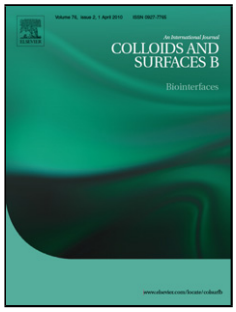
Accepted Manuscript



Title: Insights into Gliadin Supramolecular Organization at Digestive pH 3.0

Authors: M.G. Herrera, D.S Vazquez, R. Sreij, M. Drechsler, Y. Hertle, T. Hellweg, V.I. Dodero

PII: S0927-7765(18)30128-0
DOI: https://doi.org/10.1016/j.colsurfb.2018.02.053
Reference: COLSUB 9193

To appear in: *Colloids and Surfaces B: Biointerfaces*

Received date: 21-10-2017
Revised date: 28-12-2017
Accepted date: 23-2-2018

Please cite this article as: M.G.Herrera, D.S Vazquez, R.Sreij, M.Drechsler, Y.Hertle, T.Hellweg, V.I.Dodero, Insights into Gliadin Supramolecular Organization at Digestive pH 3.0, Colloids and Surfaces B: Biointerfaces https://doi.org/10.1016/j.colsurfb.2018.02.053

Insights into Gliadin Supramolecular Organization at Digestive pH 3.0.

M. G. Herrera [1], D. S Vazquez [2], R. Sreij [3], M. Drechsler [4], Y. Hertle [3], T. Hellweg [3], V. I. Dodero[1#]

[1] Department of Chemistry, Organic Chemistry III, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany.

[2] Instituto de Investigaciones Biotecnológicas (IIB)-Instituto Tecnológico de Chascomús (INTECH), Campus Miguelete Universidad de San Martín (UNSAM), Av. 25 de Mayo y Francia, (1650) San Martín, Buenos Aires, Argentina.

[3] Department of Chemistry, Physical and Biophysical Chemistry, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany.

[4] Bavarian Polymer Institute, KeyLab Electron and Optical Microscopy, Bayreuth University, Universitässtr. 30, 95447 Bayreuth.

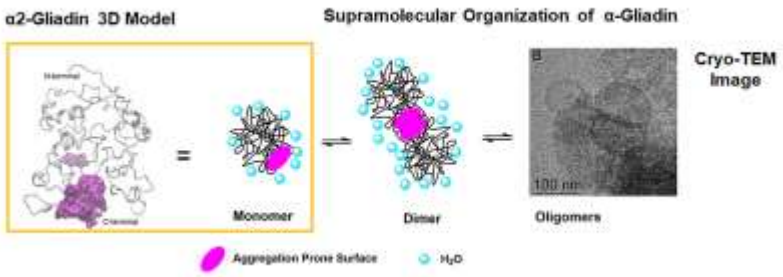
[#]Correspondence should be addressed to veronica.dodero@uni-bielefeld.de

Keywords: gliadin, oligomers, SAXS, molecular simulation, celiac disease, gluten-related disorders.

Manuscript Statistics

Text and Image Headings: 4699 words
Abstract: 237 words
Figures: 6
Tables: 2
References: 41

GRAPHICAL ABSTRACT



Highlights

Gliadin is an immunogenic protein present in wheat gluten which is not fully degraded by humans.

At digestive pH, gliadin solution is composed of prolate ellipsoidal dimers, and oligomers.

A molecular understanding of oligomerization propensity is provided.

Abstract

Alpha-gliadin is a highly immunogenic protein from wheat, which is associated with many human diseases, like celiac disease and non-celiac gluten sensitivity. Because of that, gliadin solution is subject to intense biomedical research. However, the physicochemical nature of the employed gliadin solution at physiological pH is not understood. Herein, we present a supramolecular evaluation of the alpha-gliadin protein in water at pH 3.0 by dynamic light scattering (DLS), cryo-transmission electron microscopy (cryo-TEM) and small-angle-.X-ray scattering (SAXS). We report that at 0.5 wt % concentration (0.1 mg/ml), gliadin is already a colloidal polydisperse system with an average hydrodynamic radius of 30 ± 10 nm. By cryo-TEM, we detected mainly large clusters. However, it was possible to visualise for the first time prolate oligomers of around 68 nm and 103 nm, minor and major axis, respectively. SAXS experiments support the existence of prolate/ rod-like structures. At 1.5 wt % concentration gliadin dimers, small oligomers and large clusters coexist. The radius of gyration (R_{g1}) of gliadin dimer is 5.72 ±0.23 nm with a dimer cross-

section (R_c) of 1.63 nm, and an average length of around 19 nm, this suggests that gliadin dimers are formed longitudinally. Finally, our alpha-gliadin 3D model, obtained by *ab initio* prediction and analysed by molecular dynamics (MD), predicts that two surfaces prone to aggregation are exposed to the solvent, at the C-terminus. We hypothesise that this region may be involved in the dimerisation process of alpha-gliadin.

Introduction

Gliadin is an immunogenic protein present in wheat gluten which is not entirely degraded by humans. It is composed of different isoforms which are classified as α, β, ω, ϒ depending on their electrophoretic mobility [1]. Even though the whole gliadin system exhibits low solubility in water, it can be solubilised using different salts and alcohols [2,3]. Moreover, it has been demonstrated that alpha-gliadin and its partially degraded fragments trigger an adaptive and innate immune response in some individuals which causes the development of celiac disease and non-celiac gluten sensibility [4,5]. These disorders are highly prevalent in the population, affected 1 % and 7 %, respectively. Up to now, the physicochemical understanding of gliadin under physiological conditions is limited. Nevertheless, biomedical studies confirmed their pathogenic role [6]. When food is ingested, it is retained for 120 minutes in the stomach, where the pH can be as low as a 1, but then a buffering effect is induced, leading to pH values in the range of 2.0 to 6.0 [7]. In the case of commercial gliadin, the buffering effect occurs at pH 3.0 when it is dissolved in 0.001 M HCl. Also, mainly the alpha isoform remains in the aqueous phase [8]. Under dilute concentration, gliadin is self-organised forming micelle-type aggregates as characterised by spectroscopic methods and conventional electron microscopy [8]. This study aims to get a better insight into gliadin supramolecular organisation in dilute aqueous solution at pH 3.0 and 10 mM NaCl using dynamic light scattering (DLS), cryogenic transmission electron microscopy (cryo-TEM) and small angle X-ray scattering (SAXS) experiments. Our findings reveal that gliadin under the mentioned conditions is highly polydisperse, composed of prolate ellipsoidal dimers, and oligomers of different sizes. Finally, taking into consideration that the gliadin structure has not been elucidated yet, we perform an *ab initio* modelling of alpha-gliadin and a classical molecular dynamics (MD) simulation to provide insights into the 3D structure of alpha-gliadin and its oligomerisation capability.

Materials and Methods

Sample preparation

Gliadin stock solution was obtained and characterised as described before [8]. Briefly, 10 mg of gliadin (Sigma Aldrich) was dissolved in 10 ml of HCl 0.001 M and NaCl 10 mM obtaining a final pH of 3.0. Gliadin solution was homogenised at room temperature, stabilised for 12 hours at 10°C and centrifuged at 21130 g for 15 min; the supernatant was taken as the working solution. The composition of the supernatant was mainly alpha gliadin and a minor proportion of the beta isoform. The supernatant has a final concentration of 0.5 wt% as determined by a Bradford assay. For SAXS experiments, the final concentration of 1.5 wt % was reached using a centrifugal filter with a 30.000 Da cut off (Amicon Ultra, Millipore).

Dynamic light scattering (DLS)

DLS experiments were carried out using a 3DLS spectrometer (LS Instruments, Fribourg, Switzerland). The light source was a He-Ne laser of 632.8 nm with a constant output power of 30 mW. The scattering experiments were performed with an angular range of $40° > \theta > 110°$. Samples were kept in the thermostated goniometer of the DLS instrument at 25 °C between 15 to 20 minutes to reach thermal equilibrium. The measured intensity time autocorrelation functions were converted to the respective field correlation functions and analysed using the CONTIN software of S. Provencher [9]. The outcome is an intensity weighted relaxation rate distribution which can be converted into a distribution of the corresponding apparent diffusion coefficients D_{app} of the present structures. By the Stokes-Einstein equation ($R_h=k_BT/6\pi\eta D_{app}$), the radii of hydrodynamically equivalent spheres (R_h) are calculated. Here, kB is the Boltzmann constant, T the temperature and η the solvent viscosity. Gliadin working solution was filtered through 0.2 µm syringe filter (Millex, low protein binding membrane) before DLS measurements.

Cryo-transmission electron microscopy (Cryo-TEM)

Gliadin solution was deposited on lacey carbon filmed copper grids by blotting with a filter paper. The resulting thin film was vitrified by quickly plunging the grids into liquid ethane at its freezing

point. Specimens were examined at a temperature around 90 K with a Zeiss/LEO EM922 Omega TEM (Zeiss SMT, Oberkochen, Germany). Images were recorded by a CCD digital camera (Ultrascan 1000, Gatan) and analysed using the GMS 1.8 software (Gatan). The size distribution was obtained using Image J [10].

Small angle X-ray scattering (SAXS)

SAXS of gliadin was performed on a Xeuss SAXS/WAXS System (Xenocs, Sassenage, France). The X-ray source consists of a 30 W CuK_α (λ=1.54 Å) microfocus tube with an ultra-low divergence mirror optics (GeniX, Sassenage, France) and a Pilatus 300K, 20 Hz hybrid pixel detector (DECTRIS, Baden-Deatwil, Switzerland). The scattering curves were normalised by the integrated intensity incident on the sample, exposure time, sample thickness, transmission and background. The focus of the sample was 1 mm². The samples were measured in a 1 mm flow-through Kapton capillary (Goodfellow GmbH, Bad Nauheim, Germany) placed in a Linkam stage (Linkam Scientific, Tadworth, UK) at 10 °C. The sample-to-detector distance was for SAXS 2772 mm, calibrated with silver behenate. The range of the scattering vector $q = \frac{4\pi}{\lambda}\sin(\theta)$ was around q=0.7 - 0.13 nm^{-1} at a scattering angle of 2θ. The absolute calibration of the scattering data was done with glassy carbon type 2 [11]. The scattered intensity $I(q)=N\cdot I(0)\cdot(\Delta\rho)^2\cdot V^2\cdot P(q)\cdot S(q)$ is dependent on: the incident intensity $I(0)$, the scattering volume of the colloid V, the electron density difference $\Delta\rho$ between particle and solvent and the particle form factor $P(q)$. The particle structure factor $S(q)$ equals to one for dilute systems.

Prediction of alpha-gliadin structure and classical molecular dynamics simulation (MD)

The alpha-gliadin sequence Q9M4L6 from the UniProt database was selected for the bioinformatics and molecular dynamics (MD) analysis. For the *ab initio* model of alpha-gliadin, the I-TASSER online server was employed to get the best structure prediction [12]. The structure with the best energy-score was then subjected to classical MD simulation. The initial coordinates were solvated with 15363 molecules of the 3-particle TIP3P water model [13] in a 10 Å cubic box (7.7

x 8.2 x 7.5 nm dimension). In our model, the protonation state of each residue was set to the corresponding state at neutral pH. One chloride ion was added to neutralise the net charge of the system. A minimisation was applied to the resulting structure to remove any clashes for 1000 steps of steepest descent followed by 2000 steps of conjugate gradient minimisation procedure using constant volume periodic boundary conditions. The system was subsequently heated from 10 K to 300 K, using the Berendsen thermostat [14] to suppress fluctuations of the kinetic energy of the system with a relaxation constant of 2 *ps.* Subsequently, the system was switched to constant isotropic pressure for allowing the density to equilibrate. The SHAKE algorithm [15] was applied to all bonds involving hydrogen atoms; then a 2 fs time step was settled. An 8 Å cutoff radius for range-limited interactions with the particle mesh *Ewald* summation for long-range electrostatic interactions was used. Harmonic positional restraints of the strength of 20 kcal mol^{-1} $Å^{-2}$ on C*α* atoms were applied during minimisation and equilibration. These restraints were removed in successive simulation stages (15, 10, 5, 2 and 0 kcal mol^{-1} $Å^{-2}$) of 5 ns each. Production unrestrained MD simulation was performed for 500 ns with the Amber14 suite using the *pmemd.*CUDA engine for the GPU code [16] and the *ff14SB* force field [17]. Post-processing analysis was done using the *cpptraj* routine of AmberTools15 [18]. Protein representations were prepared in VMD 1.9.2 [19].

Bioinformatics prediction of gliadin hydrophobicity, intrinsic disorder regions and aggregation tendency.

Q9M4L6 sequence from UniProt database was analysed using the IUPred server to predict its intrinsical disorder [20]. For that, the long prediction scale type was selected. The hydrophobicity of the protein was studied using the Kyte-Doolittle scale [21] from ProtScale of the ExPasy server [22]. Finally, the aggregation propensity of this protein was assessed and compared using the following servers: Tango and AGGRESCAN [23-25].

Results and Discussion

DLS and cryo-TEM evaluation of gliadin oligomers

In the present work, we use non-invasive methods such as DLS and cryo-TEM experiments to characterise gliadin supramolecular structures. DLS experiments reveal a broad distribution of hydrodynamic radii (R_h) in the gliadin solution, indicating the presence of different oligomeric species in solution. The multiple species are probably generated by the dynamic interactions of gliadin monomer and oligomers (**Fig. 1A**). The majority of the particles have a R_h of 30 ± 10 nm, which was in agreement with the isolated oligomers that we previously observed by TEM [8]. However, the hydrodynamic size distribution is asymmetric and shows a long tail. This kind of behaviour was previously observed in TiO_2 particles which do not have a spherical shape, suggesting that some bigger oligomers or interacting small ones are present in solution, producing non-spherical aggregates [26].

In our previous TEM and scanning electron microscopy (SEM) [8] results, we detected gliadin oligomers. However, these methods require drying steps and staining which could generate changes in the morphology and oligomerisation state of the specimen [27]. In this context, cryo-TEM was successfully used to characterise other proteins aggregates such as *α*-synuclein oligomers [28], and amyloid *β*-peptide channel-like protofibrils [29,30]. In the case of alpha-gliadin, cryo-TEM experiments allow us to visualise gliadin oligomers in a native-like state. By this procedure, we mainly observed clusters of oligomers (*see* Fig.S1) which contained in some cases detectable ellipsoidal/prolate structures, as presented in **Fig. 1B**. The minor and major axis of the ellipsoidal oligomers is 68.13 ± 7.92 nm and 103.7 ± 18.9 nm, respectively. This morphology was previously reported by Sato *et al.* [31] and Thomson *et al.* [32], for different experimental conditions. However, to the best of our knowledge, such prolate structures are visualised by cryo-TEM for the first time in the present study. It was not possible to visualise low-molecular weight oligomers, probably due to the high cluster formation [33].

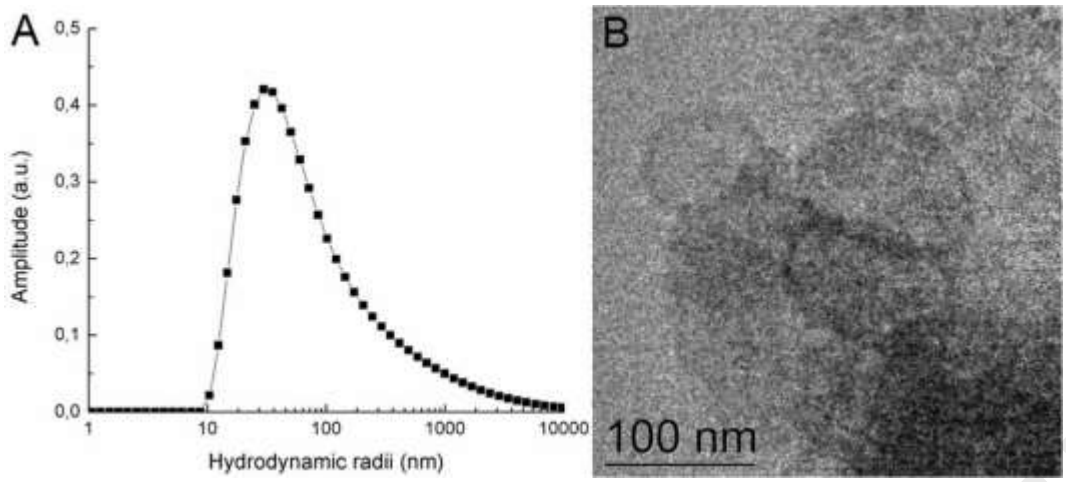


Figure 1. DLS experiment and cryo-TEM of alpha-gliadin in HCl 0.001M, 10 mM NaCl at pH 3.0. **A**) R_h distribution obtained by DLS. **B**) High magnification image of gliadin oligomers obtained by cryo-TEM (*see* Fig.S1).

SAXS of gliadin oligomers

Sato *et al.* showed that α and γ gliadins exist in aqueous media at low concentration as monomers whereas increasing the concentration led to the formation of oligomers as determined by SAXS [31]. This report was mainly focused on high concentration experiments in water and with different salts (the pH of the water was not mentioned). We are more interested in the early stages of gliadin oligomerisation at a dilute concentration at pH 3.0 because it is the concentration range and pH where digestion takes place. We tested the system at two different gliadin concentrations, 0.5 and 1.5 wt%. The morphology of the oligomers present in the solution can be approximated by a scaling law, which yields a relation between the scattering intensity *I(q)* and the scattering vector *q* Eq. (1) in the low *q* regime of the scattering curve.

$$I(q) \propto \frac{1}{q^{\alpha}} \Leftrightarrow \ln\big(I(q)\big) \propto -\alpha \cdot \ln(q) \qquad (1)$$

Here, *α* is the scattering exponent indicating the morphology of the structure, i.e. for a spherical object *α*=0, for a cylinder *α*=1 and a disk *α*=2 [34]. The scattering profiles of gliadin samples with the concentrations of 0.5 and 1.5 wt% are shown in **Fig. 2A**. The thick lines represent linear fits with the slope -*α*, indicating the linear dependency between ln(*I(q)*) and ln(*q*) (*see* Eq. (1)). Here, *α*=1.0 at 0.5 wt% confirms a rod-like structure in the dilute system. At 1.5 wt%, the exponent was

found to be α=1.2, confirming the presence of mainly cylindrical particles at this concentration, too. The deviation from the theoretical value of 1.0 may arise either from aggregation effects of the oligomers or is caused by contributions in the scattering profile from an underlying structure factor [35]. In the following, we calculate the structural parameters of the samples by a Guinier analysis. **Fig. 2B** shows a Guinier plot with ln($I(q)$) against q^2 for gliadin with a concentration of 0.5 and 1.5 wt% under the aforementioned experimental conditions. Due to the complexity of the system, the scattering profiles could not be fitted with a simple Guinier relationship. This result is in agreement with a polydisperse system. Hence, we applied a similar analysis scheme as Sato *et al.* which consists of a three-component Guinier (Eq. (2)) to describe the $I(q)$ data.

$$I(q) = A_1 \cdot \exp\left(-\frac{R_{g1}^2 q^2}{3}\right) + A_2 \cdot \exp\left(-\frac{R_{g2}^2 q^2}{3}\right) + A_3 \cdot \exp\left(-\frac{R_{g3}^2 q^2}{3}\right) \quad (2)$$

In this case, R_g denotes the radius of gyration, and A is a pre-exponential factor. Being aware of the validity limit of the Guinier approach given by qR_g=1.3, we show the fit for the entire scattering curve. The analysis of the scattering curves by Eq. (2) reveals the presence of three main species (**Table 1**). The minor radius (R_{g1}) represents minimal gliadin structure. This value increases with increasing gliadin concentration in solution (3.62 ± 2.63 nm at 0.5 wt% to 5.72 ± 0.23 nm at 1.5 wt%). Our results are in agreement with the results of Sato *et al.* [31]. The large error of R_{g1} at 0.5 wt% arises from the poor statistics of the scattering profile. Two other radii are observed (R_{g2} and R_{g3}). The second contribution arises from the formation of gliadin oligomers with increasing protein concentration in solution. R_{g2} is computed to 21.05 ±12.76 nm for 0.5 wt% and 64.00 ± 2.87 nm for 1.5 wt%, respectively. These values are also in agreement with the average radii calculated by DLS. Moreover, the third component of Eq. (2) yields R_{g3} =240.09 ± 36.64 nm and 587.85 ± 19.06 nm at the two concentrations, respectively. These experiments confirm the presence of rather large aggregates in the sample in agreement with the observed tailing of the hydrodynamic radius distribution and the cryo-TEM findings.

The intermediate region of the scattering data can be approached by the Guinier-Porod model (Eq. (3)) which relates the Guinier and the Porod region [36] and is applicable for spherical and non-spherical objects.

$I(q) = \frac{G}{q^s}\exp\left(\frac{-q^2R_g^2}{3-s}\right)$ for $q \leq q_1$ and $I(q) = \frac{D}{q^\alpha}$ for $q \geq q_1$. (3)

Where q_1=1/R_g(3α/2)$^{1/2}$ relates the radius of gyration obtained from the Guinier analysis with the scattering exponent α from the scaling law in the Porod (low *q* region) region. G and D are the Guinier and Porod scaling factors, respectively. The *s* – parameter which is derived from the dimensionality parameter *3-s* becomes for rods *s* =1. With *α*=1 and *s*=1 for the present system, this so-called Guinier-Porod model [36] can be applied to this system for $Q_2 \leq Q \leq Q_1$ to derive information about the cross-section radius R_c and thus of the length of the gliadin monomers. R_c can then be derived by the relation $R_{gc} = \frac{R_c}{\sqrt{2}}$ and determined from the slope *m* of the intermediate-*q* Guinier plot ln(*q*·*I*(*q*)) against q^2 as shown in **Fig. 2C**. With $= -\frac{1}{2}R_{gc}^2 = -\frac{1}{2}\left(\frac{R_c}{\sqrt{2}}\right)^2 = -\frac{R_c^2}{4}$, the cross-sectional radius is calculated by $R_c = \sqrt{-4m}$, where *m* is slope of the linear fit. In **Fig. 2C**, the limit condition qR_c =1 in which this approximation holds true is underlined. The results are presented in **Table 1**. We observe that R_c remains nearly constant for both concentrations and is in accordance with the values calculated by Thomson *et al.,* when gliadin was dissolved in ethanol-water mixtures. By taking the relation between R_{g1} and R_c into account $R_{g1} = \left(\frac{\Lambda^2}{12} + \frac{R_c^2}{2}\right)$ and assuming cylindrical shape for the gliadin molecules, the rod length *Λ* of the gliadin monomers can be calculated by *Λ*= √(12 R_{g1}^2 -6 R_c^2) as presented in **Table 1**, too. By this calculation, we obtain a rod length of 12 and 19 nm for 0.5 wt% and 1.5 wt%, respectively. Sato *et. al* obtained similar values at dilute concentrations. **Fig. 2D** shows a Kratky plot (q^2·*I*(*q*) against q^2 is shown) for both concentrations which indicates rather disordered, unfolded structures present in solution, as previously reported [8,37,38].

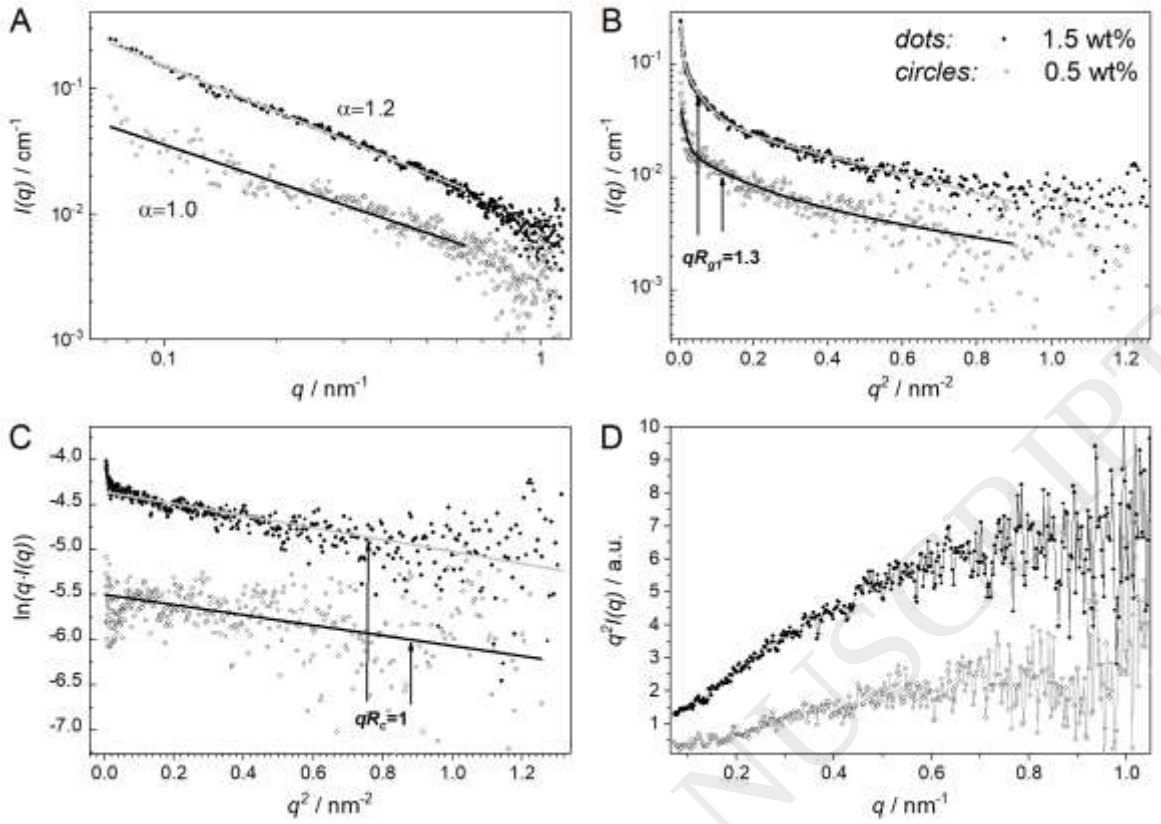


Figure 2. *SAXS data.* The solutions at 0.01 M HCl with protein concentrations of 0.5 wt% and 1.5 wt% were measured at 10 °C. **A**) log-log plot of the SAXS data of the two dilute gliadin solutions with linear fits showing the scattering exponent *α*. **B**) Guinier plots for the two dilute gliadin solutions. The scattering profiles are fitted with a three-component Guinier equation (lines). The validity condition qR_g=1.3 is shown for R_{g1}. **C**) Guinier-Porod plot for concentrations of 0.5 and 1.5 wt% of gliadin at pH 3.0 and 10 °C. The limits qR_c=1 for the model to be valid are shown. **D**) Kratky-plot ($q^2 \cdot I(q)$ against q) of the gliadin solutions at 0.5 and 1.5 wt% at pH 3.0.

Comparison of the cross-section values (R_c) of 1.50 ± 0.025 nm for gliadin monomer and 1.63 ± 0.01 nm for gliadin dimer and the average length of 11.98 ± 0.11 nm and 19.41 ± 0.01 nm, for gliadin monomer and dimer respectively, suggesting that the gliadin dimers are formed longitudinally.

Conc (wt%)	R_{g1} / nm	R_{g2} / nm	R_{g3} / nm	R_c / nm	Λ / nm
0.5	3.62 ± 2.63	21.05 ± 12.76	240.09 ± 36.64	1.50 ± 0.025	11.98 ± 0.11
1.5	5.72 ± 0.23	64.00 ± 2.87	587.85 ± 19.06	1.63 ± 0.01	19.41 ± 0.01

Table 1: Radius of gyration (R_g) of gliadin at pH 3.0 calculated by fitting the SAXS profiles in Figure 2B with a three-component Guinier equation. The cross-section radius of gyration (R_c) and the average length (Λ) were calculated from the Guinier-Porod plot of R_{g1}.

Bioinformatics analysis, structural prediction and MD simulation of α-gliadin

The high oligomerisation/aggregation propensity of gliadin complicates the elucidation of the 3D structure of all gliadin isoforms by common high-resolution techniques like NMR or X-ray crystallography. To get a better understanding of the oligomerisation behaviour of gliadin at the atomistic level, we present our results by the use of different computational tools. As a basis for the numeric calculations, the alpha-gliadin sequence from the UniProt database with code Q9M4L6 was employed (**Table 2**). The amino acid composition is mainly glutamine (34.2%) and proline (15.6%), but also valine and serine are present with 5.2% in high percentage. Tyrosine and phenylalanine are present in a 3.1% and 3.9 %, respectively; in contrast, only one tryptophan residue is present. The hydrophobic profile of the alpha-gliadin primary structure was assessed using the Kyte-Doolittle scale (**Fig. 3A**) [21]. From this hydrophobicity plot, it is shown that α-gliadin is mainly hydrophilic, in the regions 25-151 (blue box) and 225-275 (green box). In the C-terminal region, there are short hydrophobic patches in the regions from 150 to 200 and from 275 to 290 and are composed of 5 to 10 residues each.

Alpha-gliadin, >tr|Q9M4L6|Q9M4L6_WHEAT α-Gliadin OS=Triticum aestivum
MVRVPVPQLQPQNPSQQQPQEQVPLVQQQQ[30]FPGQQQPFPPQQPYPQPQPFPSQQPYL QLQ[60]PFPQPQLPYP70QPQLPYPQPQLPYPQPQPFRPQQPYPQSQP[100]QYSQPQQPISQQ QQQQQQQQQQKQQQQQQQ[130]QILQQILQQQ[140]LIPCRDVVLQQHSIAYGSSQVLQQSTYQ LV[170]QQLCCQQLWQIPEQSRCQAIHNVVHAIILH[200]QQQQQQQQQQ[210]QQPLSQVSFQQPQ QQYPSGQGSFQPSQQNP[240]QAQGSVQPQQLPQFEEIRNLALETLPAMCN[270]VYIPPYCTIA[2 80] PVGIFGTNYR[290]

Table 2. The sequence of alpha-gliadin used for the computational analysis obtained by UniProt database.

To evaluate which regions of the protein are unfolded, we employed the intrinsic disorder analysis with IUPred [20]. This server predicts regions with a lack of structure by the estimation of their total pairwise inter-residue interaction energy considering that the sequences do not fold due to their inability to form sufficient stabilising inter-residue interactions. It was computed that gliadin presents highly disordered regions along the molecule (**Fig. 3B**), which is in concordance with previous secondary structure experiments [8,37,38]. Interestingly, the unfolded regions correspond to the previously mentioned hydrophilic regions (25-125 and 225-250), suggesting that they may be highly exposed to the solvent.

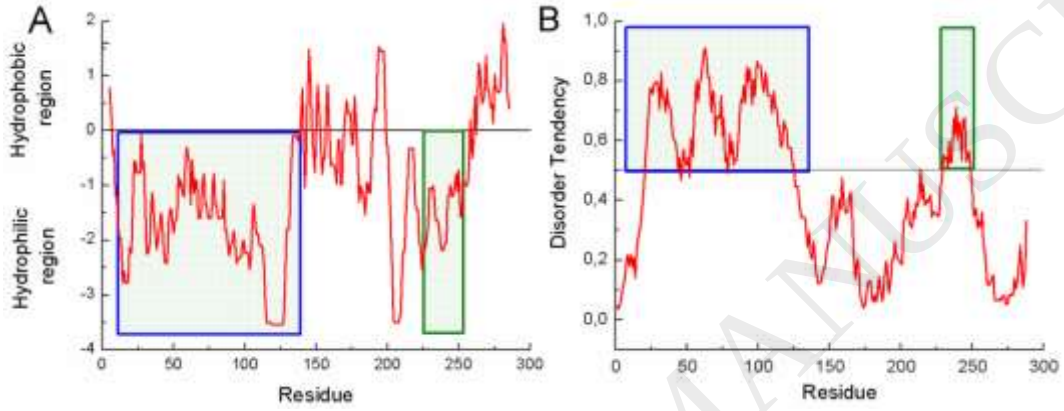


Figure 3. *Bioinformatics analysis.* A) Hydrophobicity plot of the alpha-gliadin sequence performed using Kyte-Doolittle parameters. B) Disorder regions predicted using the IUPred server. Scores above 0.5 indicate disorder. The blue and the green boxes indicate the residues 25-151 and 225 to 250 which are hydrophilic and have highest disorder tendency.

Tridimensional predictions of *α*/*β*-, *γ*- and *ω*- gliadin isoforms, but not of alpha-gliadin isoform, have been computed using a hierarchical approach based on the sequences of these proteins [39]. Here, the alpha-gliadin sequence was retrieved from UniProt as aforementioned and uploaded to the I-TASSER web server, (best-ranked protein structure prediction server in the last CASP12 edition (Critical Assessment of Techniques for Protein Structure Prediction)), to predict a set of possible structures. Then, the best energy-scored structure was subjected to a 500 ns long explicit-water MD simulation to test the stability of the predicted structure and get insights into the native ensemble of the alpha-gliadin isoform in solution.

An initial inspection of the best-scored model obtained by I-TASSER hints at an ellipsoidal conformation. A reduced number of residues with ψ / φ values in the canonical Ramachandran distribution were found where 48 residues (16.7 %) are in the outline of the probability zones (**Fig. 4A** and Fig.S2). Next, we applied a minimisation and equilibration protocol before MD production, where the backbone was slowly freed in successive harmonical restrained simulations. **Fig. 4B** shows the last frame of the unrestrained equilibration step before the MD production (**Fig. 4C**). Here, 8 residues (2.8 %) are outliers of the more favourable zones, whereas in the last frame of the MD production, only 4 residues (1.4 %) are outliers and 98.6 % are in favoured regions of the Ramachandran distribution. This result exemplifies how the molecular simulation corrects the dihedral backbone angles.

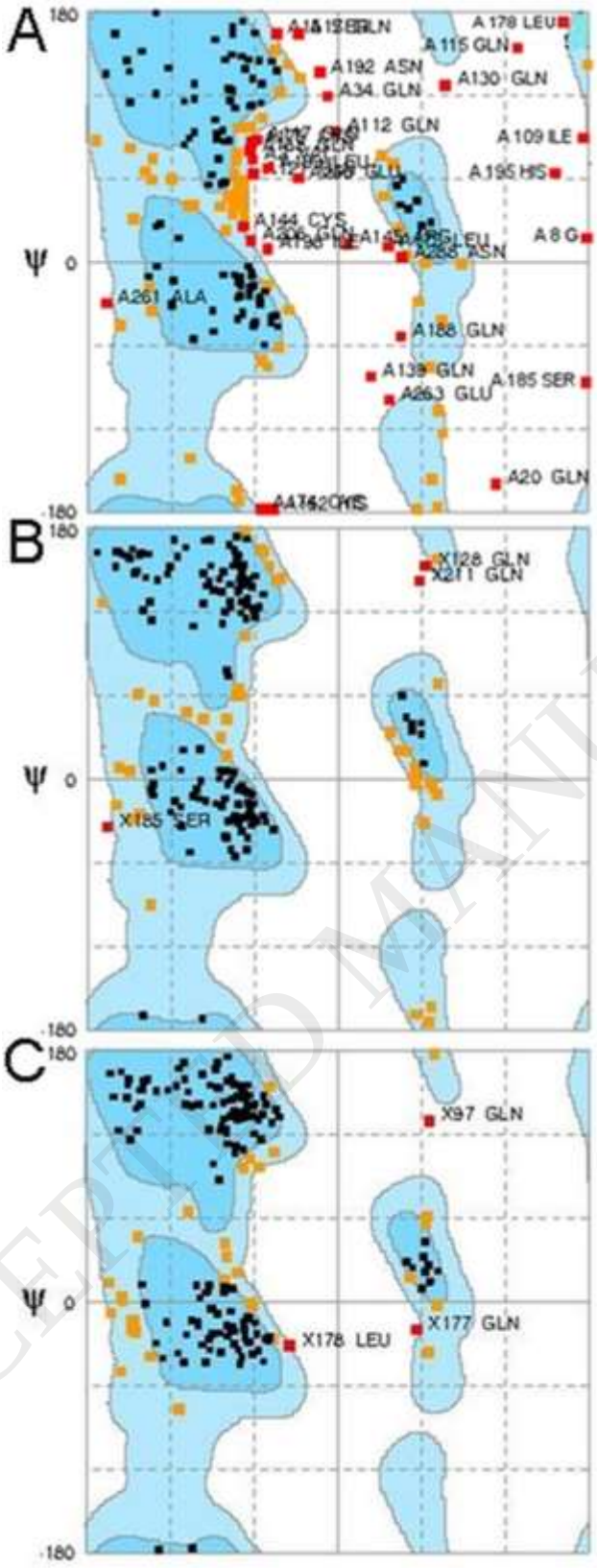


Figure 4. *Ramachandran plot*. A) The distribution of ψ/φ values for the I-TASSER best-scored mode, B) the last frame of MD equilibration and C) the last frame of MD production. The red symbols denote residues outside the probability zone, the orange symbols are residues in the less favoured probability zones, and the black symbols are those residues in the more favoured

zones. The plots were performed using the RAMPAGE online tool (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php).

More than the 90 % of the residues are in an unstructured/disordered conformation, and only ~8 % in helical structure as was calculated by the DSSP algorithm [40]. This result is in agreement with previous circular dichroism experiments [8], denoting a protein with a low secondary structure content. In the same way, a low secondary structure acquisition was observed during the MD production. Here, 86 % of the residues remain in non-canonical secondary structure conformation, and the helical content increases only to ~11 % in average over the production simulation. The remaining 3% are in *β* conformation (**Fig. 4C**).

The root-mean-square (RMS) profile of α-gliadin, achieved through the LOVO fit algorithm, shows a stable structure along the simulation. The best-aligned region has an average RMSD value of 0.2 nm and with a small increase of 0.3 nm at the end of the simulation (**Fig. 5A**) but it does not exhibit local unfolding. The plot of the RMS *per* residue shows a significant fluctuation at the end of the protein, around 30 residues in the N-terminus and the last 10 residues in the C-terminus (**Fig. 5B**). Whereas and due to the unstructured behaviour of α-gliadin, the number of native contacts (**Fig. 5D**) is unchanged during the simulation suggesting a stable and suitable α-gliadin model. The radius of gyration (R_g) shows a typical distribution along the simulation with a maximum frequency around 2.13 nm which is of the order of magnitude of the R_{g1} calculated by SAXS (*see* **Table 1** and Ref. 31). The observed differences may arise from a higher hydration of the protein in solution.

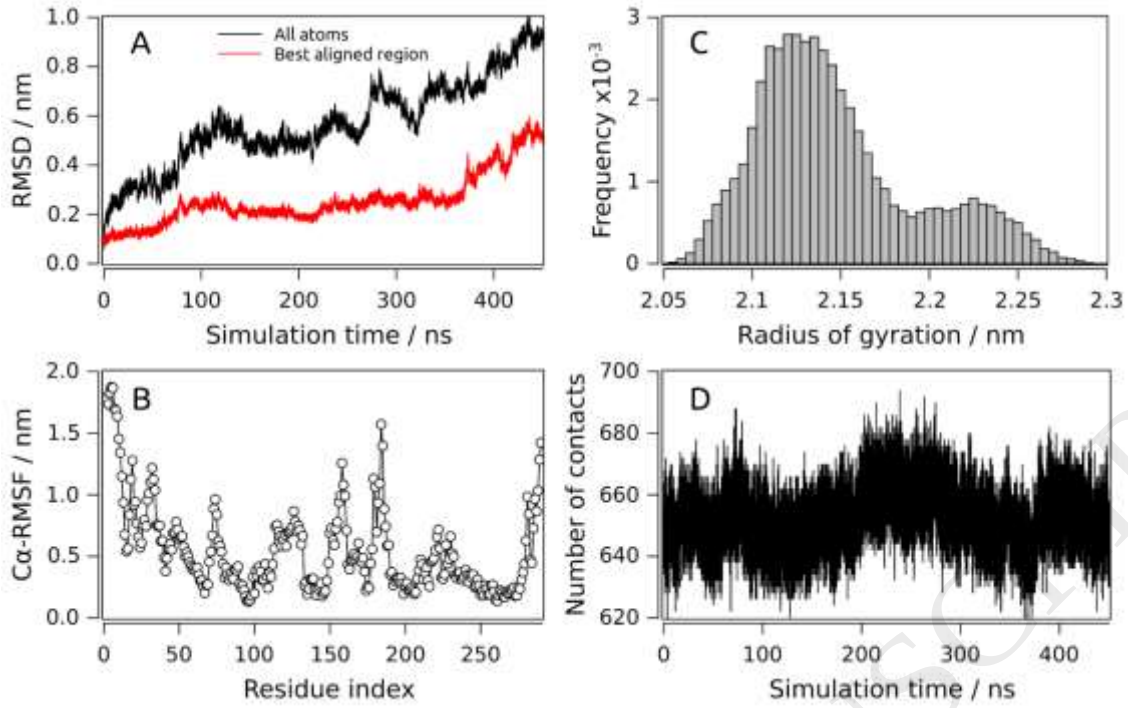


Figure 5: *MD simulation analysis*. A) RMSD for all atoms (black line) and the best-aligned fraction (red line) obtained by the LOVO fit analysis. The parameter φ used to perform the fitting was 0.65, and the RMSD profile of those residues are shown in red (best-aligned region). B) From the same analysis, the C-*α* RMSF is shown. The first 50 ns was not taken into account for RMS analysis to diminish initial conformation bias. C) Histogram of the radius of gyration (R_g) over the whole trajectory. The R_g of maximal frequency was 2.13 nm. A second maximum at 2.24 nm was also obtained. D) The number of contacts over the simulation time is shown. The cutoff radius set over the alpha carbon atoms was 5 Å.

The topology of the last frame of the molecular dynamics is shown in **Fig. 6A** (*see also* **TableS3**). Next, considering that gliadin can self-assemble, we evaluated the potential aggregation "nuclei" of this protein by sequence-based methods. For that, we employed two predictors, AGGRESCAN and Tango (see Materials and Methods), mapping the residues tendency towards aggregation. Tango predicted that residues 132-137, 147-151, 166-170, 192-199, 281-289 are prone to aggregation meanwhile AGGRESCAN predicted the residues 169-178, 190-198, 267-273, 276-280 and 282-290. Both methods predict that the C-terminal region is prone to aggregation especially in the regions 190-199 and 281-290 (**Fig. 6B**). A further analysis of the sequence and the parameters obtained by MD (*see* **TableS3**) showed that the region 143-149 is unstructured, while between residues 147 to 151 hydrophobic amino acids as valine and isoleucine are present,

which are prone to aggregation. On the other hand, the regions 171-179 and 191-200, are turns and extended structures. Both regions are predicted to aggregate. It is interesting to notice that the 171-179 region that has a high content of glutamine has been reported to interact with the CXCR3 receptor, inducing zonulin release and increasing the gut's permeability [41]. In the region 191-200 are valine, isoleucine and alanine highly frequent, contributing to the hydrophobicity of the region and favouring the observed self-assembly. In the C-terminus region, residues from 265 to 285, a turn and 3_{10} helix structure is predicted; this region has a high propensity for aggregation due to the hydrophobic nature of the amino acids, mainly alanine, valine, isoleucine and aromatic amino acids as tyrosine and phenylalanine.

Finally, a three-dimensional analysis of the regions predicted here was done taking the structure of the last frame of the MD simulation and observing the position of each region predicted to aggregate in the model obtained. We observed that these residues are exposed to the solvent generating two surfaces that are shown in **Fig. 6C**. We hypothesise that these regions may be the surfaces in which protein-protein interaction might occur favouring the formation of gliadin dimers longitudinally.

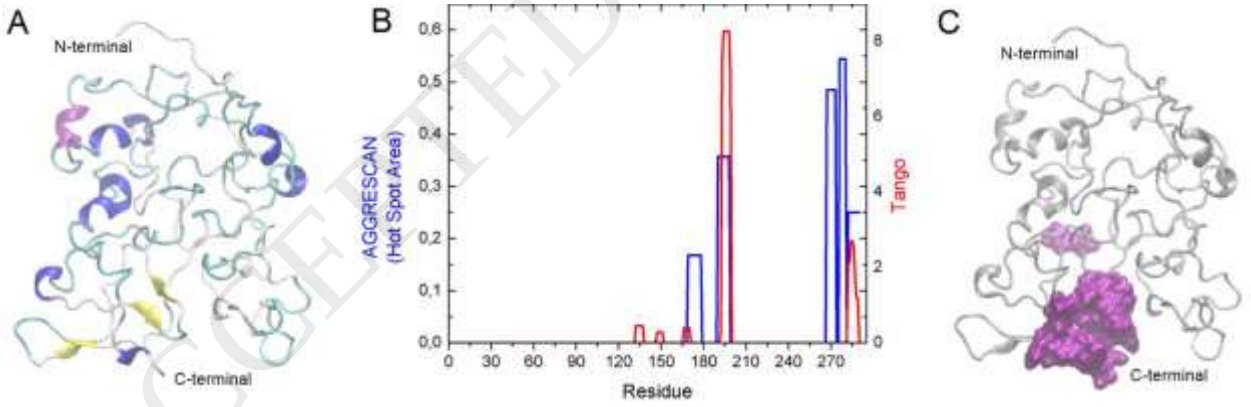


Figure 6. *Structural visualisation of gliadin*. A) The topology of the last frame of α-gliadin obtained by the MD simulation; showing α-helix (magenta), 3-10 helix (blue), β-structure (yellow), turns (green), random coils (white), (*see* TableS3). B) Graphical representation of aggregation profile of α -gliadin, obtained by Tango (red) and AGGRESCAN (blue). C) Surface representation of the residues predicted to aggregate (magenta) in the alpha-gliadin structure by AGGRESCAN and Tango, showing the existence of two possible regions of protein-protein interaction.

Conclusion

In this work, we analyse the supramolecular assembly of gliadin in diluted solution at pH 3.0. This pH is a typical value in the digestive process, and therefore the behaviour of this immunogenic and proteolytically stable protein has a paramount importance. Under these conditions at close to physiological ionic strength, gliadin is found to form aggregates with a broad range of sizes and a tendency to form elongated structures which were visualised by cryo-TEM and computed by SAXS experiments (the scattering exponent of ≈1). To better understand the molecular nature of the initial steps in the self-assembly process, we have performed MD simulations to generate energy minimised structures of the protein and used different bioinformatics predictors to obtain information about hydrophobicity and aggregation propensity. Our alpha-gliadin 3D model obtained by *ab initio* and MD predicts that two surfaces prone to aggregation are exposed to the solvent, at the C-terminus. The occurrence of a hydrophobic surface in the regions of 199-198 and 281-289 amino acids allows the dimerisation process at the C-terminus. These regions are composed mainly of hydrophobic amino acids with a defined extended secondary structure (beta and turns). Altogether this can explain the high propensity of oligomerisation observed, showing that the C-terminus of alpha-gliadin can trigger the oligomerisation process. The occurrence of gliadin oligomers even at diluted conditions can explain the up to now unexplained proteolytical resistance, and thus the hypothesis of incapability of digestive enzymes to access to the degradation sites of gliadin is reasonable. Moreover, considering the reported innate immunological response triggered by gliadin, the hypothesis of colloidal nanostructures of gliadin as real triggers instead of randomly disperse gliadin may explain the reported immune and tissue stress. Our findings of the gliadin oligomerisation might be a first step towards the understanding of gliadin *in vitro* and *in vivo* behaviour.

Acknowledgments

This work was supported by Alexander von Humboldt Foundation and DAAD (German Academics Exchange Service). T. H. and R. S. further acknowledge the financial support by the Deutsche

Forschungsgesellschaft (DFG) (INST 215/432-1 FUGG). V. I. D. thanks, Prof. N. Sewald from Bielefeld University for his support.

References:

[1] K.A. Caldwell, A Comparative Study of Disulphide-Bonding and Molecular Weights of Purified Fractions from Secalin, Gliadin, and Hordein, J. Exp. Bot. 34 (1983) 1411-1420.
[2] T.B. Osborne, The Proteins of Wheat-Kernel; Carnegie Institution of Washington, Washington, 1907.
[3] T. Ukai, Y. Matsumura, R. Urade, Disaggregation and reaggregation of gluten proteins by sodium chloride, J. Agric. Food Chem. 56 (2008) 1122-1130.
[4] L. Shan, O. Molberg, I. Parrot, F. Hausch, F. Filiz, G.M. Gray, L.M. Sollid, C. Khosla, Structural basis for gluten intolerance in celiac sprue, Science 297 (2002) 2275-2279.
[5] F. Hausch, L. Shan, N.A. Santiago, G.M. Gray, C. Khosla, Intestinal digestive resistance of immunodominant gliadin peptides, Am. J. Physiol. Gastrointest. Liver. Physiol. 283 (2002) G996-G1003.
[6] D. Bernardo, J.A. Garrote, L. Fernandez-Salazar, S. Riestra, E. Arranz, Is gliadin really safe for non-coeliac individuals? Production of interleukin 15 in biopsy culture from non-coeliac individuals challenged with gliadin peptides, Gut 56 (2007) 889-890.
[7] S. Bellmann, D. Carlander, A. Fasano, D. Momcilovic, J.A. Scimeca, W.J. Waldman, L. Gombau, L. Tsytsikova, R. Canady, D.I. Pereira, D.E. Lefebvre, Mammalian gastrointestinal tract parameters modulating the integrity, surface properties, and absorption of food-relevant nanomaterials, Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol. 7 (2015) 609-622.
[8] M.G. Herrera, T.V. Veuthey, V.I. Dodero, Self-organization of gliadin in aqueous media under physiological digestive pHs, Colloids Surf. B. 141 (2016) 565-575.
[9] S.W. Provencher, CONTIN: A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations, Comput. Phys. Commun. 27 (1982) 229-242.
[10] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, Nat. Methods 9 (2012) 671-675.
[11] F. Zhang, J. Ilavsky, G.G. Long, J.P.G. Quintana, A.J. Allen, P.R. Jemian, Glassy Carbon as an Absolute Intensity Calibration Standard for Small-Angle Scattering, Metall. Mater. Trans. A 41 (2010) 1151-1158.
[12] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: protein structure and function prediction, Nat. Methods 12 (2015) 7-8.
[13] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926-935.
[14] H.J.C. Berendsen, J.P.M. Postma, W.F.v. Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684-3690.
[15] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys. 23 (1977) 327-341.
[16] D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, Walker, J. Wang, R.M. Wolf, X. Wu, P.A. Kollman, AMBER 14, 2014.
[17] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, J. Chem. Theory Comput. 11 (2015) 3696-3713.
[18] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, J. Chem. Theory Comput. 9 (2013) 3084-3095.
[19] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph 14 (1996) 33-8, 27-8.
[20] Z. Dosztányi, V. Csizmok, P. Tompa, I. Simon, IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, Bioinformatics 21 (2005) 3433-3434.

[21] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105-132.
[22] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: The proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31 (2003) 3784-3788.
[23] A.M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz, L. Serrano, Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins, Nat. Biotechnol. 22 (2004) 1302-1306.
[24] O. Conchillo-Sole, N.S. de Groot, F.X. Aviles, J. Vendrell, X. Daura, S. Ventura, AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides, BMC Bioinformatics 8 (2007) 65.
[25] M. Oliveberg, Waltz, an exciting new move in amyloid prediction, Nat. Methods 7 (2010) 187-188.
[26] W.L. Vos, T.W. Tukker, A.P. Mosk, A. Lagendijk, W.L. Ijzerman, Broadband mean free path of diffuse light in polydisperse ensembles of scatterers for white light-emitting diode lighting, Appl. Opt. 52 (2013) 2602-2609.
[27] M. Winey, J.B. Meehl, E.T. O'Toole, T.H. Giddings, Jr., Conventional transmission electron microscopy, Mol. Biol. Cell. 25 (2014) 319-323.
[28] S.W. Chen, S. Drakulic, E. Deas, M. Ouberai, F.A. Aprile, R. Arranz, S. Ness, C. Roodveldt, T. Guilliams, E.J. De-Genst, D. Klenerman, N.W. Wood, T.P. Knowles, C. Alfonso, G. Rivas, A.Y. Abramov, J.M. Valpuesta, C.M. Dobson, N. Cremades, Structural characterization of toxic oligomers that are kinetically trapped during alpha-synuclein fibril formation, Proc Natl Acad Sci U S A 112 (2015) E1994-2003.
[29] Y. Miller, B. Ma, C.J. Tsai, R. Nussinov, Hollow core of Alzheimer's Abeta42 amyloid observed by cryoEM is relevant at physiological pH, Proc Natl Acad Sci U S A 107 (2010) 14128-14133.
[30] R.F. Thompson, M. Walker, C.A. Siebert, S.P. Muench, N.A. Ranson, An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology, Methods, 100 (2016) 3-15.
[31] 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.
[32] 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: alpha-, gamma- and omega-gliadins and the high molecular weight (HMW) subunits of glutenin, Biochim. Biophys. Acta 1430 (1999) 359-366.
[33] N. Roos, A.J. Morgan, Cryo-preparation of thin biological specimens for electron microscopy, Oxford Science Publications, 1990.
[34] G. Porod, Theorie der diffusen Röntgenkleinwinkelstreuung an kolloiden Systemen, Z. Naturforsch. A (1949), 401.
[35] O.Glatter, O. Kratky, Small angle x-ray scattering, Academic Press, London, 1982.
[36] B. Hammouda, A new Guinier-Porod model, J. Appl. Crystallogr. 43 (2010) 716-719.
[37] D.D. Kasarda, J.E. Bernardin, W. Gaffield, Circular dichroism and optical rotatory dispersion of .alpha.-gliadin, Biochemistry 7 (1968) 3950-3957.
[38] Y.V. Wu, J.E. Cluskey, Optical rotatory dispersion studies on wheat gluten proteins: Gluten, glutenin, and gliadin in urea and hydrochloric acid solutions, Arch. Biochem. Biophys. 112 (1965) 32-36.
[39] F. Rasheed, W.R. Newson, T.S. Plivelic, R. Kuktaite, M.S. Hedenqvist, M. Gallstedt, E. Johansson, Structural architecture and solubility of native and modified gliadin and glutenin proteins: non-crystalline molecular and atomic organization, RSC Adv. 4 (2014) 2051-2060.
[40] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577-2637.
[41] K.M. Lammers, R. Lu, J. Brownley, B. Lu, C. Gerard, K. Thomas, P. Rallabhandi, T. Shea-Donohue, A. Tamiz, S. Alkan, S. Netzel-Arnett, T. Antalis, S.N. Vogel, A. Fasano, Gliadin induces

an increase in intestinal permeability and zonulin release by binding to the chemokine receptor CXCR3, Gastroenterology 135 (2008) 194-204.