Q., Zhang, B., Wang, Q., & Huang, Q. (2012). Understanding the dissolution of α-Zein in aqueous ethanol and acetic acid solutions. *The Journal of Physical Chemistry B, 116*(39), 12057–12064. https://doi.org/10.1021/jp305709y

Li, Y., Zhang, C., Zheng, W., Zhou, X., Bell, E. W., Yu, D.-J., & Zhang, Y. (2021). Protein inter-residue contact and distance prediction by coupling complementary coevolution features with deep residual networks in CASP14. *Proteins: Structure, Function, and Bioinformatics, 89*(12), 1911–1921. https://doi.org/10.1002/prot.26211

Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., Smetanin, N., Verkuil, R., Kabeli, O., Shmueli, Y., dos Santos Costa, A., Fazel-Zarandi, M., Sercu, T., Candido, S., & Rives, A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science, 379*(6637), 1123–1130. https://doi.org/10.1126/science.ade2574

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., & Shaw, D. E. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics, 78*(8), 1950–1958. https://doi.org/10.1002/prot.22711

Martínez, L. (2022). ComplexMixtures.Jl: Investigating the structure of solutions of complex-shaped molecules from a solvent-shell perspective. *Journal of Molecular Liquids, 347*, Article 117945. https://doi.org/10.1016/j.molliq.2021.117945

Nassar, R., Brini, E., Parui, S., Liu, C., Dignon, G. L., & Dill, K. A. (2022). Accelerating protein folding molecular dynamics using inter-residue distances from machine learning servers. *Journal of Chemical Theory and Computation, 18*(3), 1929–1935. https://doi.org/10.1021/acs.jctc.1c00916

Pearce, R., Li, Y., Omenn, G. S., & Zhang, Y. (2022). Fast and accurate ab initio protein structure prediction using deep learning potentials. *PLoS Computational Biology, 18*(9), Article e1010539. https://doi.org/10.1371/journal.pcbi.1010539

Prabhu, N., & Sharp, K. (2006). Protein-solvent interactions. *Chemical Reviews, 106*(5), 1616–1623. https://doi.org/10.1021/cr040437f

Rost, B., & Sander, C. (1993). Prediction of protein secondary structure at better than 70% accuracy. *Journal of Molecular Biology, 232*(2), 584–599. https://doi.org/10.1006/jmbi.1993.1413

Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., … Khosla, C. (2002). Structural basis for gluten intolerance in celiac sprue. *Science, 297*(5590), 2275–2279. https://doi.org/10.1126/science.1074129

Shewry, P. R., & Belton, P. S. (2024). What do we really understand about wheat gluten structure and functionality? *Journal of Cereal Science, 117*, Article 103895. https://doi.org/10.1016/j.jcs.2024.103895

Shewry, P. R., Halford, N. G., Belton, P. S., & Tatham, A. S. (2002). The structure and properties of gluten: An elastic protein from wheat grain. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 357*(1418), 133–142. https://doi.org/10.1098/rstb.2001.1024

Shewry, P. R., & Tatham, A. S. (1997). Disulphide bonds in wheat gluten proteins. *Journal of Cereal Science, 25*(3), 207–227. https://doi.org/10.1006/jcrs.1996.0100

Singh, A., Vanga, S. K., Orsat, V., & Raghavan, V. (2018). Application of molecular dynamic simulation to study food proteins: A review. *Critical Reviews in Food Science and Nutrition, 58*(16), 2779–2789. https://doi.org/10.1080/10408398.2017.1341864

Sollid, L. M. (2002). Coeliac disease: Dissecting a complex inflammatory disorder. *Nature Reviews Immunology, 2*(9), 647–655. https://doi.org/10.1038/nri885

Tatham, A. S., & Shewry, P. R. (1985). The conformation of wheat gluten proteins. The secondary structures and thermal stabilities of α-, β-, γ- and ω-gliadins. *Journal of Cereal Science, 3*(2), 103–113. https://doi.org/10.1016/S0733-5210(85)80021-7

Timasheff, S. N. (2002). Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proceedings of the National Academy of Sciences of the United States of America, 99*(15), 9721–9726. https://doi.org/10.1073/pnas.122225399

Tolmachev, D. A., Malkamäki, M., Linder, M. B., & Sammalkorpi, M. (2023). Spidroins under the influence of alcohol: Effect of ethanol on secondary structure and molecular level solvation of silk-like proteins. *Biomacromolecules, 24*(12), 5638–5653. https://doi.org/10.1021/acs.biomac.3c00637

Urade, R., Sato, N., & Sugiyama, M. (2018). Gliadins from wheat grain: An overview, from primary structure to nanostructures of aggregates. *Biophysical Reviews, 10*(2), 435–443. https://doi.org/10.1007/s12551-017-0367-2

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research, 35*(suppl_2), W407–W410. https://doi.org/10.1093/nar/gkm290

Wieser, H. (2007). Chemistry of gluten proteins. *Food Microbiology, 24*(2), 115–119. https://doi.org/10.1016/j.fm.2006.07.004

Wu, R., Ding, F., Wang, R., Shen, R., Zhang, X., Luo, S., Su, C., Wu, Z., Xie, Q., Berger, B., Ma, J., & Peng, J. (2022). *High-resolution de novo structure prediction from primary*

sequence (p. 2022.07.21.500999). bioRxiv. https://doi.org/10.1101/2022.07.21.500999

Yang, J., Kuang, H., Kumar, N., Song, J., & Li, Y. (2024). Changes of structure properties and potential allergenicity of ovalbumin under high hydrostatic pressures. *Food Research International, 190*, Article 114658. https://doi.org/10.1016/j.foodres.2024.114658

Yang, J., Kumar, N., Kuang, H., Li, Y., & Song, J. (2025). Modification of structural properties and potential allergenicity of ovalbumin induced by ultrasound treatment during *in vitro* digestion. *Food Chemistry, 493*, Article 146015. https://doi.org/10.1016/j.foodchem.2025.146015

Yang, J., Kumar, N., Kuang, H., Song, J., & Li, Y. (2025). Changes of digestive stability and potential allergenicity of high hydrostatic pressure-treated ovalbumin during *in vitro* digestion. *Food Chemistry, 473*, Article 142962. https://doi.org/10.1016/j.foodchem.2025.142962

Yang, J., & Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research, 43*(W1), W174–W181. https://doi.org/10.1093/nar/gkv342

Yu, J.-T., Li, X.-Y., Huang, J.-H., Yu, M.-Y., Wu, Z.-Y., & Cao, S.-L. (2023). Molecular dynamics simulation of α-gliadin in ethanol/aqueous organic solvents. *International Journal of Food Science & Technology, 58*(2), e1–e9. https://doi.org/10.1111/ijfs.15754

Yu, Y., Xu, S., He, R., & Liang, G. (2023). Application of molecular simulation methods in food science: Status and prospects. *Journal of Agricultural and Food Chemistry, 71*(6), 2684–2703. https://doi.org/10.1021/acs.jafc.2c06789

Zheng, W., Zhang, C., Li, Y., Pearce, R., Bell, E. W., & Zhang, Y. (2021). Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell Reports Methods, 1*(3), Article 100014. https://doi.org/10.1016/j.crmeth.2021.100014