

Clustering and cross-linking of the wheat storage protein α -gliadin: A combined experimental and theoretical approach

Joel Markgren^{a,1}, Faiza Rasheed^{a,b,1,2}, Mikael S. Hedenqvist^b, Marie Skepö^c, Eva Johansson^{a,*}

^a Department of Plant Breeding, Swedish University of Agricultural Sciences, P.O. Box 190, SE-234 22 Lomma, Sweden

^b Department of Fibre and Polymer Technology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

^c Theoretical Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden

ARTICLE INFO

Keywords:

Disulphide bonds
Monte Carlo simulations
Polymers
Synthetic peptides

ABSTRACT

Our aim was to understand mechanisms for clustering and cross-linking of gliadins, a wheat seed storage protein type, monomeric in native state, but incorporated in network while processed. The mechanisms were studied utilizing spectroscopy and high-performance liquid chromatography on a gliadin-rich fraction, in vitro produced α -gliadins, and synthetic gliadin peptides, and by coarse-grained modelling, Monte Carlo simulations and prediction algorithms. In solution, gliadins with α -helix structures (dip at 205 nm in CD) were primarily present as monomeric molecules and clusters of gliadins (peaks at 650- and 700-s on SE-HPLC). At drying, large polymers (R_g 90.3 nm by DLS) were formed and β -sheets increased (14% by FTIR). Trained algorithms predicted aggregation areas at amino acids 115–140, 150–179, and 250–268, and induction of liquid-liquid phase separation at P- and Poly-Q-sequences (Score = 1). Simulations showed that gliadins formed polymers by tail-to-tail or a hydrophobic core (Kratky plots and R_{ee} = 35 and 60 for C- and N-terminal). Thus, the N-terminal formed clusters while the C-terminal formed aggregates by disulphide and lanthionine bonds, with favoured hydrophobic clustering of similar/exact peptide sections (synthetic peptide mixtures on SE-HPLC). Mechanisms of clustering and cross-linking of the gliadins presented here, contribute ability to tailor processing results, using these proteins.

1. Introduction

The gluten proteins in the wheat grain form one of the most extensive polymeric aggregates in Nature [1], a characteristic that contributes to the unique properties of wheat, e.g. viscoelastic behavior in the bread-baking process. The gluten proteins are known to impact both the strength and the viscosity of the dough [2]. Recent studies have also indicated a value of these protein characteristics for production of materials with absorption properties [3] and for polymeric materials with properties resembling plastics [4,5].

Approximately 80% of the wheat grain proteins are gluten proteins, also referred to as seed storage proteins, and they serve as an amino acid depot for the developing seed embryo. Similar homologous proteins are found in all grass species [6–9], although the proteins from other species

have significantly less ability to form large protein polymers due to slight but critical differences in structure, primary amino acid sequence compositions, and molecular weights. Based on the solubility while extracted, the gluten proteins are divided into two types, the glutenins and the gliadins [10–12]. The glutenins are present as polymers in the seed, aggregated through covalent disulfide bonds, while the gliadins are monomeric with limited intra-molecular disulphide bonds [13–15]. The gliadins are further divided into four homologous types: α/β , γ and ω [10–12], based on their mobility on acid gels [16]. The different types of gliadins share dominating regions of repeating amino acid motifs and homologous conserved cysteine motifs, except for ω -gliadins that lacks cysteines [11,17]. Therefore, the present study uses the α -gliadins as a model system to describe gliadins properties regarding clustering by weak forces and aggregation by covalent bonds.

* Corresponding author.

E-mail addresses: joel.markgren@slu.se (J. Markgren), faizar@kth.se (F. Rasheed), mikaelhe@kth.se (M.S. Hedenqvist), marie.skepö@teokem.lu.se (M. Skepö), eva.johansson@slu.se (E. Johansson).

¹ Authors are considered to have contributed equally.

² Present address: Department of Biotechnology, Quaid-i-Azam University, Islamabad, Pakistan

<https://doi.org/10.1016/j.ijbiomac.2022.05.032>

Received 20 December 2021; Received in revised form 30 April 2022; Accepted 4 May 2022

Available online 14 May 2022

0141-8130/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Although the gliadins are monomeric in their native state, they can form clusters by weak forces and cross-links through disulphide and irreversible bonds while processed with heat and pressure [17]. The α -gliadins contain hydrophobic amino acids, primarily in evolutionarily conserved parts, and in parts of its repetitive sequences, which are known to form either disordered hydrophobic cores or self-assembling clusters [18–20]. Most of the gliadins also contain sequences, which may be associated with protein coagulation into small clusters, known e.g. for proteins that occur in Huntingtins disease [21–23]. During dough mixing and materials production, it has been shown that the intra-molecular disulfide bonds of the gliadins are broken and some of them are re-formed into inter-molecular cross-links with surrounding gluten proteins [24,25]. Besides the general knowledge about protein clustering and cross-linking described above, information is scarce about the behavior for gliadins. An increased understanding would contribute to opportunities to improve the rheological impact from gliadins and tailor protein structures in e.g. dough systems, and for production of polymeric or absorbing materials.

Gliadins belong to the family of intrinsically disordered proteins (IDPs), with various conformational and structural properties [18,26,27]. Due to their length, \approx 250 amino acids, they are challenging to study on an atomistic level from a sampling and convergence perspective. Here Monte Carlo (MC) simulations in combination with coarse-grained modelling play an important role where artificial moves can be implemented. Computational methods have previously been utilized in a few studies to evaluate α -gliadin and gliadin structures in general [18,28–30], however, regarding simulations on peptide fragments, mostly fragments associated with celiac disease [31,32] and polyglutamine (but not in a gliadin context) [33] have been evaluated. From an experimental perspective, gliadins are also laborious to separate from other gliadins and glutenins, to allow proper structural and image analysis of a pure protein fraction. Previous experimental studies have also shown that the synthesis and evaluation of gluten peptides provide valuable information about the properties of the protein [20,34–38]. Gliadin fractions have been produced heterogeneously in tobacco and *Xenopus Oocytes* systems [39–41], and for α -gliadin in particular, *Escherichia coli* (*E. coli*) [42] has been adopted. However, until now, studies on heterogeneously produced gliadins have not focused on structural characterisation. The recent advances in computational simulation techniques, both in terms of enhanced sampling techniques and computer power, combined with protein production in heterogenous systems, provide innovative opportunities to study structure and cross-linking behavior in gliadins [43].

The aim of this study was to understand the mechanisms behind, and properties of cross-linking and aggregation (covalent bonds), as well as the clustering (weak forces) of the gliadins. Thus, experimental studies were carried out using gliadin-rich fractions, *E. coli* produced α -gliadin and five different synthetic peptides designed by us. Furthermore, MC simulations on a coarse-grained level were used to understand the underlying physics and inter/intra-molecular interactions. The second aim was to evaluate possibilities for large scale fractionation/production of gliadins to be used in plastic materials. The present study is unique, as this is the first time *E. coli* produced α -gliadin and synthetic peptides of wheat have been utilized to understand the clustering and cross-linking ability of the wheat storage proteins. Also, the use of simulation tools to understand mechanisms behind clustering and cross-linking of gliadins has hitherto been limited. Thus, this study contributes unique opportunities to contribute state-of-the-art novel understanding to the field.

2. Material and methods

Materials and methods utilized in this study are summarized in Table 1, and further described below.

Table 1

Materials utilized in the present study and analyses carried out on each of the materials.

Material	Analyses carried out		
	Secondary structure	HPLC	Simulations
Gliadin-rich fraction		SE-HPLC, RP-HPLC	
<i>E. coli</i> produced α -gliadin	CD, FTIR, DLS	SE-HPLC, RP-HPLC	Course-grain model, Monte Carlo simulations, Prediction algorithms
Synthetic peptide		SE-HPLC, RP-HPLC	Course-grain model, Monte Carlo simulations, Prediction algorithms

CD=Circular Dichroism, FTIR = Fourier Transform Infrared, DLS=Dynamic Light Scattering, SE-HPLC=Size Exclusion-High Performance Liquid Chromatography, RP-HPLC = Reversed Phase-High Performance Liquid Chromatography.

2.1. Materials

The following materials, as described below, have been used to study the effect of cross-linking, clustering, and aggregation behavior of gliadins; (1) gliadin-rich fractions, (2) *E. coli* produced α -gliadin, and (3) synthetic peptides.

2.1.1. Production of gliadin-rich fractions

The gliadin-rich fraction (91% protein) was produced from wheat gluten (77.7% protein, Lantmännen Reppe AB, Sweden), following the method described by Muneer et al. [44], through extraction with 70% ethanol followed by evaporation at reduced pressure before lyophilisation. Thus, 16 g wheat gluten powder was slowly dispersed in 200 ml of 70% ethanol under constant stirring, and the solution was then shaken (IKA-KS 500, IKA, Germany) for 30 min at 300 rpm and centrifuged (Sorvall RC 6+ centrifuge, Thermo Scientific, USA) for 10 min at 26,413g (12 000 rpm). Thereafter, the gliadin in ethanol solution was decanted, evaporated (Bichi, Switzerland) at 75 ± 5 °C under vacuum, lyophilized and ground (IKA A10, IKA, Germany) into powder.

2.1.2. *E. coli* production of gliadin

The purified (purity further described in results and discussion) α -gliadin produced and used in this study was based on the primary sequence described at www.uniprot.org under the accession number Q9ZP09. The primary sequence and the charge and hydrophobicity distribution, are shown in supplementary A. For the protein production in vitro, the plasmid DNA PMR191 containing the α -gliadin gene [45] was used, kindly provided by the group headed by Dr. Rossi at the Institute of Food Sciences, the National Research Council of Italy. The production of α -gliadin in *E. coli* was carried out, as described below, according to Senger et al. [42] with some minor modifications.

2.1.2.1. Preparation of *E. coli* cells for α -gliadin production. Bacterial *E. coli* BL21(DE3) pLysS cells were activated on liquid Lysogeny broth (LB) medium containing 10 mg/l tetracycline at 37 °C. Activated cells were further cultured in 25 ml liquid LB medium with 10 mg/l tetracycline and shaking at 37 °C overnight. Five ml of these activated and cultured cells were diluted in 50 ml pre-warmed LB medium (without antibiotics) in a 250 ml flask with shaking (250 rpm). The OD₆₀₀ of the growing bacterial culture was measured every 20 min, and when reaching 0.4, the flask with the culture was rapidly transferred to an ice water bath for 15–30 min. The bacterial culture was then transferred to ice-cold sterile centrifuge tubes and centrifuged at 2500 rpm for 15 min at 4 °C. The supernatant was decanted, and the cell pellet was resuspended in 50 ml of ice-cold sterile water, followed by centrifugation at 2500 rpm for 15 min at 4 °C. Thereafter, the supernatant was again decanted, and the pellet was resuspended in 10 ml of ice-cold sterile

10% glycerol, followed by centrifugation at 2500 rpm for 20 min at 4 °C. Then, the pellet was resuspended again in 1 ml of ice-cold 10% glycerol. The cell suspension was diluted to a concentration of 2×10^{10} to 3×10^{10} cells/ml (1 OD = approx. 2.5×10^8 cells / ml) with ice-cold 10% glycerol. Aliquots (40 μ l) of the competent cells were prepared into sterile ice-cold 0.5 ml microfuge tubes and dropped into liquid nitrogen and stored at -80 °C.

2.1.2.2. Transformation with plasmid DNA. Electroporation cuvettes were placed on ice, and the competent cells were thawed slowly on ice. Approximately 30 ng (1 μ l) of Plasmid DNA (PMR191) was added to the microfuge tube containing the competent bacterial cells, and the tube was incubated on ice for 2 min. The electroporation of plasmid DNA to bacterial cells was carried out under an electrical pulse of 25 μ F capacitance, 1.8 kV, and 200- Ω resistance for 4.5 milliseconds. The cuvette was quickly removed, and 0.5 ml of LB medium was added at room temperature in a laminar flow hood. The cells were added to a sterile tube and incubated at 37 °C for 1 h. The cultured solution was poured onto a solid LB medium with 50 mg/l ampicillin. The solution was evenly dispersed with a sterile disposable “Z-shaped” rod, and cells were cultured at 37 °C until single colonies appeared. The plates were stored at 4 °C for further use.

2.1.2.3. Transformation validation. A two-step transformation confirmation was carried out, i.e. through the growth of single bacterial colonies in selection medium (ampicillin 50 mg/l) and Polymerase Chain Reaction (PCR). Colony PCR was performed in Biometra Thermocycler AKSDF Model (Analytik Jena, USA) using the α -gliadin-F-AJ130948 (forward primer) GCCCATAGCATGACGATCA and α -gliadin-R-AJ130948 (reverse primer) GTTGGAAGGAGACCTGGCTC. High fidelity DNA polymerase was used for the amplification of the gene with its buffer and oligonucleotides. The following PCR conditions were used: 95 °C of denaturation for 5 min, and a thirty cycle lap of 95 °C for 15 s of denaturation, 55 °C for primer annealing and 72 °C for 30 s of extension, and a final extension of 72 °C for 5 min.

2.1.2.4. Large scale preparation. Initially, each single transformed bacterial colony was cultured in 10 ml of LB media containing 30 mg/l ampicillin. When the OD₆₀₀ reached 0.6, the inoculant was transferred to a 500 ml flask, in which it was further cultured until the OD₆₀₀ again reached 0.6. At this point, 0.4 mM isopropyl β -D-thiogalactoside (Sigma-Aldrich, Saint Louis, USA) was added to the culture medium to induce recombinant protein expression, and the inoculant was thereafter cultured for 18 h. The bacterial cells were then centrifuged at 2500 g for 15 min, where the bacterial cells were harvested through the precipitate. The precipitated bacterial cells were washed in 20 mM Tris-HCl (pH 8.0), resuspended in 70% ethanol, and incubated at 60 °C for 2 h, whereafter the samples were centrifuged to remove the bacterial cell debris, the supernatants containing the proteins were collected. The proteins were then precipitated over night by the addition of two times their volume of 1.5 M NaCl. The protein precipitate was collected by centrifugation at 14,000 g for 30 min, and followed by washing with *n*-butanol to remove residual bacterial lipids. Then, 0.1 M Tris-HCl (pH 8.5) containing 6 M urea was used to solubilize the pellet, and the samples were stored at 4 °C before further purification, applying reversed phase-high performance liquid chromatography (RP-HPLC; methods description see below). Before the purification on HPLC, multi-step dialysis was performed to remove the access urea using 20 K and 25 K dialysis tubes against water. Protein fractions collected were lyophilised and thereafter dissolved in 70% ethanol before they were analysed by size exclusion (SE)-HPLC (described below) to evaluate the purity of the extracted α -gliadin proteins.

2.1.3. Peptide design and production

From the primary sequence of the same α -gliadin as was used for in

vitro (*E. coli*) production (www.uniprot.org under the accession number Q9ZP09 and at European Molecular Biology Laboratory under the accession number AJ130948 [45]), five peptide sequences, see Table 2, were selected based on their potential aggregation/clustering behavior. Two of the peptide sequences (“Q” & “P”) represents repetitive motifs of glutamines and prolines that have homologues counterparts in most gliadins [46]. Peptide sequence “P” only includes the most reoccurring patterns found in the repetitive sequences of α -gliadin, hence describing these sections from a low complexity perspective [47]. The three remaining sequences represent sections with several hydrophobic amino acids and cysteine, known to participate in intra-disulphide bonds in most α -gliadins [13,14,48]. TAG Copenhagen A/S (Copenhagen, Denmark) produced the selected peptides of 99% purity (mixed with trifluoroacetic acid (TFA) salt). The terminal amino acids were capped with acylate or amide groups [49], except for the C-terminal in peptide “A” which terminal corresponds to the C-terminal in α -gliadin, to avoid interference from electrostatics related to the terminal amino acids.

2.1.4. Peptide sequences

The complete amino acid sequence (supplementary A) of the α -gliadin and the peptide sequences were used in the MC simulations to determine the structural properties of α -gliadin and the derived peptides. The α -gliadin contains 268 amino acids with a positive net charge of one. Studies have previously been conducted on this protein, characterising its existence and primary structure experimentally [45], and it has also been used in a simulation study [18] by the authors of this paper. The protein sequence is considered as representative for gliadins in general due to the close relationship and similarities among most gliadins [12]. Similarly, as for the *E. coli* produced gliadins, the N-terminal section of 20 amino acids was omitted, because these amino acids are known to code for a transport peptide that is removed during synthesis [50,51].

2.2. Secondary structure and hydrodynamic radius estimation of in vitro produced α -gliadin

The secondary structure and the size of the in vitro produced α -gliadins were evaluated by spectral analysis through Circular Dichroism (CD), Fourier Transform Infrared Spectroscopy (FTIR), and Dynamic Light Scattering (DLS). A Chiracscan from Applied Photophysics was used for CD measurements performed at 25 °C using a cell with 1 mm optical path length. Measurements were done in the spectral region of 190 to 260 nm with 0.5 nm intervals, a measurement time of 0.5 s per point and a bandwidth of 1 nm, according to Josefsson et al. [52]. An average of five scans were used after background subtraction. The Bestsel software [53] was used to estimate the secondary structure of the in vitro produced α -gliadins in water and ethanol solution.

The FTIR analysis was conducted in a PerkinElmer Spectrum 2000 FT-IR spectrometer (PerkinElmer) equipped with a single reflection ATR accessory, Golden Gate from Specac. The sample was dried for at least two days in a desiccator with silica gel before being analysed. The IR absorption spectrum was recorded in the ATR mode in the region 600 and 4000 cm^{-1} , with a resolution of 4 cm^{-1} . The FTIR spectrum was resolved into 9 Gaussian peaks in the amide I band (1580 to 1700 cm^{-1}) using fixed peak positions, and assignments of these according to Cho et al. [54]. Before the fitting of the peaks, the FTIR data was deconvoluted using an enhancement factor (γ) of 2 and a smoothing filter of 70%, baseline-corrected to have a horizontal baseline and the total absorbance in this region was normalized to 1. The fitting was performed with the Fityk software [55].

The DLS was performed with a Malvern zetasizer nano series instrument. The samples were loaded into the instrument with a 1 mm cuvette at room temperature. The built-in software processed the scattering light and provided the hydrodynamic radius of the present particles in the sample according to the Stokes-Einstein equation [56].

Table 2

Primary sequences of selected peptides and their amino acid range in the α -gliadin primary sequence. Letters in black represent neutral/hydrophilic amino acids, yellow represent cysteine, green represents negatively charged amino acids, purple represents positively charged amino acids, and blue indicates amino acids considered to be hydrophobic, according to the Kyte-Doolittle or Rose index [54,55].

Peptide acronym	Primary sequence	α -gliadin residues
Q	QNPSQQQPQEQVPLVQQQFLGQQPFPP	10–40
P	PQYPQPQFPQPYPQPFQYPYQ	40–92
I	ILQQILQQQLIPCMDVVLQQHIAH	115–140
Y	YQLLQELCCQHLWQIPEQSQCQAIHKVVHA	150–179
A	ALQTLPAMCNVYIPPYCTITPFGIFGTN	240–268

2.3. HPLC analysis of gliadin-rich fraction, *E. coli* produced α -gliadin, and peptides

To understand the aggregation/clustering abilities of the α -gliadins, the gliadin-rich fraction and the peptides were examined by SE- and RP-HPLC. Wheat gluten proteins, to which the α -gliadins belong have traditionally been analysed by HPLC methods in a range of studies [57–59] and methods are well described [59]. As the wheat gluten profile in a genotype consists of hundreds of proteins that resembles each other, they are difficult to separate into single peaks when HPLC is carried out and also purification of the proteins is difficult and therefore, there are no standards available. Thus, quantification of separate proteins has never been used in the literature, but instead comparisons using relative chromatograms are utilized [57,59]. In the present study, we followed previously established HPLC methods [59] where chromatograms are relatively compared to characterize the proteins, and analyses methods for the different evaluations are described below. For both the HPLC analyses, treatments/extractions of gliadins and peptides were carried out using polar solvents, detergents, reducing, and chaotropic agents, according to Rasheed et al. [27]. Thus, to samples of 1 mg of the gliadin-rich fraction, 1 ml of five separate extraction buffers; i.e. 1) 70% ethanol, 2) 50% isopropanol, 3) 50% isopropanol with 1% dithiothreitol (DTT; 95 μ l sample volume), 4) 50% isopropanol with 0.5% sodium dodecyl sulfate (SDS) (95 μ l sample volume), and 5) urea 6 M with 1% DTT and 0.5% SDS, were subsequently added [27]. Ethanol and isopropanol are commonly used to extract gliadins as monomeric proteins [10,27,58] and as being less polar than water, it influences the hydrophobic interactions of the protein. DTT disrupts covalent disulphide bonds [60], and SDS is a surfactant that can break secondary bonds including those in clusters related to hydrophobicity [61], while urea denatures the protein [62]. The propanol extraction samples were heated in a 60 °C oven for 30 min, and the urea treated samples were heated at 100 °C for 5 min. At the end of the treatments, the samples were vortexed and centrifuged at 12000 rpm for 4 min, and the supernatant was transferred to 200 μ l inserts in vials for HPLC analyses.

For the peptide samples, the trifluoro acetic acid (TFA) salt was removed through dialysis. Thus, 10 mg of each peptide was dissolved in duplicates into 1 ml of milli-Q water, which were then transferred to dialysis tubes with a 2000 MWCO size cut-off. The samples were dialysed in 1 l of milli-Q water, which was changed after 4 h and further dialysed overnight (minimum 8 h). The samples (100 μ l sample volume) were then diluted in each of the same five extraction buffers as was used for the gliadin-rich fraction, and the samples were thereafter prepared for HPLC analyses as described above.

A mixture of the peptides I, Y, and A was also evaluated by HPLC in order to understand if and how different peptides formed cross-links with each other. Approximately 7 mg of the individual peptides were weighed in duplicates and mixed with 1 ml water containing 1% DTT, hindering the peptides from forming disulphide bonds. The mixtures were then vortexed, followed by dialysis in the same manner as described above, to remove the DTT and to allow cross-linking. The samples were treated with the five extraction buffers and prepared for HPLC analyses as described above.

A Waters Alliance e2695 separation module HPLC system with a DAD Waters 2998 photodiode array detector (Waters, Milford, USA) was used to perform both SE- and RP-HPLC analysis. Data was collected and analysed by Empower software (Waters, Empower3). UV signals were recorded at 210 nm, and UV scans were collected at 190–220 nm.

A Biosep-SEC 2000 LC Column 300*7.8 mm (Phenomenex, Torrance, USA) was used with an isocratic flow rate of 500 μ l per minute when performing SE-HPLC. Two different solutions in equal proportions (50% each) were used as elution buffers; acetonitrile with 1% TFA and water with 1% TFA. The injection volume was 20 μ l per sample, and the absorbance at 210 nm was recorded for 30 min.

For RP-HPLC, a Phenomenex Synergi 4 μ m Hydro-RP 80 Å LC Column 250*4.6 mm (Phenomenex, Torrance, USA) was used at a flow rate of 700 μ l per minute. The injection volume was set to 20 μ l. Absorbance at 210 nm was recorded. For the HPLC gradient, elution buffers of 18,2 M Ω water with 1% TFA (mobile phases A) and acetonitrile with 1% TFA (mobile phase B) were used. Initially, 10% B was used, followed by a linear increase to 90% B at 50 min, which completed the gradient. Thereafter the gradient decreased linearly back to 10% B after 55 min, and then kept at 10% B to 60 min.

Additionally, SE-HPLC analyses on the *E. coli* produced gliadins were carried out as described above and RP-HPLC was carried using a SUPELCO column (discovery bio wide pore C8, 5 mm 25 cm \times 4.6 mm, catalogue no. 568323–4; Supelco, Sigma-Aldrich, Saint Louis, USA). A SUPELCO pre-column (guard column, discovery bio wide pore C8, 5 mm, 2 cm \times 4.0 mm) was used together with the main column. The solvent system for elution in RP-HPLC was based on two solvents: water (A) and gradient grade acetonitrile (B) containing 0.1% TFA. The solvent flow rate was maintained at 0.8 ml min⁻¹, and the temperature of the column was maintained at 70 °C at a gradient flow of 28–72% for 1–40 min extraction time. Thus, initially 28% of B was used, followed by a linear increase to 72% B at 40 min, a linear decrease back to 28% B at 50 min and then kept at 28% to 60 min.

2.4. Statistical analysis and plotting

All HPLC chromatograms were blank-separated and integrated according to a valley-to-valley procedure and plotted. All plots and calculations were performed with the statistical software R version 3.4.4 “Someone to Lean On” [63] with the package ggplot2.

2.5. The coarse-grained model

The monomers of the proteins, i.e. the amino acids, were represented by hard spheres (beads) and connected via harmonic bonds. The bead radius was set to 2 Å providing a realistic contact separation between the charges and an accurate Coulomb interaction, including the hydration layer. For non-bonded spheres, a short-ranged attractive interaction as well as electrostatic interactions were used. The simulations were performed at constant pH with point charges. The beads were positive, negative, or neutral, depending on the nature of the amino acid that it represents. Furthermore, the beads could carry a hydrophobic or polar property, were beads representing amino acids with a Kyte-Doolittle index [64] value >0 (I,V,L,F,C,M,A) were considered to be hydrophobic, as illustrated in Appendix B, Fig. A1.

The total potential energy of the simulated system contains non-bonded and bonded contributions, and is given by:

$$U_{tot} = U_{nonbond} + U_{bond} = U_{hs} + U_{el} + U_{short} + U_{hydrophob} + U_{bond} + U_{angular}, \quad (1)$$

where the non-bonded energy is assumed to be pairwise additive according to:

$$U_{nonbond} = \sum_{i<j} u_{ij}(r_{ij}), \quad (2)$$

where $r_{ij} = |\mathbf{R}_i - \mathbf{R}_j|$ is the center-to-center distance between two monomers, and \mathbf{R} refers to the coordinate vector. The hard-sphere potential, U_{hs} , is used to take account of the excluded volume, and is given by:

$$U_{hs} = \sum_{i<j} u_{ij}^{hs}(r_{ij}), \quad (3)$$

which sums up over all amino acids. The hard-sphere potential, $u_{ij}^{hs}(r_{ij})$, between two monomers in the model is given by:

$$u_{ij}^{hs}(r_{ij}) = \begin{cases} 0, & r_{ij} \geq R_i + R_j \\ \infty, & r_{ij} < R_i + R_j \end{cases} \quad (4)$$

where R_i and R_j denote the radii of the beads. The electrostatic potential (U_{el}), is given by an extended Debye-Hückel potential according to:

$$U_{el} = \sum_{ij} u_{ij}^{el}(r_{ij}) = \sum_{i<j} \frac{Z_i Z_j e^2}{4\epsilon_0 \epsilon_r} \frac{\exp[-k(r_{ij} - (R_i - R_j))]}{(1 + kR_i)(1 + kR_j)} \frac{1}{r_{ij}}, \quad (5)$$

where e is the strength and Z (positive, negative, or neutral) is the type of the elementary charge, k is the inverse Debye screening length, ϵ_0 is the vacuum permittivity, and ϵ_r is the dielectric constant for water 78.4 [65]. The used screening length was 7.9 Å, corresponding to 150 m salt in a 1:1 saline solution [65]. The short-ranged attractive interaction between the beads is included through an approximate arithmetic average overall amino acids, given by:

$$U_{short} = - \sum_{i<j} \frac{\epsilon}{r_{ij}^6}, \quad (6)$$

where ϵ reflects the polarizability of the proteins and thus sets the strength of the interaction. In this model an ϵ of $6 \cdot 10^3$ kJ/mol was applied, which corresponds to 0.6 kT at the closest distance, which follows previous IDP studies [66]. The hydrophobic potential (U_{hpob}) is similar to the short-ranged potential but involves only the hydrophobic beads:

$$U_{hpob} = - \sum_{i<j} \frac{\epsilon_{hpob}}{r_{ij}^6}. \quad (7)$$

In this study, two different ϵ_{hpob} were used to describe in which interval the clustering abilities of the α -gliadin and the peptides occur; 2.5 ($2.5 \cdot 10^4$ kJ/mol) and 3 kT ($3.0 \cdot 10^4$ kJ/mol). A harmonic bond potential, U_{bond} , connects the beads in the protein:

$$U_{bond} = \sum_{i=1}^{N-1} \frac{K_{bond}}{2} (r_{i,i+1} - r_0)^2, \quad (8)$$

where K_{bond} describes the stiffness of the bond, and r_0 describes the equilibrium distance. These were set to $K_{bond} = 0.4$ N/m and $r_0 = 4.1$ Å in accordance with Cragnell et al. [66]. The angular potential between three consecutive beads is described by a harmonic potential (U_{angle}):

$$U_{angle} = \sum_{i=2}^{N-1} \frac{K_{angle}}{2} (a_i - a_0)^2, \quad (9)$$

where a_0 is the equilibrium angle (180°), which represents an extended chain, and K_{angle} is the bond's flexibility. In this work, K_{angle} was set to 0 N/m and 0.0008 N/m, when ϵ_{hpob} is 2.5 kT and 3 kT, respectively, which corresponds to an α -gliadin with a radius of gyration (R_g) of approximately 40 to 43 Å, corresponding to other reports [67,68].

2.6. Simulation details

The MC simulations were performed with the Metropolis algorithm in the canonical ensemble; thus, at a constant amount of particles (N), constant volume (V), and constant temperature (T). The latter was set to at 298 K [69]. The simulation details were similar to those describing the self-association of Statherin [66,70]. Five chains were placed in a simulation box with the size of 289.5 to 307.5 Å, representing a box volume corresponding to 1 mg/ml protein concentration. Periodic boundary conditions were applied in all directions. The chains were allowed to perform single bead, pivot, slithering, full chain, and cluster moves during the MC iterations, with a frequency of 0.8, 0.05, 0.03, 0.05, 0.03, and 0.07 in accordance with previous studies [70]. The cluster moves relocate an entire chain and surrounding chains within a radius of 40 Å from the centre of mass. The simulations were equilibrated using $3 \cdot 10^5$ iteration turns followed by a production run involving $3 \cdot 10^6$ iterations divided into ten subdivisions. To ensure that the simulations were sampled accurately, the R_g values' probability distribution functions were analysed. The R_g values' uncertainty was based on the standard deviation of the total mean, and the mean from the ten subdivisions, as described in Cragnell et al. [24]. A snapshot with all particle coordinates was saved every 1000 iteration. The simulation software used was an extended version Molsim v. 4.8.8. [69,70].

2.7. SAXS analysis

Average shape characteristics of the simulated α -gliadin and derived peptides were described by a wavelength function derived from simulated small-angle X-ray scattering intensity data, depicted in a Kratky plot [71]. As described previously [66], the structure factor for a system containing N identical scattering objects can be obtained as:

$$S(q) = \left\langle \frac{1}{N} \left| \sum_{j=1}^N \exp(iq \cdot r_j) \right|^2 \right\rangle. \quad (10)$$

The total structure factor can further be decomposed into partial structure factors given by:

$$S_{ij}(q) = \left\langle \frac{1}{(N_i N_j)^{1/2}} \left| \sum_{i=1}^{N_i} \exp(iq \cdot r_i) \right| \left| \sum_{j=1}^{N_j} \exp(-iq \cdot r_j) \right| \right\rangle. \quad (11)$$

The total and the partial $S(q)$ are related through:

$$S(q) = \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \frac{(N_i N_j)^{1/2}}{N} S_{ij}(q). \quad (12)$$

For a point scatterer, the form factor is constant, inferring that the scattering intensity is proportional to the structure factor. For identical homogeneous spheres, the scattering intensity can be expressed as the product of the form factor and the structure factor, where the form factor corresponds to the intra-particle, and the structure factor to the inter-particle interference. The form factor of a point scatterer is constant, therefore the scattering intensity is proportional to the structure factor. Hence, the calculated structure factor for the point scatterers corresponds to the system's scattering intensity. If the system is composed of a single protein chain, the calculated scattering profile is due to intrachain interference only, hence, it is the protein form factor. A Kratky plot is obtained by plotting $(qR_g)^2 I/I_0$ versus qR_g .

2.8. Contact analysis

A contact analysis was performed between the beads from different chains, and they were considered to be in contact if the distance was ≤ 5 Å between two beads.

2.9. Asphericity analysis

The shape of the cluster was analysed through the asphericity, using R_g in three dimensions similar to what is described in the works of Rieloff et al. [70,72] and calculated according to the relation described below, where a value equal to zero indicates a perfect sphere and one a stiff rod:

$$a_s = \frac{(R_1 - R_2)^2 + (R_2 - R_3)^2 + (R_3 - R_1)^2}{2(R_1 + R_2 + R_3)^2}. \quad (13)$$

2.10. End-to-end distance calculation of tails

The end-to-end distance (R_{ee}) has been used to describe the effect of hydrophobicities on the structure of α -gliadin and the different chain sections. The primary sequence of α -gliadin was divided into three sections: (A) involves amino acid 1 to 100 and corresponds to the polar N-terminal of the protein, (B) the hydrophobic pseudo core involving amino acid 100 to 200, and (C) the hydrophobic and polar N-terminal involving amino acid 170 to 270, which overlaps with section (B). Hence in total 100 amino acids for each section.

2.11. Prediction of clustering and liquid-liquid phase separation

Clustering prediction, based on an algorithm derived from aggregation in amyloids called "AGGRESKAN", was performed on the primary sequence of the α -gliadin. The algorithm detects "Hot Spots" within the sequence with a higher propensity for cluster formation [73]. The LLPS of α -gliadin, the peptides, and the 20 glutamine residues ("Poly Q") derived, were predicted using the trained algorithms of "PSPredictor" [71].

2.12. Statistical details and rendering

The autocorrelation functions, and distribution of the individual chains' R_g were plotted with the R package "bio3d" [74]. The 3D rendering graphics were made with the molecular visualisation software VMD, with tachyon in-memory rendering settings [75].

3. Results

3.1. Characteristics of the gliadin-rich fraction

Analyses of the gliadin-rich fraction with SE-HPLC resulted in two peaks at 650- and 700-s retention time (with relative AU of 0.2–0.3 and 0.7–1.0, respectively), when solvents without reducing agent (70% ethanol, 50% isopropanol, 50% isopropanol + SDS) were used, whereas upon addition of solvents that reduce disulphide bonds (50% isopropanol + DTT, Urea + SDS + DTT), only one peak (relative AU 0.5) was observed at 700-s retention time, see Fig. 1a. The presence of the two peaks at non-reducing conditions indicates that both polymeric and monomeric protein are present in the gliadin-rich fraction. Addition of a reducing agent in the solvents resulted in polymers dissociating to monomers and thus only one peak was obtained.

3.2. Characteristics of the *E. coli* in vitro produced α -gliadin

3.2.1. Production and purification

A PMR191 plasmid harbouring the α -gliadin gene was successfully transformed to BL21(DE3) colony strains of *E. coli*, as was confirmed by polymer chain reaction (PCR; Fig. 2a) showing the amplified genes as a single band in R1–R4. After that, the total content of proteins obtained from the transformed *E. coli* cells was isolated. An evaluation of the content revealed the presence of a 40 kDa band of the transformed α -gliadin monomeric protein (marked by an arrow) although additional proteins were also present shown as additional bands on the gel (Fig. 2b). Therefore, the α -gliadin was purified through molecular analyses i.e. by the use of an RP-HPLC preparative column and dialysis, which resulted in a super purified fraction of an α -gliadin, shown as a single band (marked by an arrow), and no presence of additional bands by gel electrophoresis (Fig. 2c).

3.2.2. Characterisation of the secondary structure and size of the *E. coli* in vitro produced α -gliadins

SE-HPLC of the purified α -gliadin resulted in a significant peak (0.5 AU) at around 1250-s retention time (20 min), with low intensity (0.05 AU) peaks/shoulders before the major peak (from around 800 s; Fig. 1b). The large peak constituted 96% of the total area, while the corresponding number of the low-intensity peaks/shoulders constituted only 4%. Thus, the major large peak of the chromatogram most likely displays the gliadin in its monomeric form (retention time correspond to the gliadin peak of gluten reported previously [57]), and the low-intensity peaks and shoulders correspond to its aggregated forms as dimers and polymers. The RP-HPLC analysis confirmed the presence of a monomeric gliadin with a distinct peak (0.5 AU) close to 1400-s retention time, and with additional low levels of aggregated forms of the gliadin (≤ 0.1 AU) as additional small peaks at lower (800–1300 s) and higher (1500–1700 s) retention times (Fig. 1c).

When the *E. coli* produced and purified α -gliadin solubilized in water and ethanol were evaluated by CD, a spectra (Fig. 3a), with the typical dip at 205 nm for α -helix structures [76] was visible in both cases. Also, secondary structure calculation with the Bestsel software indicated a high degree of α -helix structures ($\geq 50\%$) in the samples, although the level of uncertainty in the predictions was high, making this number uncertain.

The FTIR analyses of the a dry sample of the *E. coli* produced and purified α -gliadin resulted in a fitting with nine peaks (Fig. 3b), which were used to calculate the relative content of the secondary structures in the sample. Thus, a relative content of 48 and 14%, respectively, was found for strongly (peaks 1 and 2) and weakly hydrogen-bonded (peaks 3 and 8) β -sheets. The relative content of β -turns was 15% (peaks 7 and 9), and the relative content of α -helix and unordered structures was determined to be 23% (peaks 4–6). A wet sample of the *E. coli* produced and purified α -gliadin sample was also analysed with FTIR, and the results indicated a higher degree of α -helix structures in the wet than in the

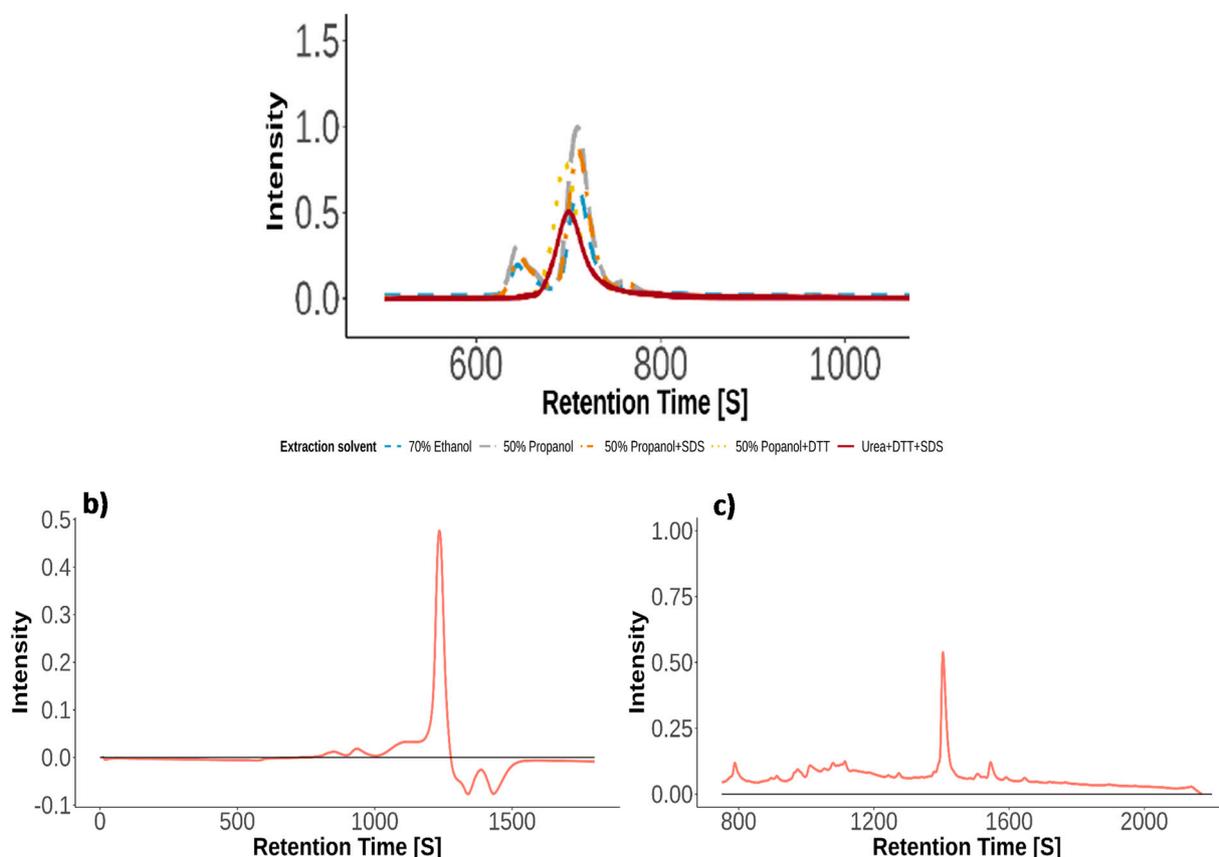


Fig. 1. HPLC chromatograms from (a) SE-HPLC of gliadin-rich fraction extracted in 70% ethanol (blue dash), 50% propanol (long grey dash), 50% propanol with 0.5% SDS (orange dot-dashed), 50% propanol with 1% DTT (yellow dotted), or 6 M Urea +1% SDS and 1% DTT (red solid). The values are mean values from duplicates, and (b) SE-HPLC of an *E. coli* produced and purified α -gliadin sample, and (c) RP-HPLC of an *E. coli* produced and purified α -gliadin sample. In (b-c) after solubilizing the sample in ethanol; the major narrow peak in both chromatograms indicates a single monomeric α -gliadin.

dry sample, although the content of water in the sample disturbed the FTIR signal, making the results uncertain.

The DLS analysis of the *E. coli* produced and purified α -gliadins revealed highly different peak positions for various ethanol dilutions (0.41 mg/ml and 0.82 mg/ml) of the sample (Fig. 3c). In general, the more dilute system resulted in peaks at lower hydrodynamic diameter (2.5, 3.5, and 7.5 nm) than in the more concentrated system. The latter showed peaks at 6.5 and 10 nm, while even resulted in a peak above 140 nm hydrodynamic diameter (corresponding to an R_g of 90.3 nm), indicating aggregation of the *E. coli* produced and purified α -gliadins.

3.3. Aggregation, clustering and polarity characteristics of the peptides

All evaluated peptides displayed intensity peaks when analysed with SE-HPLC after extraction with various solvents, although not all of the peptides displayed peaks for all extraction buffers applied, as shown in Fig. 7. Peptide A (Fig. 4a) resulted in a major peak (2.0–2.5 AU) at 1250–1300-s retention time, independent of extraction buffer used, and a smaller peak (0.4–0.6 AU) at 1100-s retention time for all extraction buffer applied, except urea + SDS + DTT. The lack of the small peak in the sample with urea + SDS + DTT buffer, indicates that urea disrupts the cross-links between A peptides, resulting in only monomeric peptide A when this extraction buffer is used. The use of the propanol and propanol + SDS extraction buffers resulted in an additional peak (0.5 AU) at 1200-s retention time. Thus, cross-links responsible for the formation of this extra peak were disrupted by ethanol as well as by urea, indicating the possibility that bonds, of other type than disulphide bonds, exist that are responsible for the cross-links of the A peptides. Peptide I (Fig. 4b) resulted in a broad early peak ((0.3–0.6 AU) at 600-s

retention time when extraction buffers not containing SDS were applied, indicating formation of some type of cross-links. For treatment with the urea + SDS + DTT a clear peak (0.3 AU) was instead revealed at 1000-s retention time, which may correspond to the monomeric form of the peptide without the presence of disulphide cross-links. For peptide P (Fig. 4c), a single dominant peak (2.0–2.3 AU) was obtained just before 1200-s retention time, independent of extraction buffer used, indicating that no disulphide bonds were present between the P peptides. Similarly, peptide Q (Fig. 4d) displayed a single monomeric peak (0.7–2.4 AU) slightly before 1200-s retention time, although with a smaller bulge (0.1 AU) preceding the large peak after 1000-s retention time, potentially indicating that this peptide form clusters or aggregates. Peptide Y (Fig. 4e) showed an early intensity peak (0.3–1.5 AU) around 600-s retention time when treated with ethanol, propanol, or propanol + DTT, followed by two peaks (0.6–0.9 AU) close to 800-s retention time when treated with ethanol, propanol, or propanol + SDS. The urea + SDS + DTT instead resulted in a large peak (2.1 AU) at around 900-s retention time, succeeded by a minor peak (0.1 AU). Also, the other extraction buffers resulted in peaks (0.4–1.5 AU) in the vicinity of 900-s retention time. Thus, the peptide Y seemed to form various types of cross-links between the peptides. A mix of the peptides A, I, and Y (Fig. 4f), resulted in three monomeric peptides represented by peaks (0.5, 0.1 and 0.4 AU, respectively) at 900-, 1000-, and 1300-s retention time when the urea + SDS + DTT extraction buffer was applied, corresponding to the single peptides (compare Fig. 4f with a,b and e). Additionally, a large peak (1.0 AU) at 600-s retention time was obtained when the propanol extraction buffer was applied, a peak (0.5 AU) at 800-s retention time was obtained when the propanol or propanol + SDS extraction buffer was used, and several peaks (0.4–2.0 AU) for most of

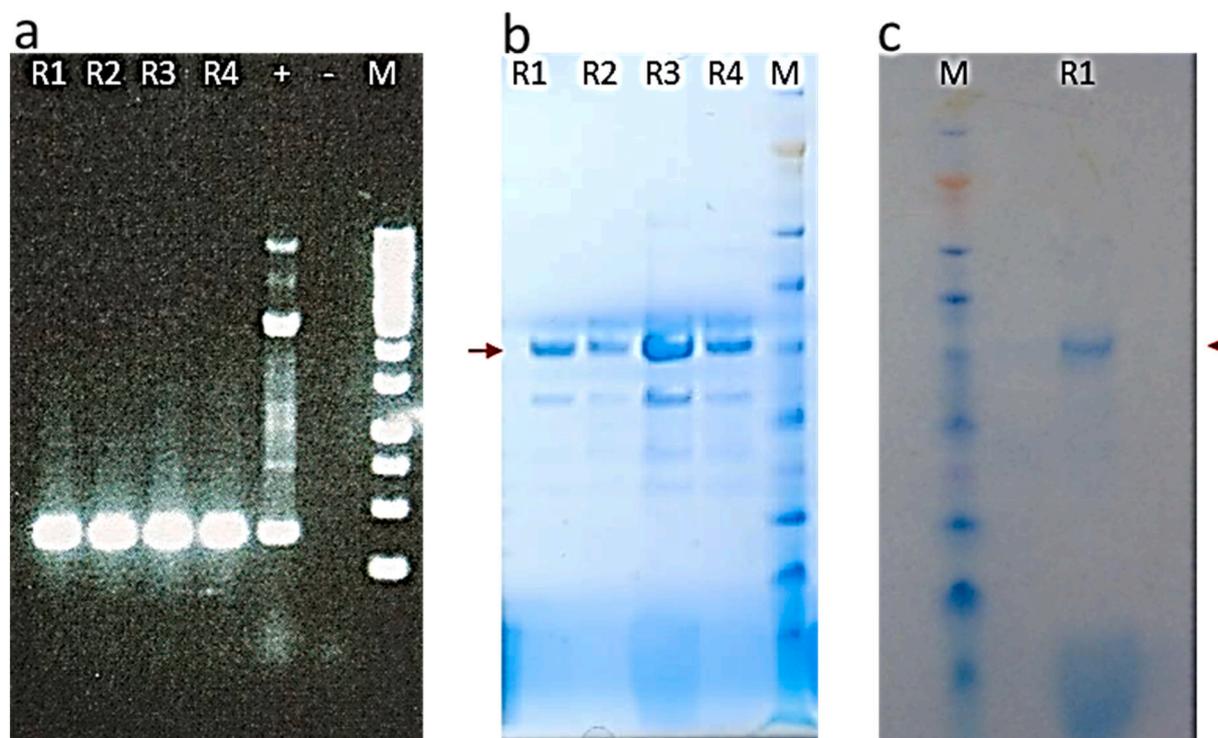


Fig. 2. Molecular analysis of BL21(DE3) pLysS *E. coli* strains transformed with α -gliadin gene. (a) R1–R4 = α -gliadin gene amplification from transformed BL21(DE3) *E. coli* strains replicate 1–4, + = pMR191 plasmid having α -gliadin gene used as a positive control, – = Blank, M = 1 kb ladder. (b) Extracted α -gliadin proteins (MW ca 40 kDa) from transformed BL21(DE3) *E. coli* strains resolved over gel electrophoresis, R1–R4 = α -gliadin gene amplification from transformed BL21(DE3) *E. coli* strains replicate 1–4, M = protein ladder. (c) α -gliadin protein (MW ca 40 kDa) isolated from transformed BL21(DE3) *E. coli* strains and purified through dialysis and RP-HPLC preparative column = R1, M = protein ladder.

the extraction buffers were obtained at around 900- and after 1200-s retention time, indicating monomeric forms to be present but also cross-links formed between various combinations of the peptides. However, comparing the peak pattern for the single peptides and mixtures of peptides show peaks at the same retention times, indicating similar di- or polypeptides to be formed independent on if only one or several different peptides are present (Fig. 4). Thus, our results indicated that the peptides favoured clustering with the same peptide sections over clustering among different types of peptides.

The RP-HPLC resulted in peaks for the peptides A, Y, P, and Q (Fig. 5), while the peptide I did not show any peaks. When propanol was used as the extraction buffer, peptide A (Fig. 5a) and Y (Fig. 5b) showed intensity peaks (0.1–2.0 AU) from 1250-s retention time until the program ended with the highest intensity close to 1500-s retention time. On the other hand, when the extraction buffer urea + DTT + SDS was used, a large peak (2.0 and 0.5 AU, respectively) at around 1700-s retention time appeared (Fig. 5a–b). For peptide P (Fig. 5c) and Q (Fig. 5d), a prominent peak (1.9–2.0 AU) and a double peak (0.8–2.0 AU), respectively, were visible after 1000-s retention time, independent of extraction buffer used.

3.4. Monte Carlo simulations of the cluster formation of α -gliadins

Both the shape of the curve of the Kratly plots (Fig. 6a) and the snapshot (Fig. 6d), showed that the formed clusters possessed both compact and extended parts. The protein became increasingly more compact when the strength of the hydrophobic potential was increased (Fig. 6a), therefore an angle potential was used to hinder unrealistically compact structures. Furthermore, the number of monomers per cluster increased slightly with an increased strength of the hydrophobic potential (Fig. 6b). The snapshots indicated two states of clustering for the α -gliadins: i) N-terminal based, resulting in extensive waving tails and ii) through cross-links in sections corresponding to the peptides Y, I, and A.

In the latter a condensed hydrophobic core was formed (Fig. 6d).

The simulations of clusters resulted in different R_{ee} for various regions of the α -gliadin chains. Furthermore, the hydrophobic potential effected the R_{ee} , specifically for the C region (amino acids 170–270) of the α -gliadin (Fig. 6c). Thus, the regions A (amino acids 1–100 of the primary sequence) and B (amino acids 100–200) resulted in an average R_{ee} of 60 and 48 nm, respectively, at the hydrophobic potential of 2.5 and 3 kT. The region C resulted in an average of 35 nm at 2.5 kT, while it collapsed to an average of 25 nm at a hydrophobic potential of 3 kT. The probability distribution function of R_{ee} for a single α -gliadin showed that the ensemble of possible conformations was differently distributed over the three regions, with a normal distribution for A and B, and a skewed distribution for C with the highest probability close to 25 nm (Fig. 6c). Hence, a more flexible structure is indicated for region A due to its wider distribution, while the hydrophobic character of region C resulted in the compactness of this region. Differences in R_{ee} between the regions correlated well with the number of defined hydrophobic amino acids in the sequence. Thus, the regions A, B, and C contained 22, 30, and 34 hydrophobic amino acids, respectively (Table 3). In principal, results obtained here corresponded well with results reported previously on other proteins [66].

3.5. Simulating cluster formation of peptides

The shape and clustering properties were altered for three (A, I, and Y) of the five investigated peptides (Table 2), when increasing the hydrophobic potential from 2.5 to 3 kT as depicted in Fig. 7. The reader should notice that for the latter potential, an angular potential is introduced into the model to prevent unrealistically compact structures. The Kratly plots of the three peptides (A, I, and Y) describe their development towards more globular structures with an increase in the hydrophobic potential (Fig. 7 a,b). Moreover, the aggregation probability increased for the three peptides with the increased hydrophobic

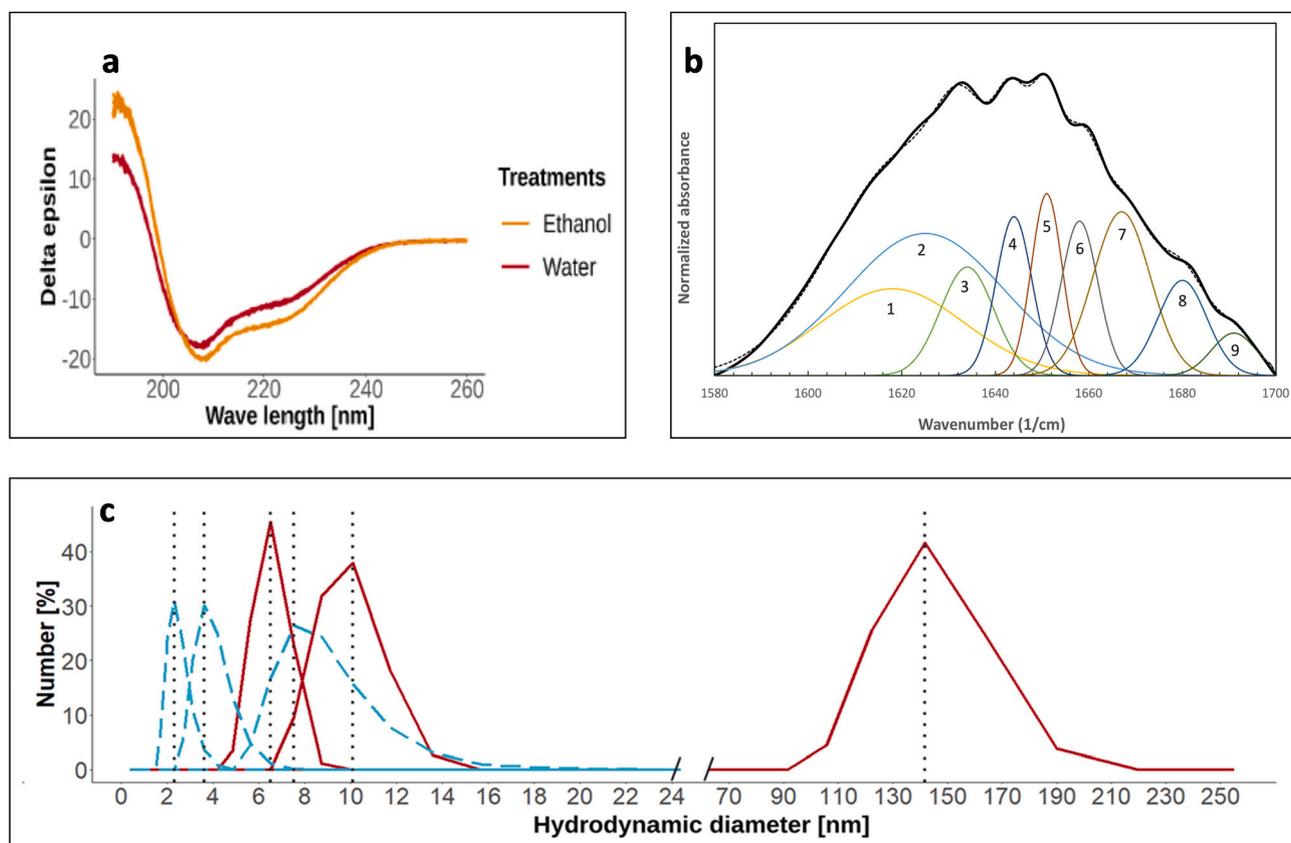


Fig. 3. Evaluation of the secondary structure of *E. coli* produced and purified α -gliadin by (a) circular dichroism (CD) spectra, (b) Fourier transform infrared (FTIR) absorbance (experimental (bold black curve), modelled (broken curve)) and (c) Dynamic light scattering (DLS). For the CD analyses, the gliadin was dissolved in ethanol and water, FTIR was carried out on a dry sample and DLS was carried out on a sample solubilized in 70% ethanol with concentrations of 0.82 mg/ml (solid red curve) and 0.41 mg/ml (broken blue curve). The resolved peaks from the FTIR correspond to the positions: 1618 (1), 1625 (2), 1634 (3), 1644 (4), 1651 (5), 1658 (6), 1667 (7), 1680 (8), and 1691 (9) cm^{-1} . In (c), the grey vertical dashed lines display the position of the curve peaks. Observe that the x-axis is truncated.

potential (Table 4). The autocorrelation functions for peptide A, I, and Y at higher hydrophobic potential (Appendix C), display similarities to previous simulation steps, indicating that the actual clustering effect occurred before 3 kT. Clustering of the peptides often occurs at their terminal ends or between multiple closely situated hydrophobic amino acids (Fig. 7). Clusters formed tend to develop a more spherical shape with the increasing number members in the cluster (Table 4, complete list in Appendix D).

3.6. Prediction of “Hot Spot” areas for clustering and propensity for liquid-liquid phase separation (LLPS)

By the use of the AGGREGSCAN analysis for prediction of aggregation in the α -gliadins, specific “Hot Spot” areas were verified (Fig. 7e). The largest “Hot Spot” area along the α -gliadin were found for the amino acids 115–140, 150–179, and 250–268, which correspond to the parts where the peptides A, I, and Y are located. Minor “Hot Spot” areas were identified on sections corresponding to those for peptides Q and P. Furthermore, a larger “Hot Spot” area was found in the N-terminal of the protein, and some small areas were found around the amino acid residue 200. The LLPS prediction indicated that the α -gliadin had a high probability of condensing into liquid droplets (see Fig. 7f), primarily originating from characters of its P and the Poly Q regions.

4. Discussion

The present study depicted unique features not previously described, behind clustering and polymer formation of gluten proteins (especially the α -gliadins), known as IDPs [18] and their possibility to form large

cross-linked networks [58]. The combination of experimental analyses and computer simulations of the proteins on various scales (from peptides to *E. coli* produced, purified α -gliadin, and gliadin-rich fractions), allowed a detailed understanding of the structural properties of the polymers formed between the α -gliadin units. Thus, the α -gliadins were found to be present both in monomeric and aggregated form, and the proportion was depending, to a large extent on, the storage solution (type and grade) of the proteins. Aggregation into polymers seemed to start with clustering of the proteins, although the clustering did not always lead to polymer formation with disulphide bonds. The clustering ability, and “Hot spot” areas for clustering, were found mainly in the hydrophobic areas of the α -gliadins, which were most prone to form aggregates through disulphide bonds. However, clustering was also found in the non-hydrophobic areas in the N-terminal region of the α -gliadins, where instead flocculants and phase-separation into liquid droplets were established.

Both the gliadin-rich fractions and the *E. coli* produced and purified α -gliadin, showed a distribution of clusters of the α -gliadins, although in various proportions in relation to amount of solvent used (dry versus wet and 0.41 versus 0.82 mg protein per ml solvent used). Here, the size of the α -gliadin in monomeric/dimeric state corresponded to the hydrodynamic diameter of 6.5 nm to 10 nm, which agrees well with previous experimental estimates of the size of the α -gliadin [67,68,77]. Oligomers of the α -gliadin were formed at a slow rate in solution, and to a higher extent if the protein was dried, but over time, this behavior resulted in a significant clustering and an extensive aggregation. The secondary structure of the α -gliadins, especially when they were in their monomeric form in solution, displayed a high proportion of disordered elements, which has also been previously reported for γ -gliadin [38,78,79].

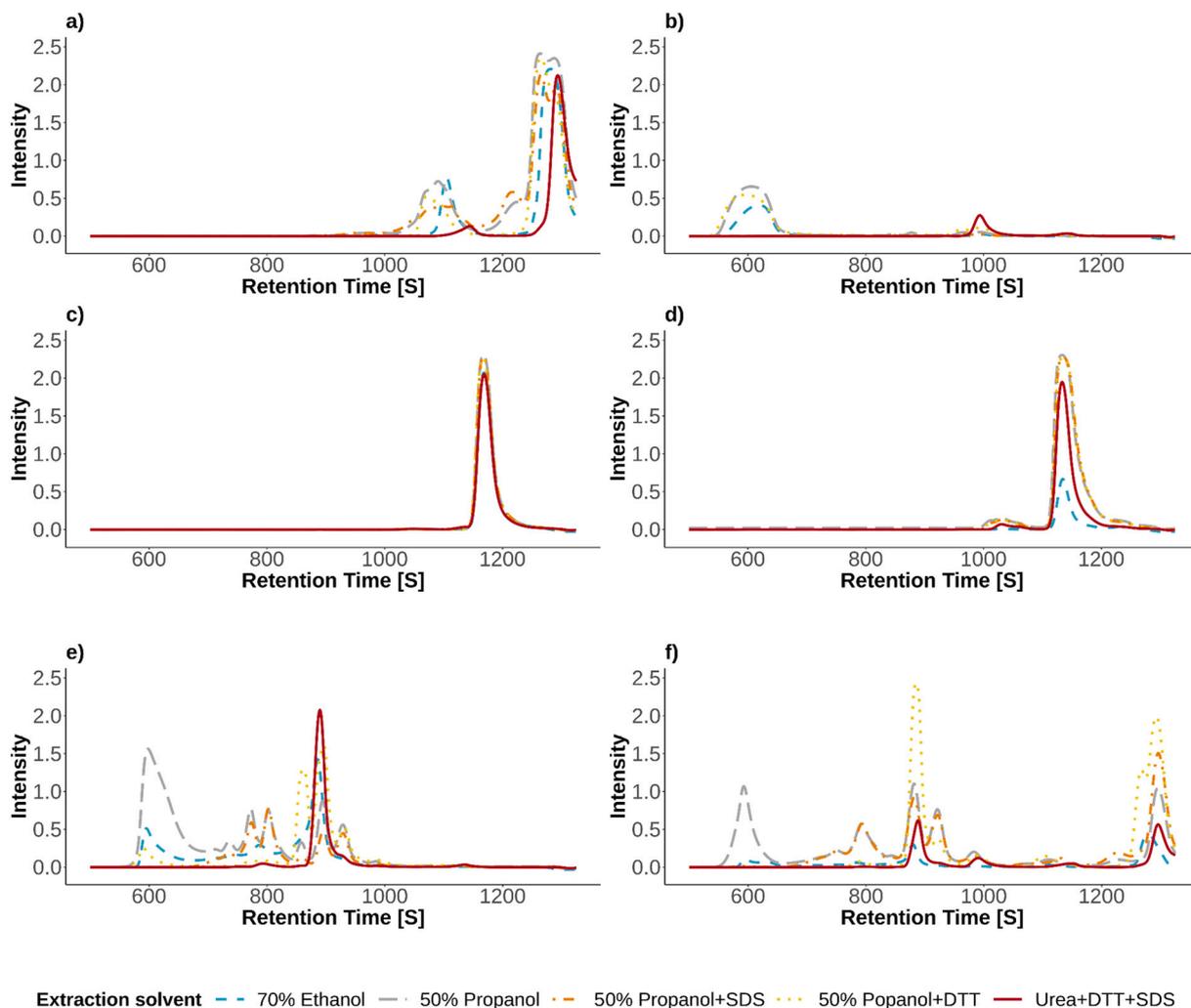


Fig. 4. SE-HPLC chromatograms of (a) peptide A, (b) peptide I, (c) peptide P, (d) peptide Q, (e) peptide Y, and (f) a mixture of the peptides I, Y, and A, extracted in 70% ethanol (blue dash), 50% propanol (long grey dash), 50% propanol with 0.5% SDS (orange dot-dashed), 50% propanol with 1% DTT (yellow dotted), or urea 6 M added with 1% SDS and 1% DTT (red solid). The values are mean values from duplicates.

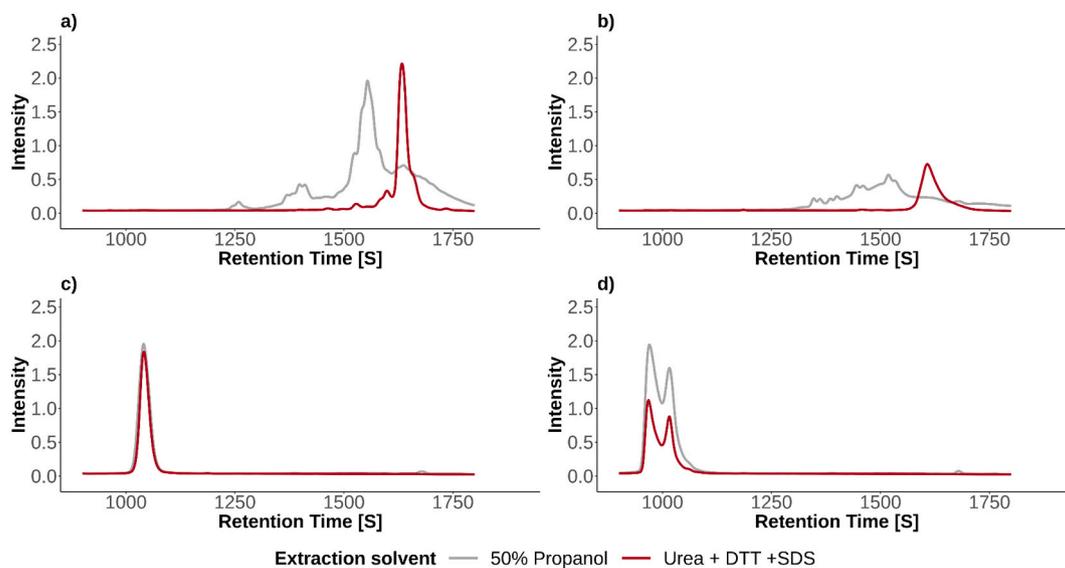


Fig. 5. RP-HPLC chromatograms of (a) peptide A, (b) peptide Y, (c) peptide P, and (d) peptide Q, extracted in 50% propanol, or Urea 6 M added +1% SDS + 1% DTT (shown as grey solid or red dashed lines). The values are mean values from duplicates.

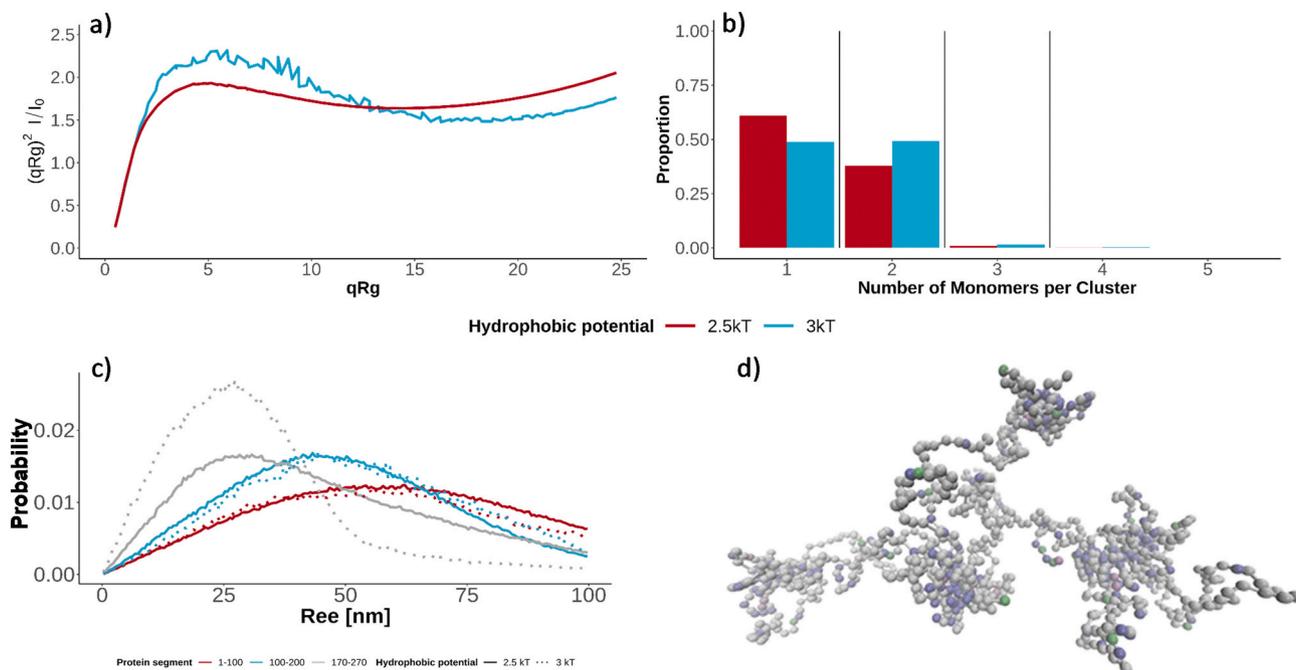


Fig. 6. Simulation of clustering ability of five α -gliadins; (a) Kratky plots at two different hydrophobic potentials, which indicates an overall extended structure with compact parts, (b) the cluster distribution from the simulations at the two different strengths of the hydrophobic potential, (c) end-to-end distance (R_{ee}) of gliadin sections 1–100, 100–200, and 170–270 amino acids (solid, dotted, and long dash lines) after cluster formation simulations, at different hydrophobic potential 2.5 kT and 3 kT (red and blue), and (d) an illustrative snapshot from the simulation of an α -gliadins cluster.

Table 3

Primary amino acid structure of three regions of the α -gliadin; A (amino acids 1–100 of the primary sequence), B (amino acids 100–200), and C (amino acids 170–270), and the summary of hydrophobic amino acids in the region (marked in green).

Region	Amino acids primary sequence	Summary of hydrophobic amino acids
A	V R V P V P Q L Q P Q N P S Q Q Q P E Q V P L V Q Q Q Q F L G Q Q Q P F P P Q Q P Y P Q P Q P F P S Q Q P Y L Q L Q P F P Q P Q L P Y S Q P Q P F R P Q Q P Y P Q P Q P Y S Q P Q Q P I S Q Q Q Q	22
B	Q Q Q Q Q Q Q Q Q Q Q Q Q Q I L Q Q I L Q Q L I P C M D V V L Q Q H N I A H G R S Q V L Q Q S T Y Q L L Q E L C C Q H L W Q I P E Q S Q C Q A I H K V V H I I L H Q Q K Q Q Q P S S V S F Q Q	30
C	Q I P E Q S Q C Q A I H K V V H A I I L H Q Q K Q Q Q P S S V S F Q Q P L Q Q Y P L G Q G S F R P S Q Q N P Q A Q G S V Q P Q L P Q F E E I R N L A L Q T L P A M C N V Y I P P Y C T I T P F G I F G T N	34

Along with the clustering and formation of oligo-/polymers (e.g. in lower solution or dry form) β -structures were formed. The formation of secondary structure in gliadins are considered to be related to disulfide bonds formed in the C-terminal, since mutants without cysteines and ω -gliadins lack stable α -helical and β -sheet content [17,80]. However, this study showed that hydrodynamic conditions had a clear impact on the structures of gliadins. In fact, this study indicated that α -gliadins showed coiled coils in aqueous solutions, and more rigid structures when dried.

The methodology used here, to analyze peptides of the α -gliadin (both experimentally and through simulations) revealed the clustering and aggregational abilities of the protein, section by section. Thus, the N-terminal sequence (i.e. with the peptides Q and P) displayed a general tendency of clustering, while aggregation by strong cross-links were basically lacking due to a high polarity, and lack of cysteines, as verified by both the experimental and the simulation results. The two mentioned peptides showed similarities to the gliadin p31–43 and p33mer peptides, respectively, which both have been shown to coalesce into nanostructures in water [19,81,82]. These two peptides (p31–43mer and p33mer) form aggregation-prone polyproline II structures in water, which are thought to be driven by hydrophobic interactions and high

content of proline and glutamine [19,81,83,84]. In the present study, only the Q peptide (of Q and P) was observed to form visible flocculants in water (data not shown). For peptide P, the divergence from the previous results [19,81,82] is most likely explained by the fact that the P peptide lacks the hydrophobic leucine residues, which were present in the p33mer [82,84]. Thereby, the P peptide was too polar to be able to cluster. The peptide Q in the present study, is most likely forming hydrophobic clusters, observed as the small bulge before the main peak of the SE-HPLC results. The size of the bulge might be a result of the elution buffer used, with acetonitrile being less polar than water, thereby partly dissolving the clusters (non-cross-linked) of proteins. The RP-HPLC results from this study described a potential spectra of oligomers and monomers for Peptide Q, due to gradual changes of the polarity of the elution buffer. Thus, the peptide Q seemed to interchange between an oligomeric and a monomeric state (resulting in two peaks while evaluated by SE-HPLC), although several different oligomeric states were also present as verified by both experimental and simulation results. Generally, the hydrophobic clusters formed by the peptide Q were weak (compared to those formed by peptides discussed below), indicating a lack of disulphide or other stronger bonds, as verified by the simulation results of the present study.

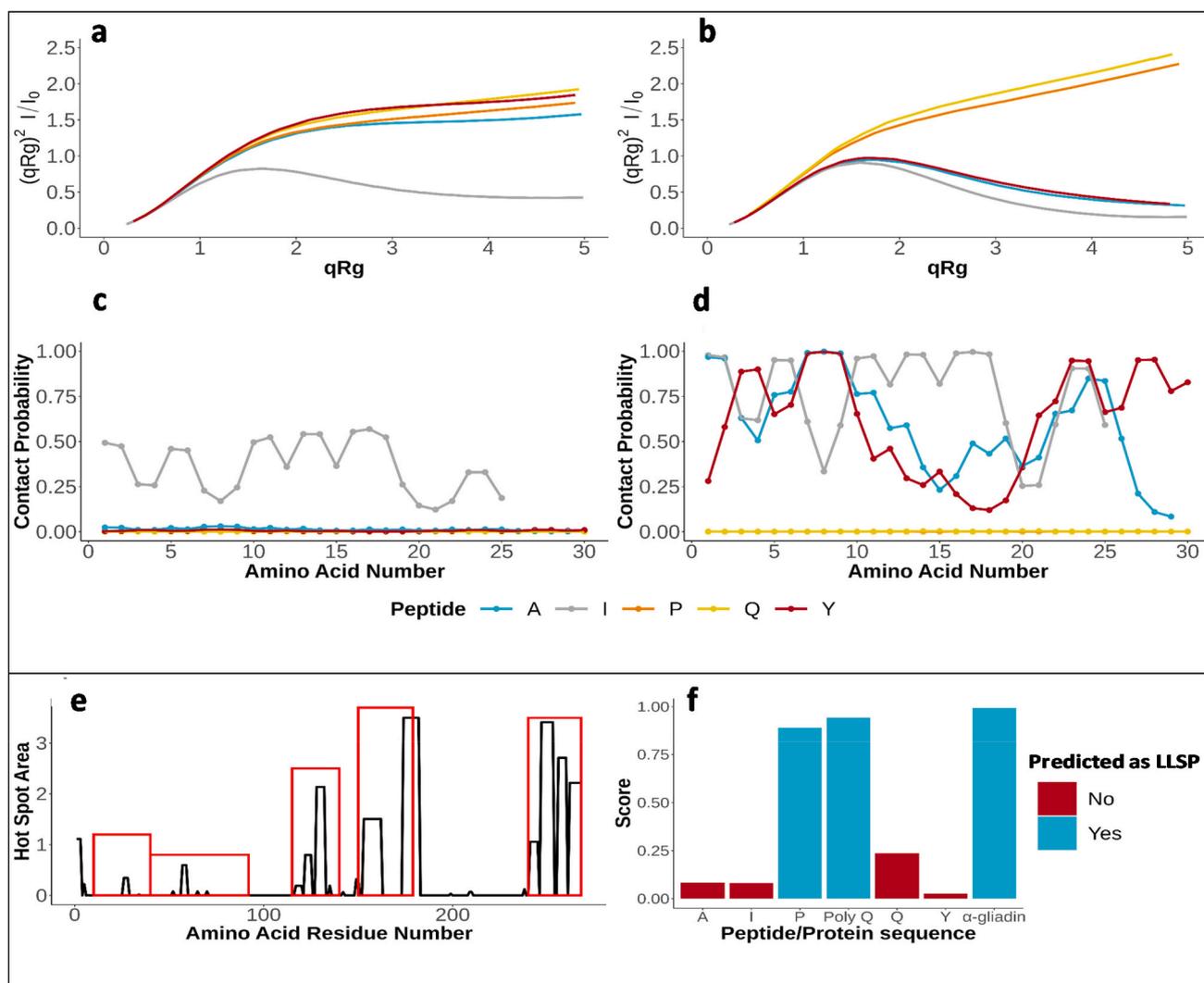


Fig. 7. Kratky plots that indicate (a, b) the shapes, and (c, d) contact probability for five different peptides (A, I, P, Q, Y) at the two different hydrophobic potentials (ϵ_{Hphob}) of (a, c) 2.5, and (b, d) 3 kT, and predicted (e) “Hot Spots” areas for clustering in α -gliadin (black line) with peptide sequences found in Table 2 as red rectangular areas, and (f) liquid-liquid phase separation (LLSP) behavior of the gliadin and its different protein sequence. The data in (a-d) displays that the peptides A, I, and Y increase their globular shape and forms larger clusters when the hydrophobic potential increases. Peptide Q and P showed a more opened size with increased hydrophobic potential, related to the increased angle potential.

Table 4
Cluster size probability, and corresponding asphericity index, is listed for each peptide at different kT. Only clusters with a probability of above 0.05 are listed, complete list is found in Appendix D.

Peptide acronym	Hydrophobic potential in kT	Cluster size	Asphericity index	Probability
A	2.5	1	0.4	0.9283
A	2.5	2	0.3	0.0651
I	2.5	1	0.36	0.3441
I	2.5	2	0.22	0.0542
I	2.5	3	0.13	0.0936
I	2.5	4	0.09	0.2567
I	2.5	5	0.09	0.2515
P	2.5	1	0.45	0.9885
Q	2.5	1	0.43	0.9851
Y	2.5	1	0.41	0.9555
A	3	5	0.11	1
I	3	5	0.04	1
P	3	1	0.49	0.9847
Q	3	1	0.47	0.9757
Y	3	5	0.07	1

The N-terminal part of the gliadin lacked the probability to form strong cross-links and showed weaker cluster ability than other parts of the α -gliadin, which could partly be the result of a low number of hydrophobic amino acids. However, the predicted self-associating properties in terms of LLPS formation were high in the N-terminal part of the protein, especially in the sequences corresponding to the peptide P and the poly Q. Simulations of this part of the α -gliadin resulted in the formation of a tail which could interact with other gliadin N-terminals, possibly through LLPS formation. Such interactions might result in an increased amount of weaker π - π interactions between aromatic residues, but might also initiate β -sheet aggregation, which has previously been reported for Alzheimer related proteins [82,84].

Both experimental (HPLC-results) and simulations results revealed that the peptides at the C-terminal part of the α -gliadin (peptides A, I, Y) were forming hydrophobic clusters. Furthermore, the HPLC-results clearly demonstrated that these peptides were forming cross-links through disulphide bonds. However, the peptide Y displayed a more complex pattern of formation of cross-links as compared to the other two, as revealed by the many different peaks obtained with various solvents. The fact that most of the peptides resulted in a single peak when 6 M urea + 1%SDS + 1%DTT was used as solvent, and the peptides

were analysed with SDS-PAGE, indicated that primarily disulphide bonds formed. However, for peptide Y, and the same solvent being used, RP-HPLC resulted in two peaks, of similar type as when 50% propanol was used as solvent, indicating the formation of also irreversible cross-links. These irreversible cross-links are most likely a lanthionine bond formation occurring between cysteines or between cysteines and serines [85], or less likely it could be histidine cross-links [86,87]. When mixing the three C-terminal peptides, the HPLC results indicated that they were primarily forming aggregates with their identical counterparts. Thus, polymerisation among gliadins seems to be biased to occur between similar sections of different gliadin molecules. These results correspond well with previous studies that indicated that the cysteines of the gliadins prefer to connect with their corresponding cysteines in the glutenins [88,89]. Simulation results on the entire α -gliadin sequences indicate that the C-terminal sequences preferred to interact with other α -gliadin C-terminals, and these are then forming a joint dense hydrophobic core. Hydrophobicity has in earlier studies been identified as an essential factor for internal cross-linking among gliadins and glutenins [22,36], which together with our results strengthen the conclusion that hydrophobicity is a vital driving force for the polymerisation of gliadin and glutenins.

The gliadins have been identified as IDPs, but their self-association into polymers seems to be an ordered procedure involving several defined steps resulting in a specific aggregated network. In a previous study, we were able to define the pre-determined process and events of specific value for internal cross-links of the α -gliadins in the wheat cell, to prevent from native polymer formation of these proteins [18]. Here, our simulation results indicated that the individual gliadin molecules in saline water form structures resembling the R1 shape model (globules and tadpoles) described by Das and Papu [90,91]. Thus, the N-terminal sequences of the α -gliadins resulted in a structural tail part of the protein and the C-terminal sequences form a hydrophobic pseudo core, similar to what is reported for elastin, which is also an IDP [92].

We also showed that individual tadpole-shaped gliadins have clustering ability in the hydrophobic C-terminal region leading to associations within their hydrophobic cores, which results in the formation of large clusters. In general, the gliadins showed interchanges between their monomeric state and their forms of hydrophobic clusters, as verified both from experimental data (e.g. HPLC and DLS) and simulation results, which also correspond well with previous studies [17,93]. The more polar part of the α -gliadins (i.e. the N-terminal part) form small arms that are reaching out of the cluster of gliadins. As described above, this part of the α -gliadin may have properties to create polyproline II aggregation and contribution of weaker $\pi \sim \pi$ interactions [94], which may result in phase separation into liquid-liquid droplets [93,95]. Previous studies have shown this behavior to correspond to an increase in the formation of β -sheets, a structure that we saw were increasing in dried gliadins. The increase in β -sheets might also lead to an increased opportunity to tightly pack the proteins previously discussed also for the hexagonal structure described for both gluten and gliadins [96,97]. In the cell, the gliadins are primarily cross-linked by intra-molecular disulphide bonds in a process previously described [18]. However, during thermal, mechanical, or chemical processing, these intra-molecular cross-links are broken, thereby opening-up for new ones to be formed. Here, we showed that inter-molecular disulfide cross-links are formed primarily between certain specific sequences inside the hydrophobic core and is mostly involving the peptides I and Y. However, aggregation behavior of the gliadins is most likely also influenced by the processing technology applied. Previous studies have shown that the gliadins are organised into supramolecular hexagonal structures when compression molded/pressed into sheets at higher temperatures [24,27]. Hexagonal formation has also been described for elastin [98–101]. These findings imply that temperature treatment on IDPs, such as the α -gliadin, may result in an ordered cross-linked network.

One aim of the present study was to evaluate opportunities for large scale fractionation/production of gliadins to be used in plastic materials.

Here, the *E. coli* system resulted in production of gliadins although also other types of proteins (most likely of *E. coli* origin) were produced simultaneously. Thus, further purification was required in order to obtain a pure α -gliadin sample. This purification was in this study carried out using RP-HPLC, which allows only small quantities. Furthermore, the *E. coli* system faced difficulties producing a higher quantity of the α -gliadin protein, leading to lower production than anticipated. Heterologous expression of wheat storage proteins such as for alpha gliadins had largely remained a challenge. The storage proteins of *Poaceae* crops such as wheat and maize differ from other proteins in their characteristics [43,58,102], and they are also exceptionally large. Also, previous attempts to use microbial systems to produce these types of proteins e.g. the α -gliadin homologous protein, Zein, and high molecular weight glutenin subunits (HMW-GS) of wheat, have resulted in low yields [103–106]. Tungekar et al. [107] experimented to understand the molecular mechanisms behind the challenges faced by the *E. coli* to express the LMW-GS and gliadins. The unusual structural properties and repetitive sequences, present in gliadin proteins probably affect their heterologous expression, which may result in unfavorable mRNA secondary structures and inhibit ribosome processivity through mRNA stem-loops [107–109]. Gliadins are clustering-prone proteins known to cause endoplasmic reticulum stress responses when synthesized in its wild type conditions, resulting in the expression of chaperones and foldases that assists in folding and packing [110]. In *E. coli*, the proteins are synthesized by ribosomes dispersed in the cytoplasm, resulting in that stress responses do not occur. The result might then be that other types of cellular stress responses appear that degrade the proteins or reduce the protein production. The present study did not investigate the exact reason for the low yield of α -gliadin from the *E. coli* system, although the reasons mentioned above are possible explanations for the low yield. However, the low yield together with the required HPLC purification to get rid of contaminants of the *E. coli* system, limited the possibility for more extensive analysis. To better upscale the production, a better understanding on how to produce aggregating proteins in in-vitro systems is a necessity, or an alternative host for the production need to be found.

5. Conclusion

The α -gliadins are present as monomers, dimers, and oligomers/polymers in solution and in the dry state. However, the proportion of monomers with less structural features is higher when suspended in water than when in dry state. Furthermore, a positive correlation exist between level of dry conditions and the amount of polymers. Oligomers and polymers are formed through diverse interactions between the gliadins. Hydrophobic interactions, formation of polyproline structures, and liquid-liquid phase separations are the main opportunities for the gliadins to form clusters. Thereafter, aggregations are formed which may result in cross-links with either disulphide or lanthionine bonds in the hydrophobic core. Different sections of the α -gliadin have different modes of clustering. Thus, C-terminal sequences, including the areas of the peptide Q and P, form clusters through disulphide or lanthionine cross-linking while areas of the Peptide P and the poly Q sequence show a high similarity to proteins that condense into phase separated liquid droplets. Furthermore, the C-terminal sequences result in high flexibility, which might result in coalescence through entanglements when in high concentrations. The N-terminal sequences are prone to hydrophobic clustering, with the preference to cluster with similar sections. The N-terminal sequences then aggregate with disulphide bonds, and the peptide Y is susceptible to the formation of dimers/oligomer by the formation of lanthionine bonds. Fig. 8 shows a possible model of gliadin molecules connected by disulphide and lanthionine bonds in their hydrophobic core and possible polyproline structures and liquid-liquid phase separations by their tails. Future studies should include experimental studies on longer peptide sequences and further simulations on the interaction between a higher number of gliadin molecules.

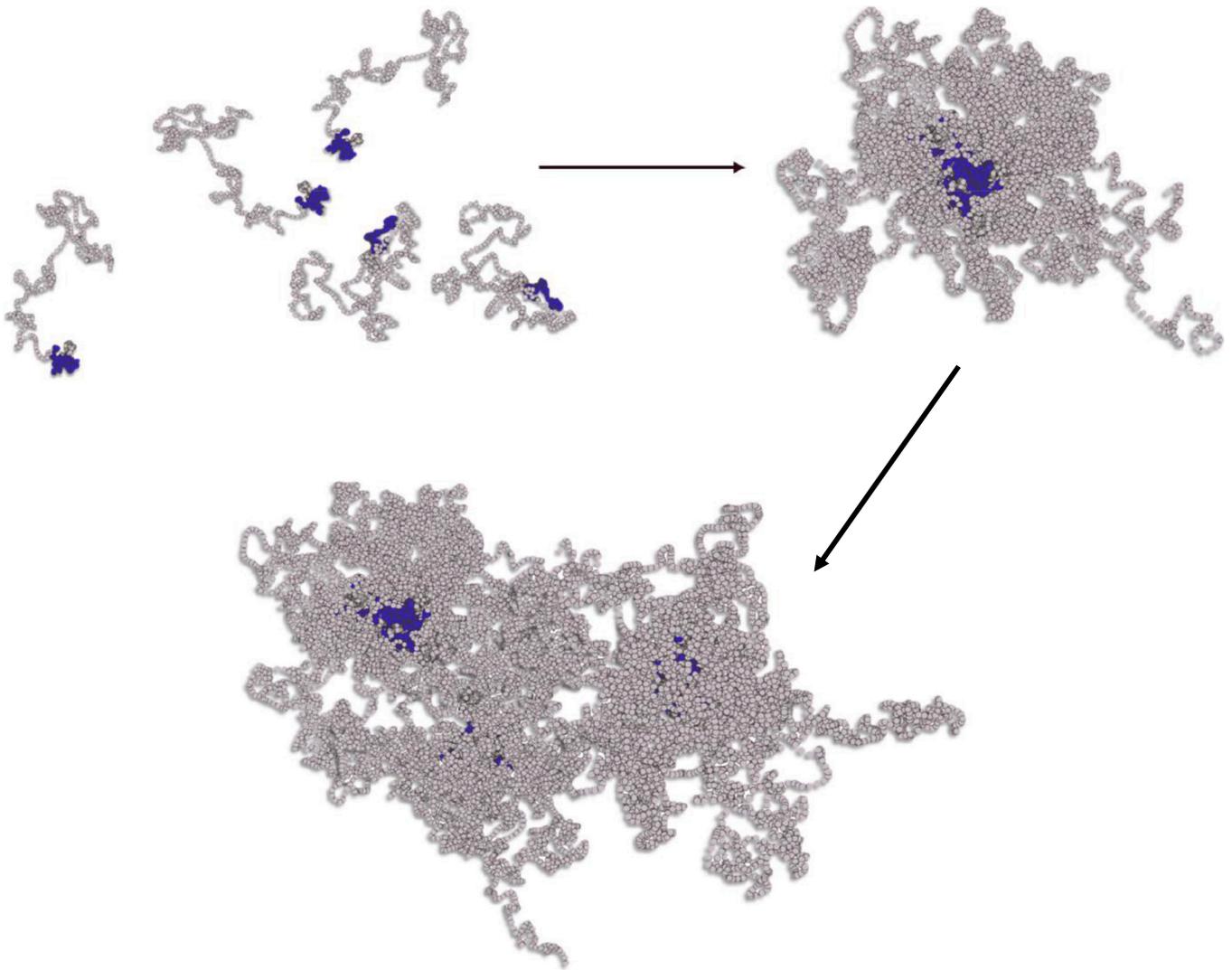


Fig. 8. Possible model of gliadin molecules that form polymers based on cross-links by their hydrophobic cores, and thereafter, these polymers interact through weaker bonds by their tails.

Funding

This research was funded by the Swedish Research Council, grant number 2013-05991, and Trees and Crops for the Future (TC4F). This research was conducted using the resources of High Performance Computing Center North (HPC2N).

Declaration of competing interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Acknowledgments

We wish to thank Senger et al. [42] for providing the α -gliadin

Appendix A

The primary structure for the α -gliadin used, the sequence in brackets belongs to signal peptide not used in this study:

plasmid and Faraz Muneer for making the gliadin-rich fractions. We also appreciate that Giulio Tesei helped elaborate the study. Anders Ekholm, Maria Luisa Prieto-Linde and Samuel Lenton for providing practical laboratory supervision. William Newson for information on how to analyze FTIR data. Finally, we also thank Christofer Lendel for helping with Dynamic Light scattering.

Autor statement

All authors have agreed on the submission to the International Journal of Biological Macromolecules. The paper has not been submitted and the content not revealed in any form in any other submission or publication.

```

" (MKTFLILALLAIVATTATTA) VRVPVPLQFPQNPSQQQPQEQVPLVQQQQLFLGQQQPF
PQQPYPPQPFPSQQPYLQLQFPFPQLPYSQPQFRPQQPYPPQFPQYSQPQQPISQQQQ
QQQQQQQQQQQQQQQIILQQILQQQLIPCMDVVLQQHNIHGRSQVLLQQSTYQLLQELCCQ
HLWQIPEQSQCQAIHKVVHAIILHQQKQQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQ
AQGSVQPQQLPQFEEIRNLALQTLPAMCNVYIPPYCTITPFGIFGTN"
    
```

Charge and hydrophobic distribution map for the α -gliadin, where + refers to positively charged, - negatively charged, * hydrophobic and . neutral amino acids:

```

"(+.....)+....._.....
.....+.....
....._.....+.....
....._.....+.....+.....+
....._.....+....."
    
```

Net charge for α -gliadin is +1.

Appendix B

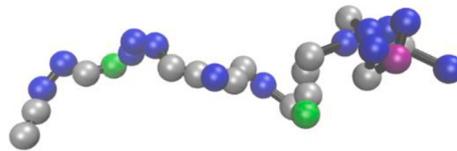
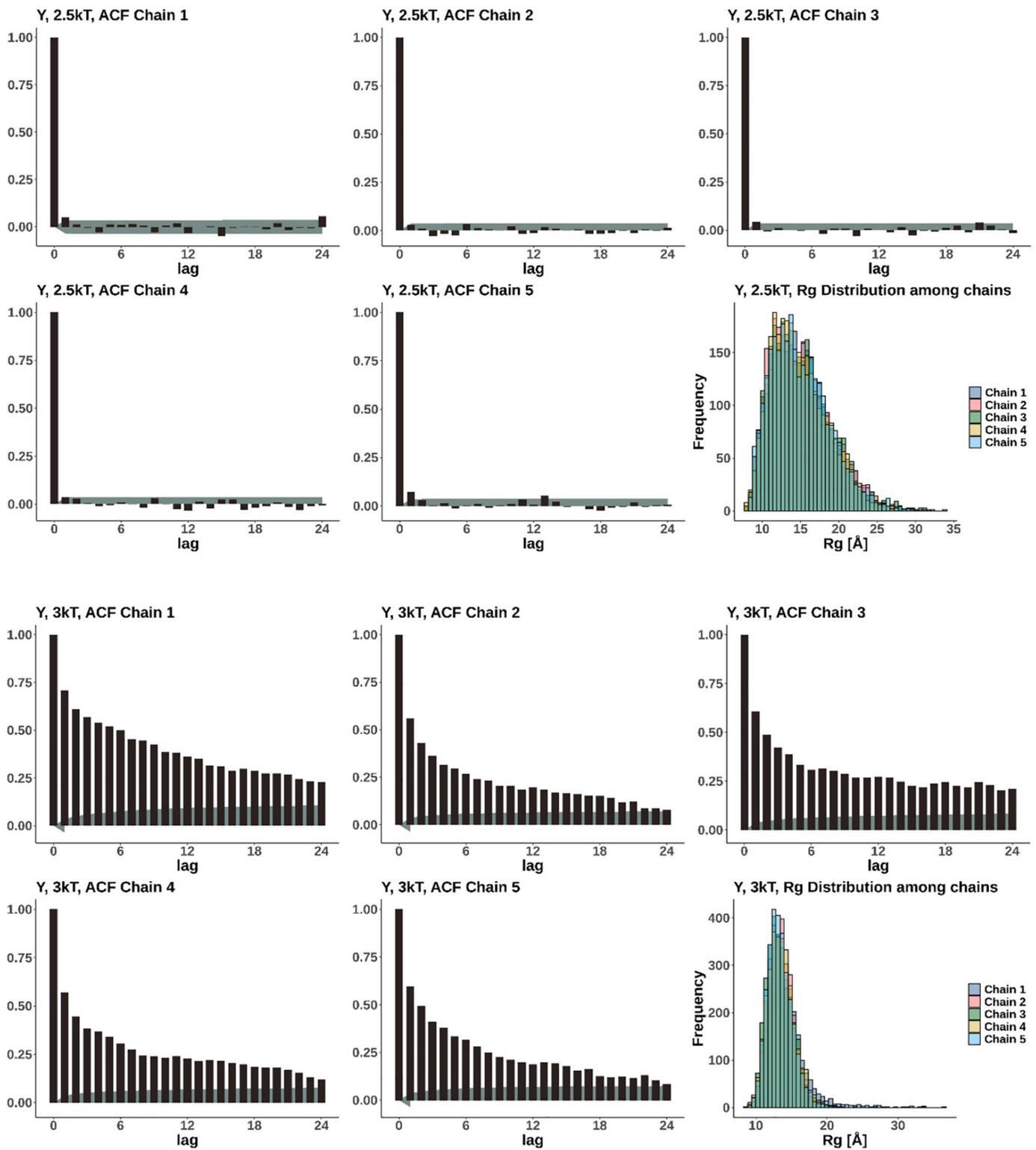
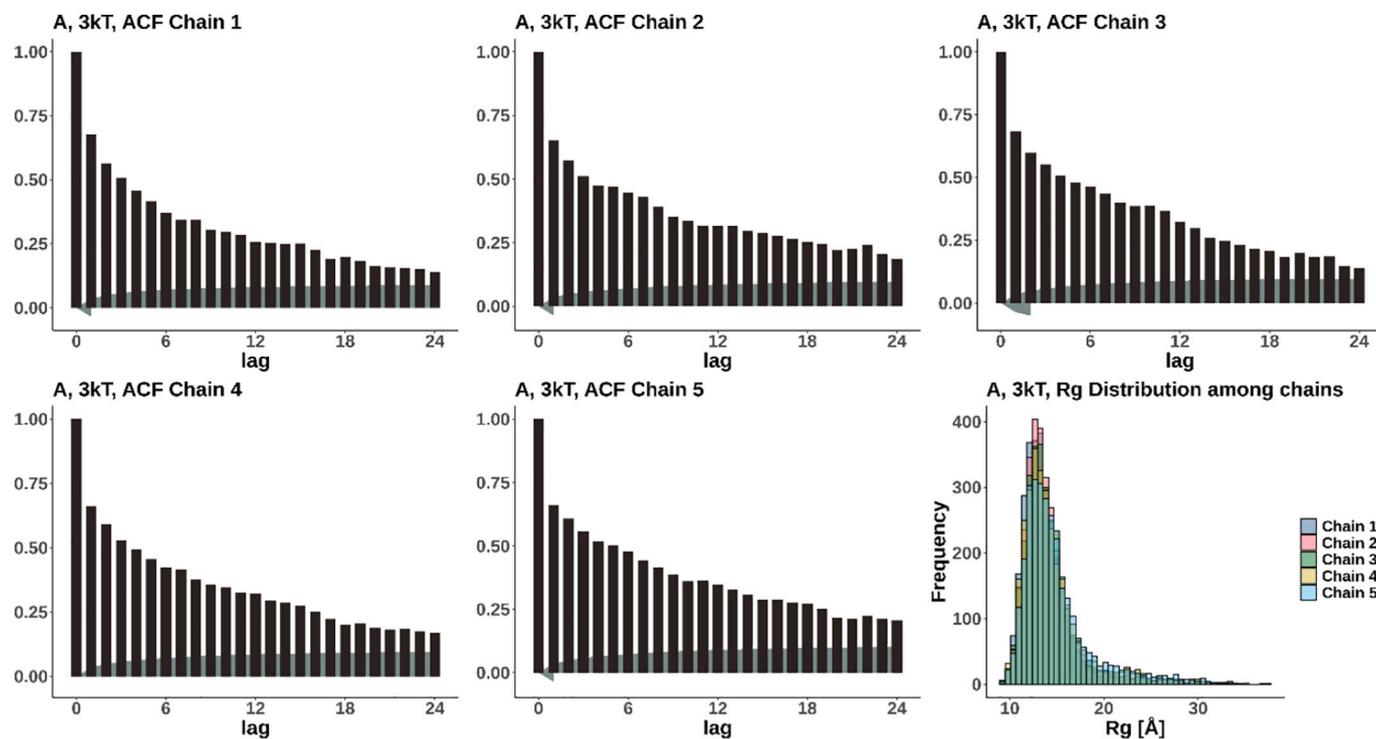
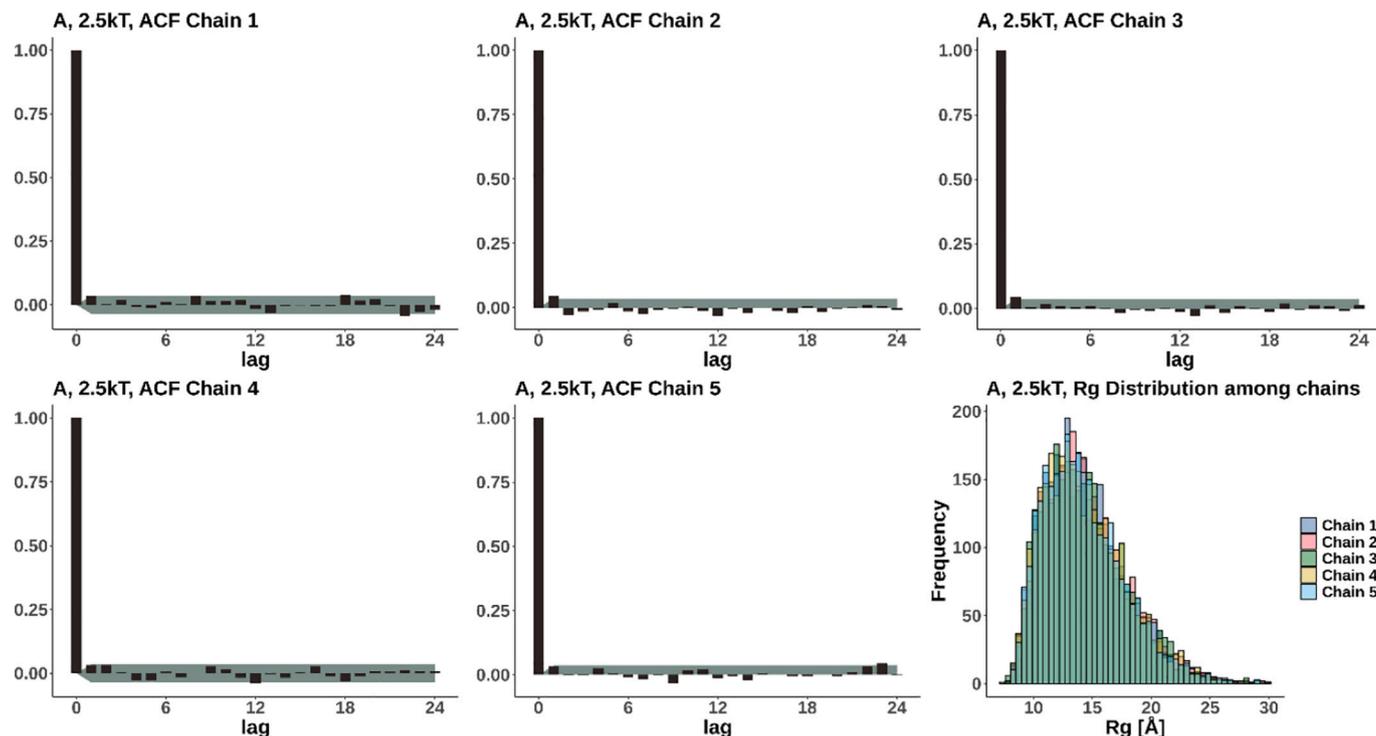


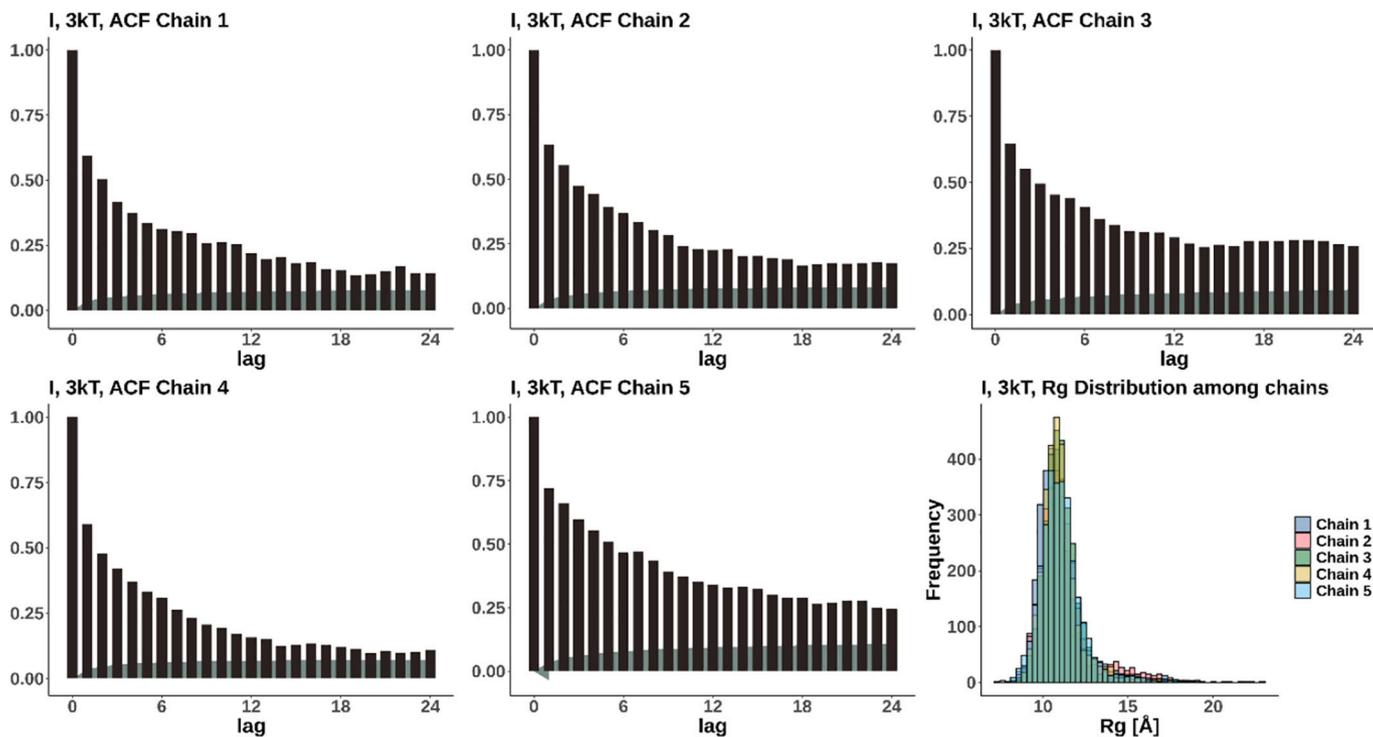
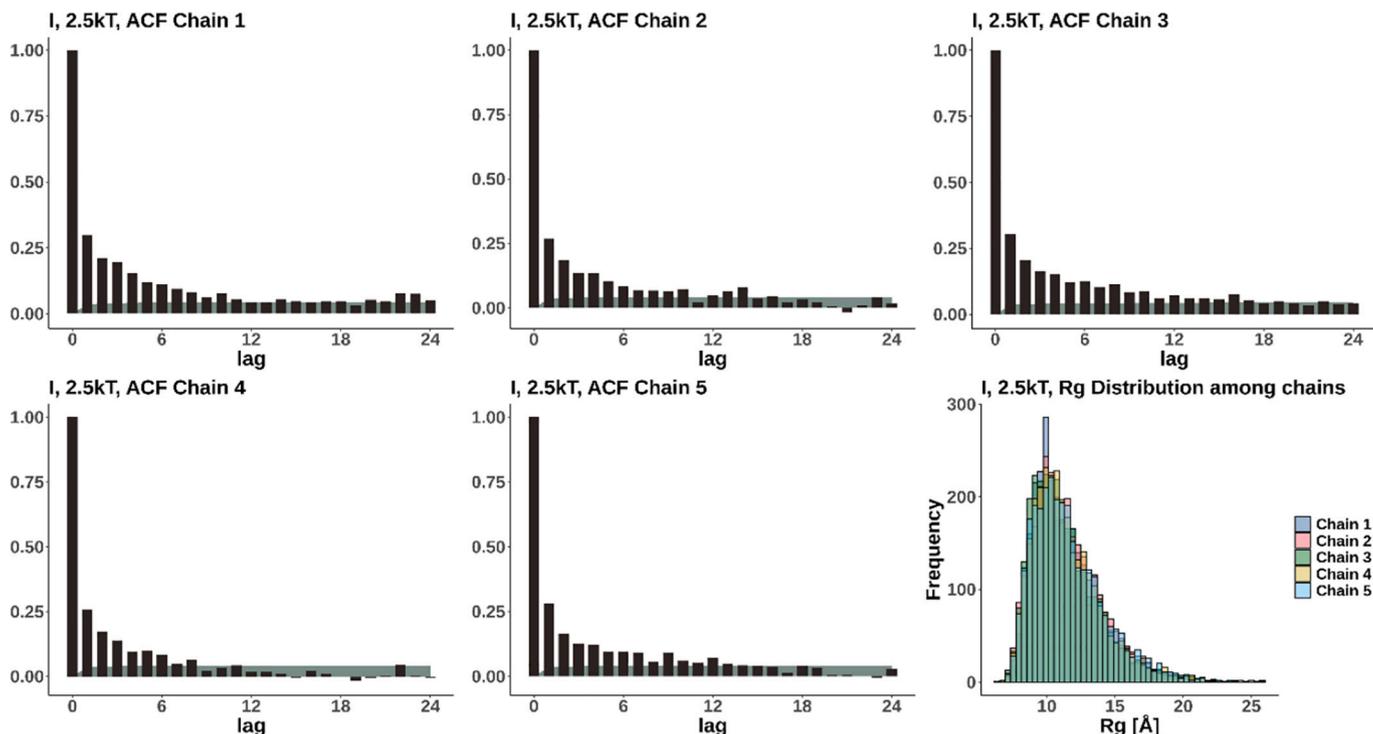
Fig. A1.

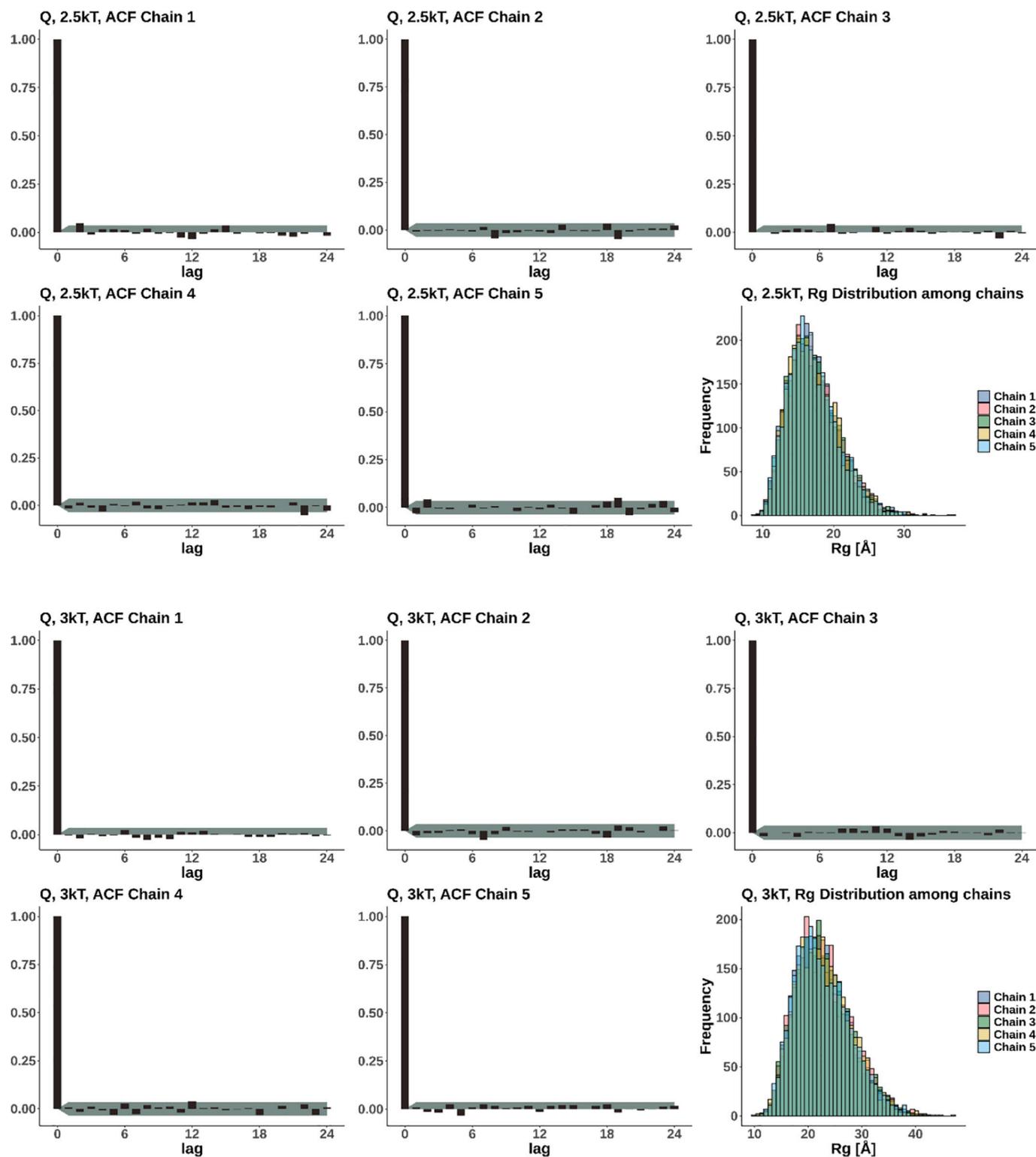
Appendix C

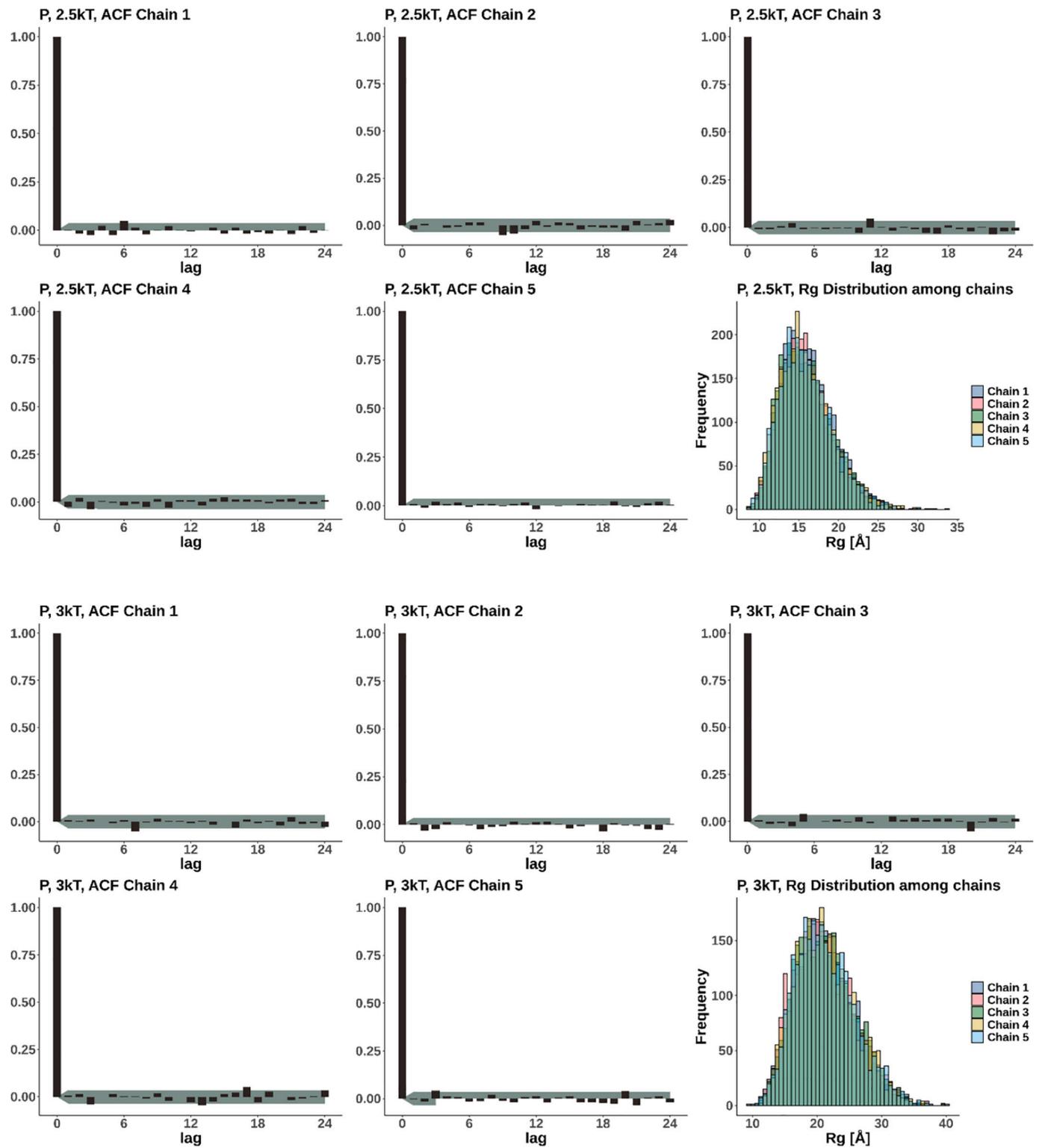
The autocorrelation-function and distribution of the Rg values for all chains in the simulations of the different peptides at both 2.5 and 3 kT are displayed in this section.

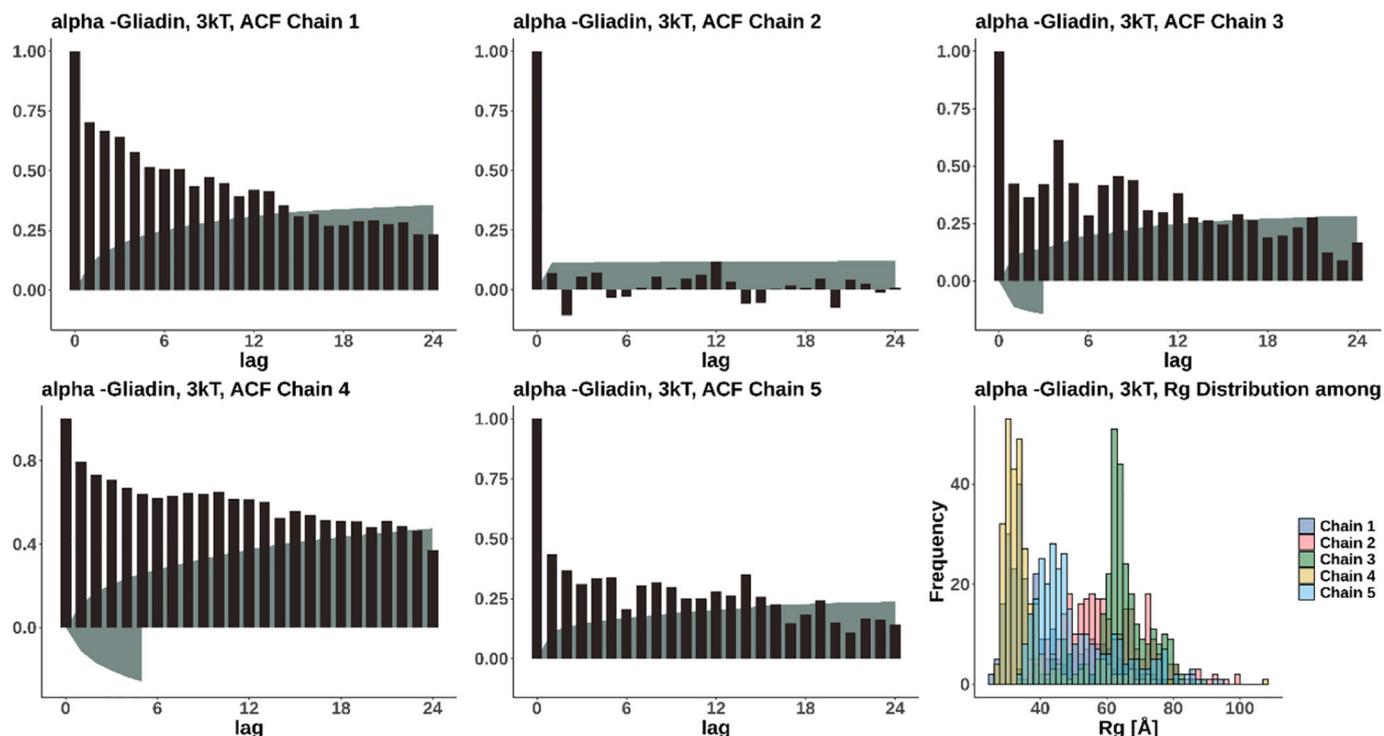












Appendix D

Cluster size probability and corresponding asphericity index, is listed for each peptide at different kT.

Peptide acronym	Hydrophobic potential in kT	Cluster size	asphericity index	probability
A	2.5	1	0.4	0.9283
A	2.5	2	0.3	0.0651
A	2.5	3	0.25	0.0062
A	2.5	4	0.22	0.0004
I	2.5	1	0.36	0.3441
I	2.5	2	0.22	0.0542
I	2.5	3	0.13	0.0936
I	2.5	4	0.09	0.2567
I	2.5	5	0.09	0.2515
P	2.5	1	0.45	0.9885
P	2.5	2	0.33	0.0113
P	2.5	3	0.27	0.0002
Q	2.5	1	0.43	0.9851
Q	2.5	2	0.31	0.0146
Q	2.5	3	0.24	0.0003
Y	2.5	1	0.41	0.9555
Y	2.5	2	0.32	0.0422
Y	2.5	3	0.30	0.0023
A	3	5	0.11	1
I	3	5	0.04	1
P	3	1	0.49	0.9847
P	3	2	0.34	0.0151
P	3	3	0.28	0.0002
Q	3	1	0.47	0.9757
Q	3	2	0.32	0.0235
Q	3	3	0.24	0.0008
Y	3	5	0.07	1

References

[1] Giant proteins with flour power | Nature, (n.d.). <https://www.nature.com/articles/381738a0> (accessed August 14, 2019).

[2] F. Ortolan, C.J. Steel, Protein characteristics that affect the quality of vital wheat gluten to be used in baking: a review, *Compr. Rev. Food Sci. Food Saf.* 16 (2017) 369–381, <https://doi.org/10.1111/1541-4337.12259>.

[3] A.J. Capezza, M. Lundman, R.T. Olsson, W.R. Newson, M.S. Hedenqvist, E. Johansson, Carboxylated wheat gluten proteins: a green solution for production of sustainable superabsorbent materials, *Biomacromolecules* 21 (2020) 1709–1719, <https://doi.org/10.1021/acs.biomac.9b01646>.

- [4] S. Bassi C.C. Maningat R. Chinnaswamy L. Nie , Modified wheat gluteins and use thereof in fabrication of films, US5747648A, 1998. <https://patents.google.com/patent/US5747648A/en> (accessed August 17, 2019).
- [5] H. Zhang, G. Mittal, Biodegradable protein-based films from plant resources: a review, *Environ. Prog. Sustain. Energy* 29 (2010) 203–220, <https://doi.org/10.1002/ep.10463>.
- [6] C. de Souza, R. Soares, T.S. Balbuena, P. Arruda, Structure, organization, and expression of the alpha prolamin multigenic family bring new insights into the evolutionary relationships among grasses, *Plant Genome* 8 (2015), <https://doi.org/10.3835/plantgenome2014.06.0027>.
- [7] J.-H. Xu, J. Messing, Amplification of prolamin storage protein genes in different subfamilies of the Poaceae, *Theor. Appl. Genet.* 119 (2009) 1397, <https://doi.org/10.1007/s00122-009-1143-x>.
- [8] M. Daly, S.N. Bromilow, C. Nitride, P.R. Shewry, L.A. Gethings, E.N.C. Mills, Mapping coeliac toxic motifs in the prolamin seed storage proteins of barley, rye, and oats using a curated sequence database, *Front. Nutr.* 7 (2020), <https://doi.org/10.3389/fnut.2020.00087>.
- [9] P.R. Shewry, B.J. Mifflin, D.D. Kasarda, The structural and evolutionary relationships of the prolamin storage proteins of barley, rye and wheat, <sb: contribution><sb:title>Philos. Trans. R. Soc. Lond.</sb:title> </sb: contribution><sb:host><sb:issue><sb:series><sb:title>Ser. B Biol. Sci.</sb: title></sb:series></sb:issue></sb:host> 304 (1984) 297–308.
- [10] R. Kuktaite, H. Larsson, E. Johansson, Variation in protein composition of wheat flour and its relationship to dough mixing behaviour, *J. Cereal Sci.* 40 (2004) 31–39.
- [11] P.R. Shewry, J.A. Napier, A.S. Tatham, Seed storage proteins: structures and biosynthesis, *Plant Cell* 7 (1995) 945–956.
- [12] P.R. Shewry, A.S. Tatham, J. Forde, M. Kreis, B.J. Mifflin, The classification and nomenclature of wheat gluten proteins: a reassessment, *J. Cereal Sci.* 4 (1986) 97–106, [https://doi.org/10.1016/S0733-5210\(86\)80012-1](https://doi.org/10.1016/S0733-5210(86)80012-1).
- [13] B. Keck, P. Kohler, H. Wieser, Disulphide bonds in wheat gluten: cystine peptides derived from gluten proteins following peptic and thermolytic digestion 200 (1995) 432–439, <https://doi.org/10.1007/BF01193253>.
- [14] P.R. Shewry, A.S. Tatham, Disulphide bonds in wheat gluten proteins, *J. Cereal Sci.* 25 (1997) 207–227, <https://doi.org/10.1006/jcrs.1996.0100>.
- [15] P. Kohler, H.-D. Belitz, H. Wieser, Disulphide bonds in wheat gluten: further cystine peptides from high molecular weight (HMW) and low molecular weight (LMW) subunits of glutenin and from γ -gliadins, *Z. Lebensm. Unters. Forsch.* 196 (1993) 239–247, <https://doi.org/10.1007/BF01202740>.
- [16] J.H. Woychik, J.A. Boundy, R.J. Dimler, Starch gel electrophoresis of wheat gluten proteins with concentrated urea, *Arch. Biochem. Biophys.* 94 (1961) 477–482, [https://doi.org/10.1016/0003-9861\(61\)90075-3](https://doi.org/10.1016/0003-9861(61)90075-3).
- [17] R. Urade, N. Sato, M. Sugiyama, Gliadins from wheat grain: an overview, from primary structure to nanostructures of aggregates, *Biophys. Rev.* 10 (2017) 435–443, <https://doi.org/10.1007/s12551-017-0367-2>.
- [18] J. Markgren, M. Hedenqvist, F. Rasheed, M. Škepč, E. Johansson, Glutenin and gliadin, a piece of their structural properties in the cell described through Monte Carlo simulations, *Biomolecules* 10 (2020) 1095, <https://doi.org/10.3390/biom10081095>.
- [19] M.F. Gómez Castro, E. Miculán, M.G. Herrera, C. Ruera, F. Perez, E.D. Prieto, E. Barrera, S. Pantano, P. Carasi, F.G. Chirido, p31–43 gliadin peptide forms oligomers and induces NLRP3 inflammasome/Caspase 1-dependent mucosal damage in small intestine, *Front. Immunol.* 10 (2019), <https://doi.org/10.3389/fimmu.2019.00031>.
- [20] M.G. Herrera, M.F.G. Castro, E. Prieto, E. Barrera, V.I. Dodero, S. Pantano, F. Chirido, Structural conformation and self-assembly process of p31–43 gliadin peptide in aqueous solution. Implications for celiac disease, *FEBS J.* 287 (2019) 2134–2149, <https://doi.org/10.1111/febs.15109>.
- [21] K. Kar, M. Jayaraman, B. Sahoo, R. Kodali, R. Wetzal, Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent, *Nat. Struct. Mol. Biol.* 18 (2011) 328–336, <https://doi.org/10.1038/nsmb.1992>.
- [22] J. Wen, D.R. Scoles, J.C. Facelli, Molecular dynamics analysis of the aggregation propensity of polyglutamine segments, *PLOS ONE* 12 (2017), e0178333, <https://doi.org/10.1371/journal.pone.0178333>.
- [23] S. Polling, A.R. Ormsby, R.J. Wood, K. Lee, C. Shoubridge, J.N. Hughes, P. Q. Thomas, M.D.W. Griffin, A.F. Hill, Q. Bowden, T. Böcking, D.M. Hatters, Polyalanine expansions drive a shift into α -helical clusters without amyloid-fibril formation, *Nat. Struct. Mol. Biol.* 22 (2015) 1008–1015, <https://doi.org/10.1038/nsmb.3127>.
- [24] F. Rasheed, W.R. Newson, T.S. Plivelic, R. Kuktaite, M.S. Hedenqvist, M. Gällstedt, E. Johansson, Macromolecular changes and nano-structural arrangements in gliadin and glutenin films upon chemical modification: relation to functionality, *Int. J. Biol. Macromol.* 79 (2015) 151–159, <https://doi.org/10.1016/j.ijbiomac.2015.04.033>.
- [25] B. Lagrain, K. Brijis, J.A. Delcour, Reaction kinetics of gliadin–glutenin cross-linking in model systems and in bread making, *J. Agric. Food Chem.* 56 (2008) 10660–10666, <https://doi.org/10.1021/jf801894r>.
- [26] V.N. Uversky, A.K. Dunker, Understanding protein non-folding, <sb: contribution><sb:title>Biochim. Biophys. Acta</sb:title> </sb: contribution><sb:host><sb:issue><sb:series><sb:title>Proteins Proteomics</sb: title></sb:series></sb:issue></sb:host> 1804 (2010) 1231–1264, <https://doi.org/10.1016/j.bbapap.2010.01.017>.
- [27] F. Rasheed, W.R. Newson, T.S. Plivelic, R. Kuktaite, M.S. Hedenqvist, M. Gällstedt, E. Johansson, Structural architecture and solubility of native and modified gliadin and glutenin proteins: non-crystalline molecular and atomic organisation, *RSC Adv.* 4 (2014) 2051–2060, <https://doi.org/10.1039/C3RA45522J>.
- [28] M.G. Herrera, D.S. Vazquez, R. Sreij, M. Drechsler, Y. Hertle, T. Hellweg, V. I. Dodero, Insights into gliadin supramolecular organisation at digestive pH 3.0, *Colloids Surf. B: Biointerfaces* 165 (2018) 363–370, <https://doi.org/10.1016/j.colsurfb.2018.02.053>.
- [29] F. Yaşar, S. Çelik, H. Köksel, Molecular modeling of various peptide sequences of gliadins and low-molecular-weight glutenin subunits, *Food/Nahrung* 47 (2003) 238–242, <https://doi.org/10.1002/food.200390056>.
- [30] E.P.G. Arêas, M.M. Cassiano, Folding interpenetration in a gliadin model: the role of the characteristic octapeptide motif, *Biophys. Chem.* 90 (2001) 135–146, [https://doi.org/10.1016/S0301-4622\(01\)00138-7](https://doi.org/10.1016/S0301-4622(01)00138-7).
- [31] M. Julia Amundarain, M. Georgina Herrera, F. Zamarreño, J. Francisco Viso, M. D. Costabel, V.I. Dodero, Molecular mechanisms of 33-mer gliadin peptide oligomerisation, *Phys. Chem. Chem. Phys.* 21 (2019) 22539–22552, <https://doi.org/10.1039/C9CP02338K>.
- [32] M.G. Herrera, M.F.G. Castro, E. Prieto, E. Barrera, V.I. Dodero, S. Pantano, F. Chirido, Structural conformation and self-assembly process of p31–43 gliadin peptide in aqueous solution. Implications for celiac disease, *FEBS J.* 287 (2020) 2134–2149, <https://doi.org/10.1111/febs.15109>.
- [33] V.H. Man, Y. Zhang, C. Roland, C. Sagui, Structure and stability of amyloid protofibrils of polyglutamine and polyasparagine from molecular dynamics simulations, in: *Biomolecular Simulations in Structure-based Drug Discovery*, John Wiley & Sons, Ltd, 2018, pp. 301–324, <https://doi.org/10.1002/9783527806836.ch12>.
- [34] The Central Domain of High Molecular Weight Glutenin Subunits is Water-Soluble | Elsevier Enhanced Reader, (n.d.). doi:10.1006/jcrs.1998.0223.
- [35] Dissecting the Disulfide Linkage of the N-Terminal Domain of HMW 1Dx5 and Its Contributions to Dough Functionality | Journal of Agricultural and Food Chemistry, (n.d.). <https://pubs.acs.org/doi/10.1021/acs.jafc.7b02449> (accessed June 27, 2020).
- [36] Adsorption of the High Molecular Weight Glutenin Subunit 1Dx5 Compared to the 58-kDa Central Repetitive Domain and -gliadins | Elsevier Enhanced Reader, (n.d.). doi:10.1006/jcrs.2001.0365.
- [37] M. Julia Amundarain, M. Georgina Herrera, F. Zamarreño, J. Francisco Viso, M. D. Costabel, V.I. Dodero, Molecular mechanisms of 33-mer gliadin peptide oligomerisation, *Phys. Chem. Chem. Phys.* 21 (2019) 22539–22552, <https://doi.org/10.1039/C9CP02338K>.
- [38] L. Sahli, A. Boire, V. Solé-Jamault, H. Rogniaux, A. Giuliani, P. Roblin, D. Renard, New exploration of the γ -gliadin structure through its partial hydrolysis, *Int. J. Biol. Macromol.* 165 (2020) 654–664, <https://doi.org/10.1016/j.ijbiomac.2020.09.136>.
- [39] M. Francin-Allami, A. Boudier, Y. Popineau, Comparative study of wheat low-molecular-weight glutenin and α -gliadin trafficking in tobacco cells, *Plant Cell Rep.* 32 (2013) 89–101, <https://doi.org/10.1007/s00299-012-1343-8>.
- [40] F.C. Greene, In vitro synthesis of wheat (*Triticum aestivum* L.) storage proteins, *Plant Physiol.* 68 (1981) 778–783, <https://doi.org/10.1104/pp.68.3.778>.
- [41] Y. Altschuler, G. Galili, Role of conserved cysteines of a wheat gliadin in its transport and assembly into protein bodies in xenopus oocytes, *J. Biol. Chem.* 269 (1994) 6677–6682.
- [42] S. Senger, F. Maurano, M.F. Mazzeo, M. Gaita, O. Fierro, C.S. David, R. Troncone, S. Auricchio, R.A. Siciliano, M. Rossi, Identification of immunodominant epitopes of α -gliadin in HLA-DQ8 transgenic mice following oral immunization, *J. Immunol.* 175 (2005) 8087–8095, <https://doi.org/10.4049/jimmunol.175.12.8087>.
- [43] F. Rasheed, J. Markgren, M. Hedenqvist, E. Johansson, Modeling to understand plant protein structure-function relationships—implications for seed storage proteins, *Molecules* 25 (2020) 873, <https://doi.org/10.3390/molecules25040873>.
- [44] F. Muneer, E. Johansson, M.S. Hedenqvist, M. Gällstedt, W.R. Newson, Preparation, properties, protein cross-linking and biodegradability of plasticizer-solvent free hemp fibre reinforced wheat gluten, glutenin, and gliadin composites, *Bioresources* 9 (2014) 5246–5261.
- [45] D.D. Kasarda, R. D’Ovidio, Deduced amino acid sequence of an α -gliadin gene from spelt wheat (*Spelta*) includes sequences active in celiac disease, *Cereal Chem. J.* 76 (1999) 548–551, <https://doi.org/10.1094/CHEM.1999.76.4.548>.
- [46] H. Wieser, Chemistry of gluten proteins, *Food Microbiol.* 24 (2007) 115–119, <https://doi.org/10.1016/j.fm.2006.07.004>.
- [47] K.M. Ruff, R.V. Pappu, A.S. Holehouse, Conformational preferences and phase behavior of intrinsically disordered low complexity sequences: insights from multiscale simulations, *Curr. Opin. Struct. Biol.* 56 (2019) 1–10, <https://doi.org/10.1016/j.sbi.2018.10.003>.
- [48] S. Müller, H. Wieser, The location of disulphide bonds in α -type gliadins, *J. Cereal Sci.* 22 (1995) 21–27, [https://doi.org/10.1016/S0733-5210\(05\)80004-9](https://doi.org/10.1016/S0733-5210(05)80004-9).
- [49] M. Andreasen, K.K. Skeby, S. Zhang, E.H. Nielsen, L.H. Klausen, H. Frahm, G. Christiansen, T. Skrydstrup, M. Dong, B. Schjøtt, D. Otzen, The importance of being capped: terminal capping of an amyloidogenic peptide affects fibrillation propensity and fibril morphology, *Biochemistry* 53 (2014) 6968–6980, <https://doi.org/10.1021/bi500674u>.
- [50] O.D. Anderson, F.C. Greene, The α -gliadin gene family. II. DNA and protein sequence variation, subfamily structure, and origins of pseudogenes, *Theor. Appl. Genet.* 95 (1997) 59–65, <https://doi.org/10.1007/s001220050532>.
- [51] D.D. Kasarda, Structure and properties of alpha -gliadins, *Ann. Technol. Agric.* 29 (1980) 151–173.

- [52] L. Josefsson, M. Cronhamn, M. Ekman, H. Widehammar, Å. Emmer, C. Lendel, Structural basis for the formation of soy protein nanofibrils, *RSC Adv.* 9 (2019) 6310–6319, <https://doi.org/10.1039/C8RA10610J>.
- [53] A. Micsonai, F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers, J. Kardos, Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy, *PNAS* 112 (2015) E3095–E3103, <https://doi.org/10.1073/pnas.1500851112>.
- [54] S.-W. Cho, M. Gällstedt, E. Johansson, M.S. Hedenqvist, Injection-molded nanocomposites and materials based on wheat gluten, *Int. J. Biol. Macromol.* 48 (2011) 146–152, <https://doi.org/10.1016/j.ijbiomac.2010.10.012>.
- [55] M. Wojdyr, Fityk: a general-purpose peak fitting program, *J. Appl. Crystallogr.* 43 (2010) 1126–1128, <https://doi.org/10.1107/S0021889810030499>.
- [56] S. Bhattacharjee, DLS and zeta potential – what they are and what they are not? *J. Control. Release* 235 (2016) 337–351, <https://doi.org/10.1016/j.jconrel.2016.06.017>.
- [57] E. Johansson, M.L. Prieto-Linde, J.Ö. Jönsson, Effects of wheat cultivar and nitrogen application on storage protein composition and breadmaking quality, *Cereal Chem.* 78 (2001) 19–25, <https://doi.org/10.1094/CCHEM.2001.78.1.19>.
- [58] E. Johansson, A.H. Malik, A. Hussain, F. Rasheed, W.R. Newson, T. Plivelic, M. S. Hedenqvist, M. Gällstedt, R. Kuktaite, Wheat gluten polymer structures: the impact of genotype, environment and processing on their functionality in various applications, *Cereal Chem.* 90 (2013) 367–376.
- [59] M. Helguera, A. Abugalieva, S. Battenfield, F. Békés, G. Branlard, M. Cuniberti, A. Hüskén, E. Johansson, C.F. Morris, E. Nurit, M. Sissons, D. Vazquez, Grain quality in breeding, in: G. Igrejas, T. Ikeda, C. Guzmán (Eds.), *Wheat Quality for Improving Processing And Human Health*, Springer, Cham, 2020, pp. 273–307.
- [60] R. Singh, G.M. Whitesides, Reagents for rapid reduction of disulfide bonds in proteins, in: J.W. Crabbe (Ed.), *Techniques in Protein Chemistry*, Academic Press, 1995, pp. 259–266, [https://doi.org/10.1016/S1080-8914\(06\)80033-5](https://doi.org/10.1016/S1080-8914(06)80033-5).
- [61] J.A. Reynolds, C. Tanford, Binding of dodecyl sulfate to proteins at high binding ratios. Possible implications for the state of proteins in biological membranes*, *Proc. Natl. Acad. Sci. U. S. A.* 66 (1970) 1002–1007.
- [62] P.J. Rossky, Protein denaturation by urea: slash and bond, *PNAS* 105 (2008) 16825–16826, <https://doi.org/10.1073/pnas.0809224105>.
- [63] R: What is R?, (n.d.). <https://www.r-project.org/about.html> (accessed May 13, 2019).
- [64] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.* 157 (1982) 105–132, [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0).
- [65] D.G. Archer, P. Wang, The dielectric constant of water and Debye-Hückel limiting law slopes, *J. Phys. Chem. Ref. Data* 19 (1990) 371–411, <https://doi.org/10.1063/1.555853>.
- [66] C. Cragnell, E. Rieloff, M. Skepö, Utilising coarse-grained modeling and Monte Carlo simulations to evaluate the conformational ensemble of intrinsically disordered proteins and regions, *J. Mol. Biol.* 430 (2018) 2478–2492, <https://doi.org/10.1016/j.jmb.2018.03.006>.
- [67] N. Sato, A. Matsumiya, Y. Higashino, S. Funaki, Y. Kitao, Y. Oba, R. Inoue, F. Arisaka, M. Sugiyama, R. Urade, Molecular assembly of wheat gliadins into nanostructures: a small-angle X-ray scattering study of gliadins in distilled water over a wide concentration range, *J. Agric. Food Chem.* 63 (2015) 8715–8721, <https://doi.org/10.1021/acs.jafc.5b02902>.
- [68] N.H. Thomson, M.J. Miles, Y. Popineau, J. Harries, P. Shewry, A.S. Tatham, Small angle X-ray scattering of wheat seed-storage proteins: α -, γ - and ω -gliadins and the high molecular weight (HMW) subunits of glutenin, *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* 1430 (1999) 359–366, [https://doi.org/10.1016/S0167-4838\(99\)00019-9](https://doi.org/10.1016/S0167-4838(99)00019-9).
- [69] R. Jurij, L. Per, MOLSIM: a modular molecular simulation software, *J. Comput. Chem.* 36 (2015) 1259–1274, <https://doi.org/10.1002/jcc.23919>.
- [70] E. Rieloff, M.D. Tully, M. Skepö, Assessing the intricate balance of intermolecular interactions upon self-association of intrinsically disordered proteins, *J. Mol. Biol.* 431 (2019) 511–523, <https://doi.org/10.1016/j.jmb.2018.11.027>.
- [71] A.G. Kikhney, D.I. Svergun, A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins, *FEBS Lett.* 589 (2015) 2570–2577, <https://doi.org/10.1016/j.febslet.2015.08.027>.
- [72] J. Rudnick, G. Gaspari, The asphericity of random walks, *J. Phys. A Math. Gen.* 19 (1986) L191–L193, <https://doi.org/10.1088/0305-4470/19/4/004>.
- [73] O. Conchillo-Solé, N.S. de Groot, F.X. Avilés, J. Vendrell, X. Daura, S. Ventura, AGGRESCAN: a server for the prediction and evaluation of “hot spots” of aggregation in polypeptides, *BMC Bioinforma.* 8 (2007) 65, <https://doi.org/10.1186/1471-2105-8-65>.
- [74] B.J. Grant, A.P.C. Rodrigues, K.M. ElSawy, J.A. McCammon, L.S.D. Cavas, Bio3d: an R package for the comparative analysis of protein structures, *Bioinformatics* 22 (2006) 2695–2696, <https://doi.org/10.1093/bioinformatics/btd461>.
- [75] VMD - Visual Molecular Dynamics, (n.d.). <https://www.ks.uiuc.edu/Research/vmd/> (accessed May 14, 2019).
- [76] V.I. Dodero, Z.B. Quirolo, M.A. Sequeira, Biomolecular studies by circular dichroism, *Front. Biosci.* 16 (2011) 61–73.
- [77] P.R. Shewry, M.J. Miles, N.H. Thomson, A.S. Tatham, Scanning probe microscopes—applications in cereal science, *Cereal Chem.* 74 (1997) 193–199, <https://doi.org/10.1094/CCHEM.1997.74.3.193>.
- [78] D.D. Kasarda, J.E. Bernardin, William Gaffield, Circular dichroism and optical rotatory dispersion of α -gliadin, *Biochemistry* 7 (1968) 3950–3957, <https://doi.org/10.1021/bi00851a023>.
- [79] Cereal Chem 1985 | The Beta-Turn Conformation in Wheat Gluten Proteins: Relationship to Gluten Elasticity., (n.d.). <https://www.aacnet.org/publications/cc/backissues/1985/Documents/CC1985a129.html> (accessed September 27, 2018).
- [80] E.W. Blanch, D.D. Kasarda, L. Hecht, K. Nielsen, L.D. Barron, New insight into the solution structures of wheat gluten proteins from Raman optical activity, *Biochemistry* 42 (2003) 5665–5673, <https://doi.org/10.1021/bi027059y>.
- [81] L. Calvanese, M. Nanayakkara, R. Aitoro, M. Sanseverino, A.L. Tornesello, L. Falcigno, G. D’Auria, M.V. Barone, Structural insights on P31–43, a gliadin peptide able to promote an innate but not an adaptive response in celiac disease, *J. Pept. Sci.* 25 (2019), e3161, <https://doi.org/10.1002/psc.3161>.
- [82] M.G. Herrera, L.A. Benedini, C. Lonz, P.L. Schilardi, T. Hellweg, J.-M. Ruyschaert, V.I. Dodero, Self-assembly of 33-mer gliadin peptide oligomers, *Soft Matter* 11 (2015) 8648–8660, <https://doi.org/10.1039/C5SM01619C>.
- [83] A.A. Adzhubei, M.J.E. Sternberg, A.A. Makarov, Polyproline-II helix in proteins: structure and function, *J. Mol. Biol.* 425 (2013) 2100–2132, <https://doi.org/10.1016/j.jmb.2013.03.018>.
- [84] M.G. Herrera, F. Zamarreño, M. Costabel, H. Ritacco, A. Hütten, N. Sewald, V. I. Dodero, Circular dichroism and electron microscopy studies in vitro of 33-mer gliadin peptide revealed secondary structure transition and supramolecular organisation, *Biopolymers* 101 (2014) 96–106, <https://doi.org/10.1002/bip.22288>.
- [85] I. Rombouts, B. Lagrain, K. Brijs, J.A. Delcour, β -Elimination reactions and formation of covalent cross-links in gliadin during heating at alkaline pH, *J. Cereal Sci.* 52 (2010) 362–367, <https://doi.org/10.1016/j.jcs.2010.06.006>.
- [86] C.-F. Xu, Y. Chen, L. Yi, T. Brantley, B. Stanley, Z. Sosic, L. Zang, Discovery and characterization of histidine oxidation initiated cross-links in an IgG1 monoclonal antibody, *Anal. Chem.* 89 (2017) 7915–7923, <https://doi.org/10.1021/acs.analchem.7b00860>.
- [87] N. Fujieda, His-Cys and Trp-Cys cross-links generated by post-translational chemical modification, *Biosci. Biotechnol. Biochem.* 84 (2020) 445–454, <https://doi.org/10.1080/09168451.2019.1696178>.
- [88] P.L. Weegels, Theoretical considerations on the directionality in polymerisation mechanism and polymer type, *Macromol.TheorySimul.* 5 (1996) 299–303, <https://doi.org/10.1002/mats.1996.040050210>.
- [89] H. Singh, F. MacRitchie, Changes in proteins induced by heating gluten dispersions at high temperature, *J. Cereal Sci.* 39 (2004) 297–301, <https://doi.org/10.1016/j.jcs.2003.11.004>.
- [90] R.K. Das, R.V. Pappu, Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues, *PNAS* 110 (2013) 13392–13397, <https://doi.org/10.1073/pnas.1304749110>.
- [91] R.K. Das, K.M. Ruff, R.V. Pappu, Relating sequence encoded information to form and function of intrinsically disordered proteins, *Curr. Opin. Struct. Biol.* 32 (2015) 102–112, <https://doi.org/10.1016/j.sbi.2015.03.008>.
- [92] S. Rauscher, R. Pomès, The liquid structure of elastin, *eLife* 6 (2017), e26526, <https://doi.org/10.7554/eLife.26526>.
- [93] A. Boire, C. Sanchez, M.-H. Morel, M.P. Lettinga, P. Menut, Dynamics of liquid-like phase separation of wheat gliadins, *Sci. Rep.* 8 (2018), <https://doi.org/10.1038/s41598-018-32278-5>.
- [94] B. Gabryelczyk, H. Cai, X. Shi, Y. Sun, P.J.M. Swinkels, S. Salenting, K. Pervushin, A. Miserez, Hydrogen bond guidance and aromatic stacking drive liquid-liquid phase separation of intrinsically disordered histidine-rich peptides, *Nat. Commun.* 10 (2019) 5465, <https://doi.org/10.1038/s41467-019-13469-8>.
- [95] L. Sahli, D. Renard, V. Solé-Jamault, A. Giuliani, A. Boire, Role of protein conformation and weak interactions on γ -gliadin liquid-liquid phase separation, *Sci. Rep.* 9 (2019) 1–13, <https://doi.org/10.1038/s41598-019-49745-2>.
- [96] R. Kuktaite, W.R. Newson, F. Rasheed, M.S. Plivelic, M.S. Hedenqvist, M. Gällstedt, E. Johansson, Monitoring nano-structure dynamics and polymerization in glycerol plasticized wheat gliadin and glutenin films: relation to mechanical properties, *ACS Sustain. Chem. Eng.* 4 (2016) 2998–3007.
- [97] F. Rasheed, R. Kuktaite, M.S. Hedenqvist, M. Gällstedt, T. Plivelic, E. Johansson, The use of the plant as a “green factory” to produce high strength gluten-based plastics, *Green Chem.* 18 (2016) 2782–2792.
- [98] G.H. Zerze, R.B. Best, J. Mittal, Sequence- and temperature-dependent properties of unfolded and disordered proteins from atomistic simulations, *J. Phys. Chem. B* 119 (2015) 14622–14630, <https://doi.org/10.1021/acs.jpcc.5b08619>.
- [99] A. Ribeiro, F.J. Arias, J. Reguera, M. Alonso, J.C. Rodríguez-Cabello, Influence of the amino-acid sequence on the inverse temperature transition of elastin-like polymers, *Biophys. J.* 97 (2009) 312–320, <https://doi.org/10.1016/j.bpj.2009.03.030>.
- [100] D.W. Urry, Entropic elastic processes in protein mechanisms. I. Elastic structure due to an inverse temperature transition and elasticity due to internal chain dynamics, *J. Protein Chem.* 7 (1988) 1–34, <https://doi.org/10.1007/BF01025411>.
- [101] Elasticity and Inverse Temperature Transition in Elastin | The Journal of Physical Chemistry Letters, (n.d.). <https://pubs.acs.org/doi/10.1021/acs.jpclett.5b01890> (accessed December 14, 2020).
- [102] E. Johansson, G. Branlard, M. Cuniberti, Z. Flagella, A. Hüskén, E. Nurit, R. J. Peña, M. Sissons, D. Vazquez, in: G. Igrejas, T. Ikeda, C. Guzmán (Eds.), *Wheat Quality for Improving Processing And Human Health*, Springer, Cham, 2020, pp. 171–204.
- [103] J.M. Norrander, J. Vieira, I. Rubenstein, J. Messing, Manipulation and expression of the maize zein storage proteins in *Escherichia coli*, *J. Biotechnol.* 2 (1985) 157–175, [https://doi.org/10.1016/0168-1656\(85\)90036-7](https://doi.org/10.1016/0168-1656(85)90036-7).
- [104] G. Gallili, Heterologous expression of a wheat high molecular weight glutenin gene in *Escherichia coli*, *Proc. Natl. Acad. Sci. U. S. A.* 86 (1989) 7756–7760.

- [105] D. Bartels, R.D. Thompson, S. Rothstein, Synthesis of a wheat storage protein subunit in *Escherichia coli* using novel expression vectors, *Gene* 35 (1985) 159–167, [https://doi.org/10.1016/0378-1119\(85\)90168-4](https://doi.org/10.1016/0378-1119(85)90168-4).
- [106] C. Dowd, F. Bekes, Large-scale expression and purification of high-molecular-weight glutenin subunits, *Protein Expr. Purif.* 25 (2002) 97–104, <https://doi.org/10.1006/prep.2001.1614>.
- [107] A.A. Tungekar, A. Castillo-Corujo, L.W. Ruddock, So you want to express your protein in *Escherichia coli*? *Essays Biochem.* 65 (2021) 247–260.
- [108] K.E. Griswold, N.A. Mahmood, B.L. Iverson, G. Georgiou, Effects of codon usage versus putative 5'-mRNA structure on the expression of *Fusarium solani* cutinase in the *Escherichia coli* cytoplasm, *Protein Expr. Purif.* 27 (2003) 134–142.
- [109] C. Gustafsson, S. Govindarajan, J.2. Minshull, Codon bias and heterologous protein expression, *Trends Biotechnol.* 22 (2004) 346–353.
- [110] A. Vitale, R.S. Boston, Endoplasmic reticulum quality control and the unfolded protein response: insights from plants, *Traffic* 9 (2008) 1581–1588, <https://doi.org/10.1111/j.1600-0854.2008.00780.x>.