

RESEARCH ARTICLE

Viscoelastic properties of wheat gluten in a molecular dynamics study

Łukasz Mioduszeński¹, Marek Cieplak¹*

Institute of Physics, Polish Academy of Sciences, Warsaw, Poland

* mc@ifpan.edu.pl

Abstract

Wheat (*Triticum spp.*) gluten consists mainly of intrinsically disordered storage proteins (glutenins and gliadins) that can form megadalton-sized networks. These networks are responsible for the unique viscoelastic properties of wheat dough and affect the quality of bread. These properties have not yet been studied by molecular level simulations. Here, we use a newly developed α -C-based coarse-grained model to study ~ 4000 -residue systems. The corresponding time-dependent properties are studied through shear and axial deformations. We measure the response force to the deformation, the number of entanglements and cavities, the mobility of residues, the number of the inter-chain bonds, etc. Glutenins are shown to influence the mechanics of gluten much more than gliadins. Our simulations are consistent with the existing ideas about gluten elasticity and emphasize the role of entanglements and hydrogen bonding. We also demonstrate that the storage proteins in maize and rice lead to weaker elasticity which points to the unique properties of wheat gluten.

OPEN ACCESS

Citation: Mioduszeński Ł, Cieplak M (2021) Viscoelastic properties of wheat gluten in a molecular dynamics study. PLoS Comput Biol 17(3): e1008840. <https://doi.org/10.1371/journal.pcbi.1008840>

Editor: Tamar Schlick, New York University, UNITED STATES

Received: August 3, 2020

Accepted: February 28, 2021

Published: March 24, 2021

Copyright: © 2021 Mioduszeński, Cieplak. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files. All steps required to reproduce the results are described in the Manuscript and the [S1 Text](#). The simulation program together with all the input files needed to run the simulations are included as a supplementary archive "[S1_DSB_Model.tar.gz](#)". The raw simulation data is included in the supplementary Excel file "[S1_Data.xlsx](#)".

Funding: This research has been supported by the National Science Centre (NCN, www.ncn.gov.pl),

Author summary

During the breadmaking process, expanding gas bubbles cause the dough to increase volume. Gluten proteins act as an elastic scaffold in that process, allowing the wheat dough to grow more than other kinds of dough. Thus, explaining the unique viscoelastic properties of gluten at the molecular level may be of great interest to the baking industry. Assessing the structural properties of gluten is difficult because its proteins are disordered. We provide the first molecular dynamics model of gluten elasticity, that is able to distinguish gluten and proteins from different plants, like maize and rice. Our model shows the structural changes the gluten proteins undergo during their deformation, which mimics the mixing of dough during kneading. It also allows for a determination of the force required to extend gluten proteins, as during baking. The data confirms existing theories about gluten, but it also provides molecular-level information about the extraordinary elasticity of gluten.

Introduction

When a dough from the wheat flour is gently washed in water, the cohesive mass that remains is gluten. Seed storage proteins constitute over 80% of this mass [1]. They exhibit viscoelastic

Poland, under grant No. 2018/31/B/NZ1/00047 (M.C.) and the European H2020 FETOPEN-RIA-2019-01 grant PathoGelTrap (<https://cordis.europa.eu/project/id/899616/pl>) No. 899616 (M.C.). The computer resources were supported by the PL-GRID infrastructure (plgrid.pl). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

properties that are vital for the properties and quality of bread [2]. Here, we neglect other substances present in real gluten samples and conduct molecular simulations of gluten proteins only, trying to recreate and study the polymeric network they form in gluten.

Gluten proteins have sequences with repetitive low complexity fragments that are rich in glutamines and prolines [3]. As a result, they do not possess any preferred tertiary structure [4–6] and are thus (mostly) intrinsically disordered (IDPs).

They can be divided into two fractions: glutenins (which can form large complexes, bound covalently through disulfide bonds) and gliadins (monomeric proteins that constitute around 70% of the total gluten protein mass) [1, 3].

On average, glutenin chains are longer [3] (see also Table 1), and their complexes may be stabilized also by entanglements [7]. Glutenins are thought to be the key ingredients of gluten elasticity [8], which is crucial in the breadmaking process [2]. Gliadins can be considered as a viscous plasticizer of the polymer network made by glutenins [7].

The observed viscoelastic behavior of gluten implies the existence of properties that are both solid-like (an elastic response to deformation) and liquid-like such as the viscous drag that arises from irreversible deformations. These properties can be characterized by the dynamic Young modulus G^* . It is one of the few quantities that can be used to directly compare simulations to experiment. We propose to use a one-bead-per-residue coarse-grained Dynamic Structure-Based (DSB) model [9, 10] that is a generalization of the structure-based (or Go-like) models to the case of the intrinsically disordered proteins. In the DSB model, the contacts between residues are determined from the instantaneous shape of the backbone and not from any reference structure. As a result, the contacts acquire a dynamic character—they can be born and then disappear, making their list fluent.

The DSB model captures the most important features of gluten and agrees qualitatively with available data. The main limitation of the model is the usage of time scales that are substantially shorter than those used in the experiments. Therefore, we do not expect to match the actual experimental values, especially if they are frequency dependent. Our protein model enables millisecond-scale simulations of biochemically large protein systems and incorporates the amino acid specificity. Our molecular model allows us to determine the relative importance of different factors that affect the properties of gluten. This work is especially important because gluten proteins are insoluble and disordered [11] so one cannot infer their properties from any structural features. The only simulations of gluten proteins we could find considered merely single chains instead of many protein complexes [12, 13]. However, the crowding conditions that are present in gluten are expected to affect the single-chain behavior. We study systems of ~ 4000 residues consisting of about 14 chains. This allows us to check the validity of existing theories of gluten structure on a molecular level and illustrate them in simulations. These theories have identified three main factors that contribute to gluten's elasticity:

Table 1. The mean number of cysteines per chain $\langle n_{\text{Cys}} \rangle$, mean number of residues in a chain $\langle N \rangle$, total number of residues ΣN and the mean coordination number z , for all of the systems studied here. The last property was determined based on the simulations, the rest by the sequential make-up of the systems.

system	$\langle n_{\text{Cys}} \rangle$	$\langle N \rangle$	ΣN	z
maize	6.1	221	3504	5.5 ± 0.1
gliadins	6	290	4322	6.88 ± 0.03
glutenins	7.2	445	3985	6.55 ± 0.08
gluten	6.5	384	4271	6.65 ± 0.07
rice	3.6	208	3503	5.62 ± 0.09

<https://doi.org/10.1371/journal.pcbi.1008840.t001>

1. The formation of hydrogen and disulfide bonds: in every gluten protein over 30% of amino acids is glutamine, whose side chain can be both a donor and an acceptor of hydrogen bonds [2]. Disulfide bonds can hold gluten proteins together and generate a sort of a polymer gel [8];
2. The loops-and-trains model [4, 14] explains gluten elasticity in terms of “loops” that are created by neighboring polymers: they stay together because of hydrogen bonding, but when some bonds are broken (e.g. by water), a free space emerges between the polymers. Under stretching, cavities disappear and hydrogen bonds are made between elongated and nearly parallel chain fragments, forming “trains” that are formed along the direction of stretching and cause resistance to further stretching;
3. The entangled nature of the polymer network itself leads to a strain resistance, regardless of its chemical nature [7].

We can study these phenomena in our molecular dynamics model by measuring the number of inter-residue contacts described by the Lennard-Jones (L-J) potential, the number of disulfide bonds (that can form and rupture dynamically), the number and size of cavities (as determined by using the Spaceball algorithm [15, 16]) and the number of entanglements (using the Z1 algorithm [17–20]). Structural properties, like the typical distance between the ends of a protein chain, can also be evaluated. We have also computed the mechanical work required to stretch gluten.

We consider gluten and the systems of its different fractions separately in order to elucidate the differences between these proteins. In addition, we also consider a control sample of proteins that do not exhibit such extraordinary viscoelastic properties [21]. Specifically, we use proteins derived from maize and rice for this task. Thus, altogether, we study 5 different systems. We show that many of the properties of gluten are substantially affected by “kneading” which we immitate by introducing periodic deformations.

Methods

Our molecular dynamics model is described in details in ref. [9]. The description, however, has not dealt with the formation of the disulfide bonds that are important for gluten.

Briefly, we use periodic boundary conditions in two directions (X and Y) and solid walls in the third direction, denoted as Z . The starting conformations of the proteins are constructed by generating self-avoiding random walks. They are then evolved by the methods of the coarse-grained molecular dynamics that were shown to correctly recreate properties of a set of intrinsically disordered proteins [9]. Rheological information is then obtained by deforming the box containing gluten proteins and recording the resulting force of response (see Subsection, “Simulation protocol”).

We simulate gluten as a set of protein chains in an implicit solvent. Each amino acid is represented by one pseudoatom and subsequent residues in a chain are connected by a harmonic potential. Chain stiffness is introduced through bond angle and dihedral angle potentials as obtained from a random coil library [9, 22]. These angle potentials have different forms based on the identity of residues constituting a given bond or dihedral angle: proline and glycine are treated as special cases. These particular amino acids are abundant in gluten [3].

Electrostatics are governed by the Debye-Hückel potential with relative permittivity that is dependent on the distance [23]: $\kappa = 40 \text{ nm}^{-1} r$. Other nonlocal interactions are modeled by the 6-12 L-J potential, which can be attractive (in its full form) or repulsive (truncated at the minimum and shifted). Contacts between residues form and disappear dynamically based on different criteria. The criteria have been developed by studying atomic-level overlaps that occur in

contacts arising in structured proteins. We distinguish between contacts arising from side-chain-sidechain (ss), sidechain-backbone (bs) and backbone-backbone (bb) interactions. However, the resulting depth of the potentials well, ϵ , is common, no matter what is its origin. Contacts involving sidechains are amino-acid specific. The specificity is introduced primarily through the locations of the potential minima. Some gluten proteins are thought to be partially structured, so a part of the contacts are always attractive.

The DSB coarse-grained model

The Dynamic Structure-Based model has been introduced in ref. [9] and here we just give a brief summary. We label the interacting pseudoatoms by indices i and j . Their positions are \vec{r}_i and \vec{r}_j , and $r_{ij} = |\vec{r}_i - \vec{r}_j|$ is the distance between them. The mass of each residue is set to the average amino acid mass m . The positions evolve according to the Langevin equations of motion with the time unit, τ , of ≈ 1 ns, the damping coefficient $\gamma = 2m/\tau$ (corresponding to overdamped dynamics [24]), and thermal white noise $\vec{\Gamma}_i$ with variance $\sigma^2 = 2\gamma k_B T$:

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i - \gamma \frac{d\vec{r}_i}{dt} + \vec{\Gamma}_i \tag{1}$$

We solve the equations of motion by using the 5th order predictor-corrector algorithm [25]. Energy unit ϵ is of order 1.5 kcal/mol [26, 27]. The room temperature is then about $0.3-0.35 \epsilon/k_B$, where k_B is the Boltzmann constant. Here, the simulations are run at $0.3 \epsilon/k_B$. The results are presented, approximately, in the metric units.

The excluded volume of the residues (whether involved in a contact or not) is ensured by the L-J potential that is cut at $r_o = 0.5$ nm (so that $V_r(r_{ij} \geq r_o) = 0$):

$$V_r(r_{ij} \leq r_o) = \epsilon \left(4 \left(\frac{\sigma_0}{r_{ij}} \right)^{12} - 4 \left(\frac{\sigma_0}{r_{ij}} \right)^6 + 1 \right) \tag{2}$$

Where $\sigma_0 = r_o \cdot (0.5)^{\frac{1}{6}}$, so that $V_r(r_o) = 0$. $i, j = i+2$ interactions are always described by this repulsive potential. Connected $i, j = i+1$ beads interact via harmonic potential with the elastic constant of $k = 5000 \text{ nm}^{-2} \cdot \epsilon$ and minimum at $r_b = 0.38$ nm (consistent with other works [27, 28]).

The distinction between the bb, bs and ss contacts in our one-bead-per-residue model is accomplished in the following way. When two beads come sufficiently close to each other, we compute two vectors based on positions of three consecutive beads ($i-1, i, i+1$): $-\vec{n}_i$ is the negative normal vector, approximately pointing in the sidechain direction, and \vec{h}_i is the binormal vector that points in the direction of a possible backbone hydrogen bond. The relative directions of these vectors are checked, and if, for instance, \vec{h}_i and \vec{h}_j point towards each other, a bb contact is made. Each type of amino acid has a specific number of contacts it can make with its virtual “backbone” and “sidechain”. Each ss contact making a pair of amino acids corresponds to different distances at the minimum of the L-J potential. On the other hand, that minimum is 0.5 nm for the bb contacts and 0.68 nm for the bs contacts. Independent of the source of the contact, the interaction is modeled by the potential of the same form:

$$V_{attr}(r_{ij}) = 4\epsilon \left(\left(\frac{\sigma_{attr}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{attr}}{r_{ij}} \right)^6 \right) \tag{3}$$

where σ_{attr} depends on the type of interaction and specificity. However, if the contact is due to the bb interactions, its strength is doubled. Doubling the depth of the L-J potential for the bb

contacts is consistent with the earlier literature findings [29]. If criteria for creation of a given type of contact are met, the attractive part of the potential is turned on at r_{min} in a quasi-adiabatic fashion (during 10τ). When the beads are in a distance larger than $r_f = f \sigma_{attr}$ from each other, the potential is turned off in the same way. This is why we call those contacts “dynamic” as they can appear and disappear during the simulation. In references [9, 30] and in this paper, we have used $f = 1.5$ (in an analogy to our description of relevant contacts in the structured proteins [26]). This model was previously used to study polyglutamine aggregation [30], so it should be well-suited for studying glutamine-rich gluten proteins. However, our more recent tests indicate that the choice $f = 1.3$ is better. In particular, it is more consistent with the ANTON-based simulations of α -synuclein [31].

Disulfide bonds. For modeling disulfide bonds, we considered using harmonic potential that is switched on and off, following Thirumalai et al [32]. However, we have found out that it is equivalent but much simpler to use the L-J potential with the strength of 4ϵ and $r_{min}^{ss} = 0.59$ nm.

The disulfide bond potential is turned on and off in the same manner as other ss contacts, i.e. adiabatically within approximately 10τ . It should be noted that the real timescales of the disulfide bond formation and their duration are much longer [33]. The only modification besides the amplitude change is that after two cysteines form such a contact, they cannot form additional disulfide bridges (without a prior rupture of the existing bridge).

Structured parts. Some of the gluten proteins contain domains near their termini that tend to be more structured than the whole chain. For example, homology of the A domain of high molecular weight glutenins to a plant inhibitor of digestive enzymes [34] suggests that one or two pairs of cysteines in that domain form fixed intra-chain disulfide bridges. We incorporated available information about possible structured parts of gluten by using a structure-based Go model for those regions [26]. The identification of segments that are recognized as structured domains was based on the UniProt [35] partition of gluten proteins into domains (see Section 1 of the S1 Text for details). For the backbone stiffness of the structured segments we used the same potential as for the unstructured parts.

The actual structures in these segments were predicted by using the ITASSER server [36, 37]. The contact maps were generated based on those all-atom structures, using the overlap criterion (spheres corresponding to heavy atoms of the residues in contact must overlap [26]). The contacts present in the contact map in the structured regions are always attractive (they are “static”, in contrast with the “dynamic” contacts operating in the disordered parts) and the L-J potential describing them has a minimum that corresponds to the distance between the C_α atoms in the reference all-atom structure.

Residues listed in the “static” contact map can still make “disordered” contacts with the residues outside the contact map. The number of “static” contacts made by a residue is determined by the same $r < r_f$ criterion, but the potential does not switch off if $r > r_f$.

Simulation protocol

We performed simulations of 5 different systems specified in Table 1. The exact protein composition of all simulated systems is presented and described in detail in Section 1 of the S1 Text.

While cereal storage proteins are still in the grain, they are usually located in the protein bodies or protein storage vacuoles [38]. After grain maturation, these deposits can fuse and form a continuous network [1]. This is reflected in our simulation protocol: in the beginning, the proteins are in a low-density box and have time to attain their shapes separately [39]. Then, as the box dimensions get smaller, different chains connect together, forming one big

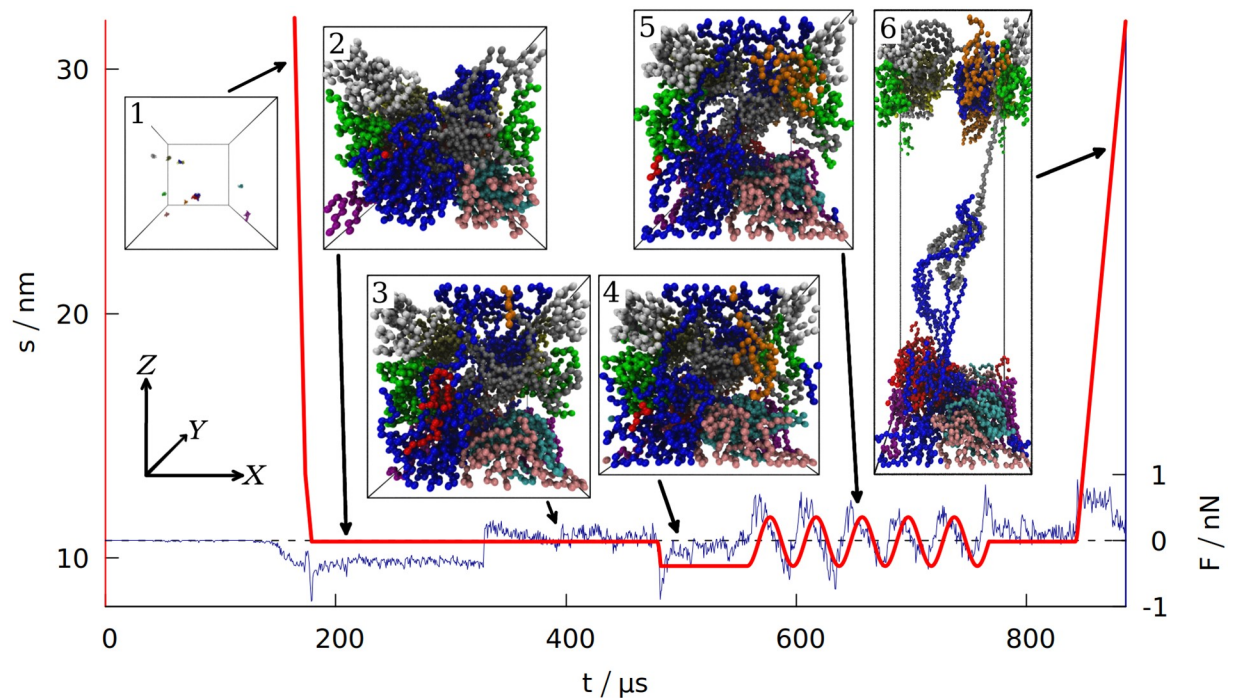


Fig 1. Time dependence of the distance s separating the two opposite box walls in the Z direction. The blue curve shows the force exerted on the system by those two walls (summed over all the residues and averaged over 100 ns time interval). The data is for gluten (defined in Section 1 of the [S1 Text](#)), with the density of 3.5 nm^{-3} . The period of oscillations is 40 ns. The six insets correspond to the snapshots of the system taken at different stages of the simulation (each bead represents one amino acid, the protein chains are shown in different colors). The same coordinate system (top left) is used for all 6 insets.

<https://doi.org/10.1371/journal.pcbi.1008840.g001>

cluster. Unfortunately we cannot reproduce drying and hydration in an implicit-solvent model, but the final density of the system reflects the density of real gluten samples [40]. Mixing and kneading of the flour and the resulting dough is represented by the periodic deformations of the box, that can also be compared to the experimental deformations of (much larger) gluten samples in a rheometer [41].

The first step of the simulations is generating the chains as self-avoiding-random-walks positioned randomly in a box with density $\rho = 0.1 \text{ res/nm}^3$ (residues per cubic nanometer). If an initial random walk configuration happens to cross the boundary of the box, the box is enlarged to keep all chains inside, resulting in a slightly lower initial density. The density of 0.1 res/nm^3 is more than 30 times lower than the density of gluten, which should ensure that proteins are in monomeric or dimeric form during the initial equilibration (see inset 1 in [Fig 1](#)). Boundary conditions along the X and Y axes are periodic, whereas the walls along the Z direction are repulsive. They are separated by the distance of s . The residues are repelled from the wall by the repulsive part of the L-J potential with the depth 4ϵ .

After equilibrating the system for $125 \mu\text{s}$ the box dimensions are reduced from the initial size with the speed of 2 mm/s, after the density ρ_0 (usually corresponding to the real density of gluten [40] $\rho_0 \approx 3.5 \text{ res/nm}^3$) is reached (see Section 3 of the [S1 Text](#) for more information about the density). At this stage, the distance between the repulsive walls has a value that is denoted by s_0 . The system is then equilibrated for another $150 \mu\text{s}$.

After the second equilibration stage (depicted in inset 2 in [Fig 1](#)) the walls along the Z direction start to attract the proteins. When the distance between any residue and the wall gets closer than $d_{\min} = 0.5 \text{ nm}$, the attractive part of the L-J potential with depth of 4ϵ is turned on

adiabatically (in the same manner as for disulfide bonds). Then the interaction between the wall and a residue attached to it is described by $V_{wall} = 4V^{LJ}(r_{i,w})$, where $r_{i,w}$ is the distance between the i th pseudoatom and an i th interaction center on the wall. The X and Y coordinates of the i th interaction center are fixed at the values of the first bead attachment (when it came to the wall closer than $d_{min} = 0.5$ nm, which is the distance corresponding to the minimum of the V_{wall} potential). The Z coordinate of an interaction center is 0 on one rigid wall and s on the opposite wall. If a wall moves so that $s \neq s_0$, the interaction centers move together with the wall. If the bead departs away from the interaction center by more than 2 nm, it is considered detached and the interaction center disappears. The bead may reattach at another place later. For $d < d_{min}$ the repulsive part of the L-J potential is always turned on to ensure that no residue can pass through the wall. The force acting on the walls is a sum of forces that individual residues exert on both walls. It is averaged over 100 ns interval to filter the thermal noise out. An alternative way to introduce interactions with the walls (involving the fcc lattice of pseudoatoms) is discussed in Section 4 of the [S1 Text](#).

After the wall attraction is enabled, we give another 150 μ s for the equilibration of the system (inset 3 in [Fig 1](#)). After this third equilibration stage, when the attraction is already turned on, the walls in the Z direction are moved to the distance of $s = s_0 - A$, where A is the amplitude of the oscillations in the normal direction (we set $A = 1$ nm). After being moved to the minimum (shown in inset 4 in [Fig 1](#)), the wall position changes periodically: $s(t) = s_0 - A \cos(\omega t)$. The wall at $Z = 0$ stays fixed. In the experiments, $\omega \sim 1$ Hz [2], which is not possible to achieve in our simulations. Our default period is 40 μ s, corresponding to $f \approx 25$ kHz. The maximum displacement is shown in inset 5 in [Fig 1](#). For shearing simulations, the walls are also displaced with the same amplitude, A , but the oscillations take place in the X direction and are described by $s'(t) = -A \cos(\omega t)$ (where $s' = 0$ corresponds to one rigid wall being directly above the other). See [S1 Fig](#) for an illustration.

After oscillations in the normal direction (or after additional 100 μ s of equilibration if no oscillations were made) the system is equilibrated again for 75 μ s and then the walls in the Z direction extend gradually to $2s_0$ so that the total work required for such an extension can be measured, as well as the distance and force required to break the system apart (recreating conditions for measuring uniaxial extension [41]). Inset 6 in [Fig 1](#) depicts this elongation. The speed of elongation is 0.5 mm/s, which is a typical pulling speed employed in AFM stretching experiments in coarse-grained simulations [26, 42]. The maximum speed during oscillations ($v_{max} = A\omega$) is even lower.

Results

Properties of gluten, maize and rice

Based on refs. [3, 8, 11, 43–49] we chose proteins that were representative for each of the five systems we study: gliadins, glutenins, gluten and storage proteins from maize and rice. We then used the UniProt database [35] to select the corresponding sequences. The resulting compositions are described in Section 1 of the [S1 Text](#). [Table 1](#) summarizes the key properties of each of the systems studied. A large number of cysteines per chain should increase the propensity to form inter-chain disulfide bonds. On the other hand, longer chains should make entanglements easier. Finally, the “stickiness” of the system is reflected in the coordination number z , which is defined as the mean number of contacts made by each residue (both “dynamic” and “static” combined).

We performed three types of simulations: a) without any oscillations; b) with oscillations with deformation in the normal (Z) direction; c) with oscillations with deformation in the shear (X) direction. All of the properties are calculated in the stage after oscillations (in the

absence of oscillations—after the equivalent time is passed). The parameter z does not seem to be affected by the oscillations.

However, the values of 5 other quantities depend on the prior history of the system. Those quantities are: the number of entanglements l_k , the maximum force F_{\max} , the mechanical work W_{\max} required to stretch the system, the number of the interchain contacts n_{inter} and the RMSF (root mean square fluctuation) as averaged over all residues and determined as a deviation from a mean location of the residue.

It should be noted that each of the systems studied is simulated in a somewhat different box (to achieve the same density for different numbers of residues) and thus comparison of the absolute values of these quantities is not very meaningful. Instead, we analyze the ratios of the five quantities to their average values in the absence of oscillations (denoted by the tildas). Fig 2 shows these ratios for the case of the shearing oscillations whereas S3 Fig for the uniaxial oscillations.

The first of the quantities studied is the total number of entanglements. We have determined l_k by using the Z1 algorithm [17]. An entanglement occurs if a path connecting the first and last residue in a chain cannot be shortened to a straight line without crossing a path from another chain [18, 19]. After minimizing the lengths of all such paths with the constraint that they cannot cross one another, each kink, where two paths intertwine, counts as an entanglement [20]. An example of an entanglement is shown in inset 6 in Fig 1, where the end of the grey chain (a low molecular weight glutenin) crosses a loop made by a blue chain (a high molecular weight glutenin). Due to the low number of the chains, each snapshot of the simulation usually has only a few entanglements. We averaged l_k over all snapshots made after oscillations.

We use a deterministic random number generator, which produces the same output when provided with same initial random number, called, “random seed” [50]). The simulations for each system were repeated at least three times with a different random seed. The same set of random seeds was used for the shearing, pulling and equilibrium simulations. The result shown in Fig 2 uses the mean values from all simulations, but the error of the mean is quite high. Nevertheless, a clear increase is observed for gluten: the oscillating deformations cause more entanglements. We did not include self-entanglements (knots) in Fig 2, but they are also present in gluten: some long proteins can form very deep knots (see S5 Fig), that can last for over 100 μs .

The differences between the studied systems are most pronounced in the case of F_{\max} , the maximum force required to stretch the system in the last stage of the simulation (illustrated in inset 6 in Fig 1). The blue curve in Fig 1 shows how the force rises and then drops near the end due to the rapture. However, different random seeds may lead to different force curves for the same system. Fig 3 shows three stretching force curves for three different simulations of gluten. Each curve has a number of force peaks. In order to determine the characteristic largest force, F_{\max} , we take 5 largest force points and calculate their mean value. These mean F_{\max} are also indicated in Fig 3 as dotted lines surrounded by strips that represent their standard deviation. The F_{\max} used in Fig 2 is a weighted average of those values (from simulations with different random seeds), where the weights are calculated from the standard deviation of the mean. Large glutenins experience the largest, 5-fold increase in F_{\max} after oscillations, while for gliadins there is even a very slight decrease. The rise in F_{\max} for gluten is substantially bigger than for corn and rice proteins.

The last stage of the simulations includes moving the Z walls apart with a constant speed. The response force may be easily integrated over distance, giving the work required to stretch the system. The maximum work W_{\max} required to stretch the system is also bigger if oscillations were introduced before. Again, the difference is largest for glutenins, this time followed

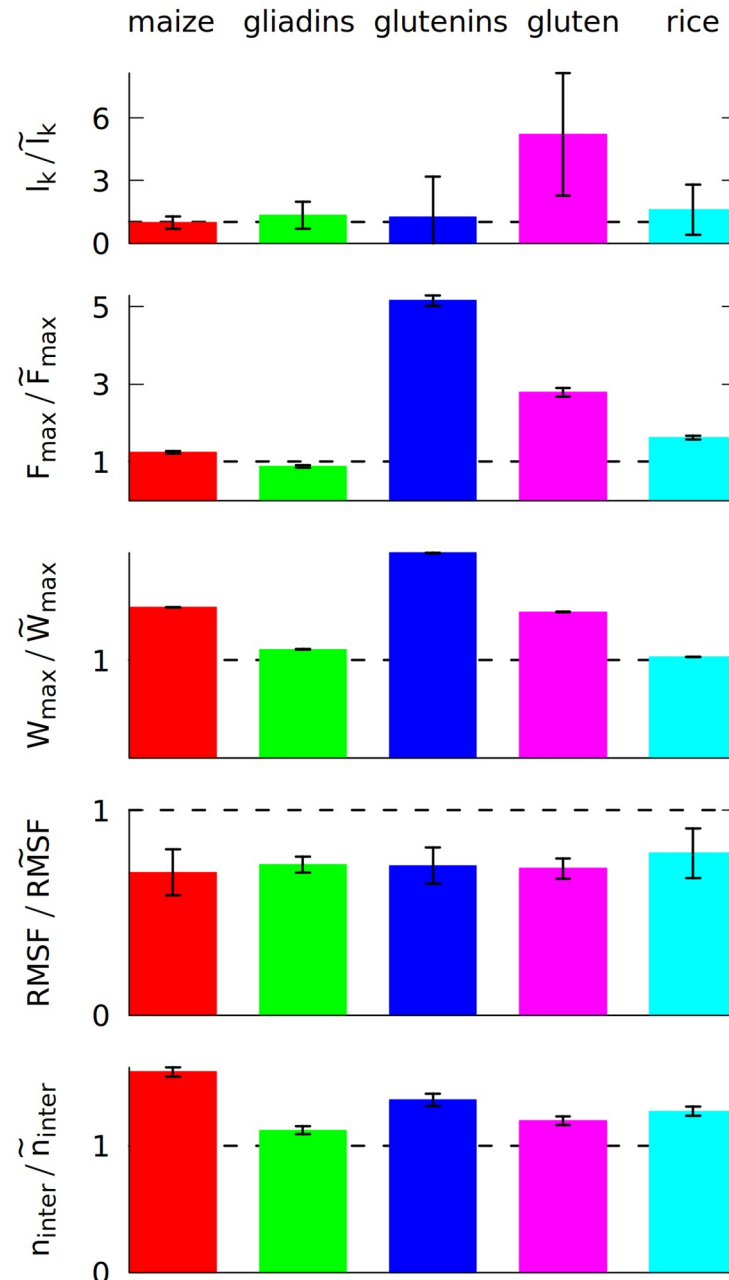


Fig 2. The ratio of 5 different properties calculated after 5 shearing oscillations (in the nominator), and without any oscillations (in the denominator, marked by ~). The properties are the average number of entanglements l_k , the maximum force F_{\max} during the elongation in the last simulation stage and the maximum work W_{\max} required to elongate the system in that stage, the number of inter-chain contacts n_{inter} and RMSF (root mean square fluctuation) averaged over all residues. The ratio of 1 is marked by a broken line. Each color corresponds to a different system, as displayed at the top, and defined in Section 1 of the [S1 Text](#). The system density is 3.5 nm^{-3} .

<https://doi.org/10.1371/journal.pcbi.1008840.g002>

by maize proteins. Sometimes the force drops to near zero after reaching the maximum (meaning that the system was ruptured) and the work stops rising (as seen for the dark blue curves on [Fig 3](#)). There are other cases where a substantial mechanical work is still required for elongating the system after reaching F_{\max} (cyan curve), but the total mechanical work W_{\max} is correlated with the height of the F_{\max} peak. The proteins apply some effective pressure on the

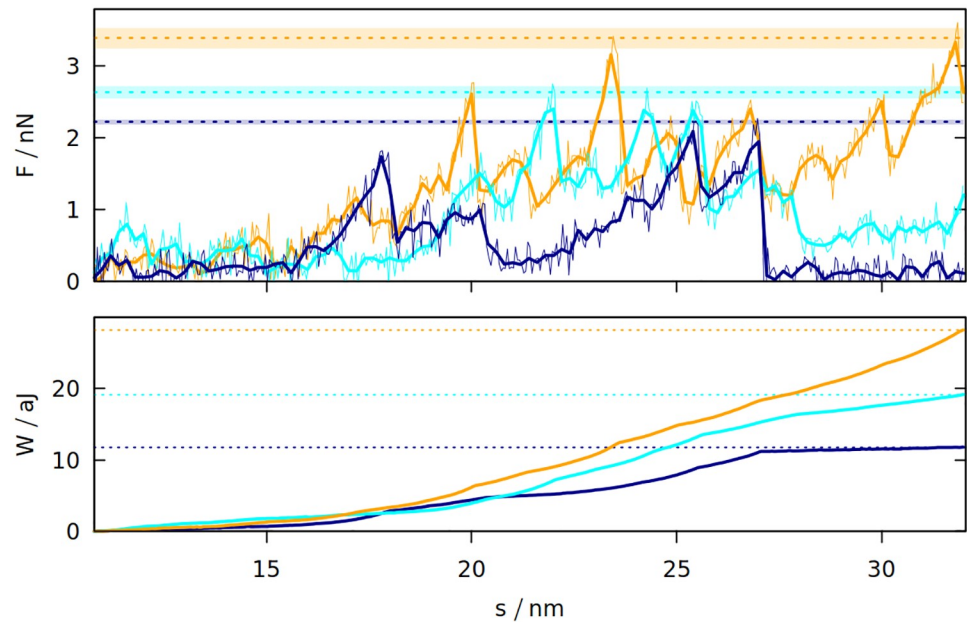


Fig 3. The force exerted on gluten with the density of 3.5 nm^{-3} (top panel) during elongation in the last stage of the simulation and the resulting work (bottom panel) as a function of the distance between the rigid walls. The three trajectories are shown in different colors. The dotted lines correspond to F_{max} and W_{max} for a given trajectory. For F_{max} the strips around the dotted lines represent the uncertainty resulting from averaging F_{max} over 5 points with the biggest force. The oscillation period is $40 \mu\text{s}$. For F , the thin lines represent raw simulation data and the thick lines are smoothed.

<https://doi.org/10.1371/journal.pcbi.1008840.g003>

walls, however it is much smaller than the mechanical stress resulting from the stretching (the atmospheric pressure would exert 0.01 nN of force on the walls).

Even if entanglements cause sharp peaks in the force curves, the resistance against deformation is also controlled by the number of inter-chain contacts, n_{inter} , that have to be ruptured during elongation. The biggest increase in n_{inter} is observed for maize proteins. They have a lower number of contacts in total (as indicated by their coordination number z shown in Table 1), but their residue composition allows them to make more contacts due to the mixing effect of the oscillating deformations.

The oscillatory mixing effect is general and it operates in each of the systems studied. It results in a stronger network of proteins that are connected together more tightly, which results in the lowering of RMSF. Less mobile proteins may be more sturdy, but, as a trade-off, the maximum force may occur for a lower inter-wall distance s [51].

The molecular explanation of gluten elasticity

According to the “loops and trains” picture, the response of gluten to deformation consists of two stages: at the first stage the chains elongate and the cavities between them shrink, which requires smaller force than disentangling the chains and moving them alongside each other that takes place in the second stage [14, 52]. We tested these predictions by measuring how 6 different quantities change during the final extension of the simulation box (see Fig 4). The snapshots shown in this figure indicate that the “trains” do form, and that shearing between the chains drawn in red and blue may lead to a substantial force by disrupting and reforming the contacts between the chains. Indeed the number of inter-chain contacts decreases in the beginning of the elongation, but then it increases again as “trains” are formed. The chains

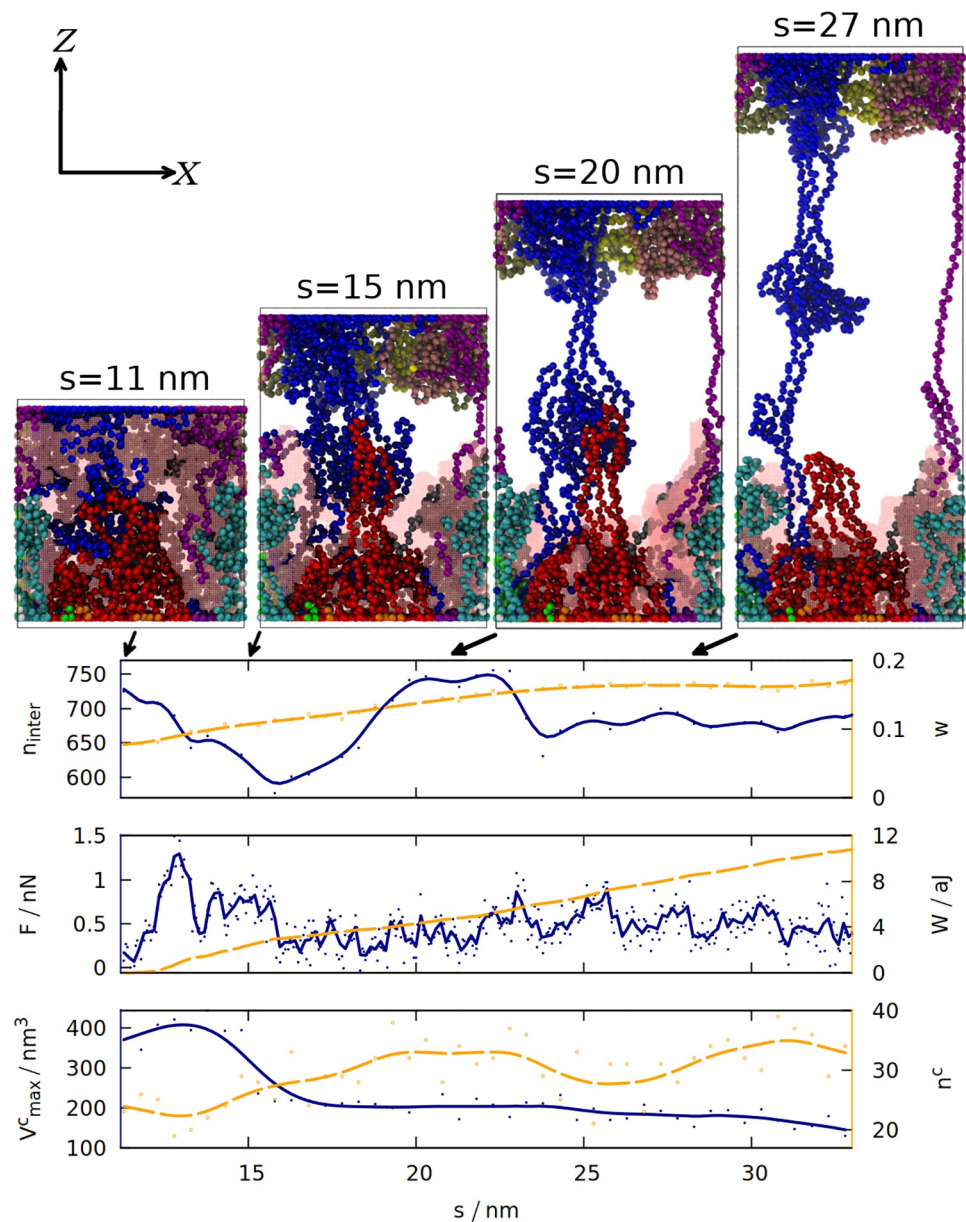


Fig 4. The time dependence of 6 different quantities that are measured for gluten (as defined in section 1 of the S1 Text) during the elongation in the last stage of the simulation. The solid blue (orange) lines correspond to the quantities listed along the left (right) y-axis. The lines are smoothed, solid (hollow) dots represent raw data from simulation. The quantities analyzed are: the number of inter-chain contacts n_{inter} and the average distortion parameter w , force F and work W , the maximum volume of a cavity V_{max}^c and the number of cavities n^c . The 4 snapshots (similar to those on Fig 1, but in the orthographic projection) correspond to the displacements s written above them and marked by the arrows. The same coordinate system (defined in the top left) is used for all snapshots. The biggest cavity is outlined by a pink mesh.

<https://doi.org/10.1371/journal.pcbi.1008840.g004>

become more rod-like and less globular, as indicated by the distortion parameter w [53] (averaged over all chains) that is equal to 0 for an ideal sphere, and 1 for an ideal rod. It is defined as $\frac{\Delta R}{\bar{R}}$, where $\bar{R} = \frac{1}{2}(R_1 + R_3)$ and $\Delta R = R_2 - \bar{R}$; R_1 is the smallest radius of inertia, R_3 —the largest and R_2 the intermediate.

Formation of a “train” (two parallel chain fragments connected by hydrogen bonds) does not coincide with the maximum of the force, because F_{\max} occurs near the beginning of the elongation and is associated with the biggest decrease in the number of contacts. Nevertheless, the force does not drop to zero and the work required to stretch the system steadily increases, because the chains from the opposite ends of the box are still connected and cause further resistance against deformation.

In order to quantify the number and size of cavities that form in gluten, we used the Spaceball algorithm [15, 16], which uses a probe with a user-defined radius to determine a three-dimensional contour of the system. The volume of the cavities is computed by filling a grid with those probes and rotating the grid with respect to the system. Our system is quite large, so we used grid with 0.1 nm lattice constant, a probe with radius 0.38 nm and one rotation of the grid. The snapshots generated show that the biggest detected cavity is dispersed over the whole simulation box, but during the elongation it becomes associated with the set of proteins attached to the $Z = 0$ wall (for the trajectory shown). As they are more and more elongated, the size of the cavity decreases. Larger cavities become divided into many smaller cavities, which is reflected in the increase in n_c .

Fig 4 illustrates just one simulation (the same properties for the same system, but with a different random seed, are shown in S6 Fig), but the volume of the biggest cavity indeed decreases after elongation in most cases (see S4 Fig). The chains become more rod-like not only during the final stretching, but also after the oscillations (see S2 Fig): deformations cause the proteins to be briefly stretched, but after the deformation ends, the proteins do not return to the exactly the same shape as before, but stay elongated in one direction (which may be called a memory effect).

Determining the dynamic shear modulus

One of the few fundamental (independent of the size and shape of the sample) rheological properties of gluten is the dynamic shear modulus $G^* = G' + iG''$, measured by deforming the gluten sample with sinusoidal deformation (shear strain for G^* and normal strain for E^*) [41]. The real part (G' , elastic modulus) represents the in-phase part of stress response, the imaginary part (G'' , loss modulus) the out-of-phase response (phase difference δ is defined as $\tan \delta = \frac{G''}{G'}$). This is a non-destructive method, because the amplitude of oscillations is small (usually up to a few percent [2]). The total force $F(t)$ that the system under a shear deformation exerts on the walls in the X direction (see S1 Fig), divided by the surface area of a wall S , defines shear stress $\phi(t) = F_X(t)/S$ (the usual symbol for shear stress, τ , is a time unit in our model, so we use ϕ instead). A strain defined as $\gamma = \gamma_0 \cos(\omega t)$ is expected to induce a stress in a similar form, shifted in phase by δ : $\phi(t) = \phi_0 \cos(\omega t + \delta)$. Dynamic shear modulus $G^* = G' + iG''$ is then defined as [54]:

$$G' = \frac{\phi_0}{\gamma_0} \cos \delta \quad G'' = \frac{\phi_0}{\gamma_0} \sin \delta \quad (4)$$

where ϕ_0 and γ_0 are the amplitudes of stress and strain accordingly. We define it by fitting to functions $s'(t) = A \cos(\omega t)$ and $F(t) = F_0 \cos(\omega t + \delta)$. Here δ is the phase difference, $\gamma_0 = A/s_0$ and $\phi_0 = F_0/S$. This is a very basic method of determining the shear modulus and more advanced methods can be used [55]. However, this method can be directly compared to experiment, where a gluten sample is also put under strain that changes periodically: a rheometer consists of two plates that move with respect to each other, and the sample is placed between them [41]. The plates may be flat or have a different geometry: we recreate the first (“plate-plate”) case (the walls in the Z direction represent the moving plates). The movement pattern

Table 2. Shear modulus of dried gluten and its fractions, found by experiment with oscillation frequency $f = 1$ Hz [2], and by our simulations (marked by subscript_{sim}) with $f \approx 14$ kHz (corresponding to the longest oscillation period we used, $70 \mu\text{s}$).

System	G' [kPa]	G'' [kPa]	$\tan \delta$	G'_{sim} [MPa]	G''_{sim} [MPa]	$\tan \delta_{\text{sim}}$
Gluten	2.5	1.3	0.5	6.0 ± 1.6	6.1 ± 1.7	1.0 ± 0.1
Gliadins	0.5	0.6	1.2	5.6 ± 1.7	8.0 ± 2.3	1.4 ± 0.1
Glutenins	40	9	0.2	8.3 ± 2.2	7.9 ± 2.1	0.95 ± 0.1

<https://doi.org/10.1371/journal.pcbi.1008840.t002>

in both cases is the same and represents the sinusoidal strain. The relative amplitude of the oscillations is also similar [2]. The rheometer records the force required to move the plates [41], whereas in the simulations the forces are computed directly. Movement in the X direction is equivalent to a measurement in a shear rheometer (that measures G^*), whereas walls moving in the Z direction reflect experiments in an extensional rheometer (that is usually used for larger and more destructive deformations, like those described in the previous subsections [41]).

In Table 2 and Fig 5 we compare the results with experiment. Table 2 shows that gliadins are responsible for the viscous properties of gluten (big δ), and glutenins for the elastic part (small δ). Despite largely different values of the results, this trend is still visible in our simulations for the longest oscillation period ($70 \mu\text{s}$): $\tan \delta_{\text{sim}}$ is biggest for gliadins and smallest for glutenins.

We use frequencies that are 4 orders of magnitude larger than in experiment, but some extrapolations can be deduced from the experimental frequency spectra of gluten's G^* . We

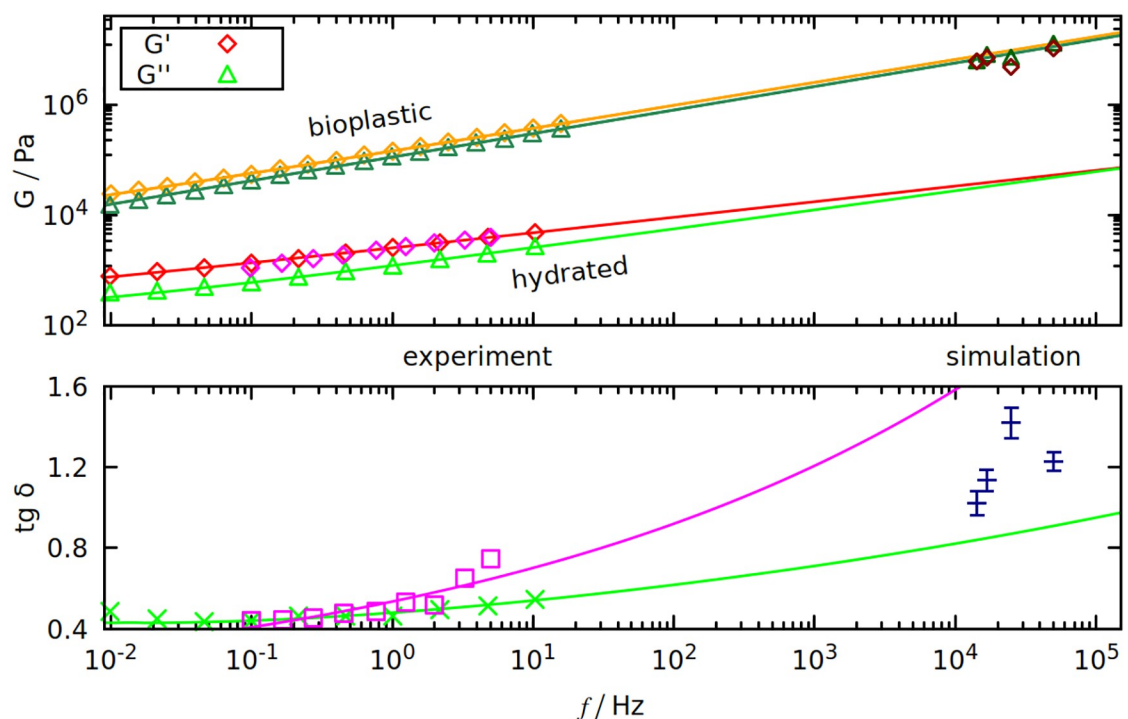


Fig 5. Dynamic shear modulus (top panel) and $\tan \delta$ (bottom panel) as a function of oscillation frequency f . Results are from our simulations (on the right) and from experiments (on the left). Purple points are from [57], red and light green from [56]. On the top panel, orange and dark green points are from [58]. All curves fitted to the data are of the form $a \cdot f^b + c$, where a , b and c are the fit parameters.

<https://doi.org/10.1371/journal.pcbi.1008840.g005>

used the results from two experimental papers [56, 57] that use hydrated gluten samples. Both predict that $\tan \delta$ should be rising with frequency—see the bottom panel of Fig 5, where $\tan \delta_{\text{sim}}$ is between the extrapolations predicted by [56] (green curve) and [57] (purple curve). Such large discrepancies are the effects of different gluten compositions, that greatly influence the results (we use the Aubaine dataset from [57]).

Growing δ means that G' and G'' converge, as seen in the top panel of Fig 5. Both G' and G'' are rising with f , however this rise is insufficient to explain why the values from simulation are 3 orders of magnitude larger. Either there is a significant non-linearity for higher frequencies and the extrapolation no longer holds, or this is a limitation of our model. We use implicit solvent which cannot account for differences in gluten's hydration that may greatly affect the results (even, “dry” gluten is slightly hydrated) [56].

On the other hand, the results of the simulation are in good agreement with the experiments done on gluten-based bioplastics (as shown on the top panel of Fig 5), where gluten proteins are mixed with glycerol and melted at over 380 K [58]. Even for bioplastics, the tensile strength (the stress needed to break the deformed system) is less than 1 MPa [58], while the F_{max} divided by the wall area (which is around 100 nm²) can be more than 10 times higher (see Fig 3).

Discussion and conclusions

We have used our coarse-grained model with the implicit solvent to provide microscopic-level understanding of gluten and the related proteins. Kneading and mixing gluten samples is known to increase their tensile strength [59]. In our simulation, we mimick these actions by introducing periodic deformations. We have found that as a result of these deformations gluten and glutenins became significantly more resistant to stretching, as indicated by the rise in F_{max} and W_{max} . Our model is also able to correctly distinguish gluten from storage proteins from other plants, and to tell the difference between durable glutenins and viscous gliadins. This difference appears to come from the combined effect of entanglements and inter-chain contacts (that represent mainly hydrogen bonds between glutamines in the case of gluten). Maize proteins formed more inter-chain contacts than rice proteins (which may justify why maize flour is more useful for making gluten-free doughs [60]). Its chains were too short to gain additional mechanical resistance from being entangled.

The importance of entanglements and inter-chain hydrogen bonding is already well-known in the experimental literature [7, 8], but our simulations provide an additional confirmation of that. We also elucidated microscopic mechanisms of the loops-and-trains picture [14]. We have found that the decrease in the volume of the biggest cavity is accompanied by the rise in the number of the inter-chain contacts during stretching.

Due to the small system size only a few inter-chain disulfide bonds were simultaneously present during our simulations (most of them were observed for maize proteins) and they did not seem to influence the results significantly. It is possible that they may become more important at larger length scales, where covalent linkage of the chains affects properties like solubility [3]. However, for about 20 gluten chains, the entanglements and non-covalent contacts are sufficient to explain elasticity of the system.

We used the periodic deformation to calculate the dynamic Young modulus G^* . We have observed that the theoretical frequency dependence shown in Fig 5 agrees with the experimental results for glycerol-plasticized gluten [58] better than with those for unrefined hydrated gluten [56, 57]. In our implicit solvent model glycerol may theoretically replace water as the solvent. However, our model was parameterized with the assumption of a watery solvent [9]. It should be noted, however, that the experimental range of the frequencies is orders of

magnitude smaller than implemented here. Further simplifications are needed to access the longer time scales.

Another reason for the very high G^* may be the coarse-graining itself: a coarse-grained pseudoatom has smaller volume than an all-atom amino acid. This may lead to different elastic behavior, as was observed previously in the case of virus elasticity [61].

Despite the deciding role of gluten in determining the dough quality, its measured G^* values are not strongly correlated with it [2]. During baking, the gas bubbles inside the dough extend it far beyond the limit of linear deformations. Therefore, measurements of extension to over 200% of the initial size are more informative [2]. Such extensions were generated in the last stage of our simulations and were discussed above. Unfortunately, properties like W_{\max} and F_{\max} depend on the system size and geometry, so they cannot be directly compared to experimental values.

Another hypothesis that can be checked in our simulation is that gluten proteins can form a β -spiral from consecutive β -turns made by repetitive fragments like QPGQ [13] (such fragments are common especially in glutenins [4]). Scanning tunneling microscopy discovered a groove with 1.5 nm periodicity in glutenins [4]. However, the sample preparation procedures used prevented the study from validating the existence of the β -spiral in bulk gluten. Models of single glutenin structures consisting of repetitive β [62] and γ [12] turns were reported in the literature, however we did not observe similar structures in our simulations.

We do not discuss here how gluten proteins interact with the human body, but this widely disputed subject [63, 64] can also be tackled by our model in the future. Simulations can also check what part of gluten proteins is solvent-accessible, and thus available for digestive enzymes in the stomach, which could shed new light on the gluten intolerance problem [65].

Supporting information

S1 Fig. Time dependence of the wall displacement s' in the X direction (top panel) and the distance s separating the two opposite box walls in the Z direction (bottom panel). The purple curve (top panel) shows the X component of the force exerted on the proteins by the walls (summed over all the residues and averaged over 100 ns time interval), the blue curve (bottom panel) shows the Z component of that force. The data is for gluten (defined in Section 1 of the S1 Text), with the density of 3.5 nm^{-3} . The period of oscillations is 40 ms. The numbers with arrows correspond to the stages of the simulation shown on Fig 1 in the main article. Snapshots 4 and 5 show the system during shearing oscillations (each bead represents one amino acid, the protein chains are shown in different colors). The coordinate system (top left) is the same as in Fig 1.

(TIF)

S2 Fig. The ratio of the average distortion parameter w [53] calculated after 5 shearing oscillations (in the nominator), and without any oscillations (in the denominator, marked by ~). Period of oscillations is $70 \mu\text{s}$ (lower periods did not result in such significant changes in w). The ratio of 1 is marked by a broken line. Each color corresponds to a different system, as displayed at the top, and defined in Section 1 of the S1 Text. The system density is 3.5 nm^{-3} .

(TIF)

S3 Fig. The ratio of 5 different properties calculated after 5 axial oscillations (in the nominator), and without any oscillations (in the denominator, marked by ~). The properties are the average number of entanglements l_k , the maximum force F_{\max} during the elongation in the last simulation stage and the maximum work W_{\max} required to elongate the system in that stage, the number of inter-chain contacts n_{inter} and RMSF (root mean square fluctuation)

averaged over all residues. The ratio of 1 is marked by a broken line. Each color corresponds to a different system, as displayed at the top, and defined in Section 1 of the [S1 Text](#). The system density is 3.5 nm^{-3} .

(TIF)

S4 Fig. A ratio showing a decrease in the volume of the biggest cavity during the last stage of the simulations (elongation). The V_{max}^c is the average taken over the first half of the elongation process, V_{max}^{c-1} is the average taken over the last half. The ratio of 1 is marked by a broken line. Each color corresponds to a different system, as displayed at the top, and defined in Section 1 of the [S1 Text](#). The system density is 3.5 nm^{-3} .

(TIF)

S5 Fig. An example of a knotted high molecular weight (HMW) glutenin (from a gluten simulation with $\rho = 4 \text{ nm}^{-3}$). Knotted area is shown in red, knot ends are indicated by black arrows.

(TIF)

S6 Fig. The time dependence of 6 different quantities (defined for [Fig 4](#) in the main article) that are measured for gluten (defined in Section 1 of the [S1 Text](#)) during the elongation in the last stage of the simulation. The solid blue (orange) lines correspond to the quantities listed along the left (right) y -axis. Solid (hollow) dots represent raw data from simulation.

(TIF)

S1 Text. Four sections providing additional information. Section 1 describes the composition of gluten and the other simulated systems, section 2 contains the supporting figures, section 3 justifies the choice of the density for simulations, section 4 is a discussion of the chosen wall interaction model.

(PDF)

S1 Data. Average values of the parameters measured during various stages of simulations.

(XLSX)

S1 DSB Model. The program used for simulations.

(GZ)

Acknowledgments

We thank Mariusz Raczkowski for his help in developing the potential for the disulfide bonds.

Author Contributions

Conceptualization: Łukasz Mioduszewski, Marek Cieplak.

Data curation: Łukasz Mioduszewski.

Formal analysis: Łukasz Mioduszewski, Marek Cieplak.

Funding acquisition: Marek Cieplak.

Investigation: Łukasz Mioduszewski, Marek Cieplak.

Methodology: Łukasz Mioduszewski, Marek Cieplak.

Project administration: Łukasz Mioduszewski, Marek Cieplak.

Resources: Łukasz Mioduszewski, Marek Cieplak.

Software: Łukasz Mioduszewski, Marek Cieplak.

Supervision: Marek Cieplak.

Validation: Marek Cieplak.

Visualization: Łukasz Mioduszewski.

Writing – original draft: Łukasz Mioduszewski.

Writing – review & editing: Łukasz Mioduszewski, Marek Cieplak.

References

1. Shewry PR. Wheat. *J Exp Bot.* 2009; 60(6): 1537–1553. <https://doi.org/10.1093/jxb/erp058> PMID: 19386614
2. Kieffer R. The role of gluten elasticity in the baking quality of Wheat. In: Lafiandra D, Masci S, D'Ovidio R, editors. *The gluten proteins.* Cambridge: Royal Society of Chemistry; 2004.
3. Wieser H. Chemistry of gluten proteins. *Food Microbiol.* 2007; 24(2): 115–119. <https://doi.org/10.1016/j.fm.2006.07.004> PMID: 17008153
4. Shewry P, Halford N, Belton PS, Tatham A. The structure and properties of gluten: an elastic protein from wheat grain. *Phil Trans Roy Soc B: Biol Sci.* 2002; 357(1418): 133–142. <https://doi.org/10.1098/rstb.2001.1024> PMID: 11911770
5. Ahmed J, Ramaswamy H, Raghavan V. Dynamic viscoelastic, calorimetric and dielectric characteristics of wheat protein isolates. *J Cereal Sci.* 2008; 47(3): 417–428. <https://doi.org/10.1016/j.jcs.2007.05.013>
6. Blanch E, Kasarda D, Hecht L, Nielsen K, Barron L. New insight into the solution structures of wheat gluten proteins from Raman optical activity. *Biochemistry.* 2003; 42(19): 5665–5673. <https://doi.org/10.1021/bi027059y> PMID: 12741823
7. Singh H, MacRitchie F. Application of Polymer Science to Properties of Gluten. *J Cereal Sci.* 2001; 33(3): 231–243. <https://doi.org/10.1006/jcrs.2000.0360>
8. Shewry P, Popineau Y, Lafiandra D, Belton PS. Wheat glutenin subunits and dough elasticity: findings of the EUROWHEAT project. *Trends Food Sci Technol.* 2000; 11(12): 433–441. [https://doi.org/10.1016/S0924-2244\(01\)00035-8](https://doi.org/10.1016/S0924-2244(01)00035-8)
9. Mioduszewski Ł, Cieplak M. Disordered peptide chains in an α -C-based coarse-grained model. *Phys Chem Chem Phys.* 2018; 20: 19057–19070. <https://doi.org/10.1039/C8CP03309A> PMID: 29972174
10. Mioduszewski Ł, Różycki B, Cieplak M. Pseudo-Improper-Dihedral Model for Intrinsically Disordered Proteins. *J Chem Theory Comput.* 2020; 16(7): 4726–4733. <https://doi.org/10.1021/acs.jctc.0c00338> PMID: 32436706
11. Rasheed F, Newson W, Plivelic T, Kuktaite R, Hedenqvist M, Gällstedt M, et al. Structural architecture and solubility of native and modified gliadin and glutenin proteins: non-crystalline molecular and atomic organization. *Roy Soc Chem Adv.* 2014; 4(4): 2051–2060. <https://doi.org/10.1039/C3RA45522J>
12. Kasarda DD. Contrasting molecular models for a HMW-GS. In: *Proceedings of the International Meeting on Wheat Kernel Proteins Molecular and Functional Aspects.* Witerbo: Universita degli Studi della Toscana; 1994.
13. Matsushima N, Creutz CE, Kretsinger RH. Polyproline, β -turn helices. Novel secondary structures proposed for the tandem repeats within rhodopsin, synaptophysin, synexin, gliadin, RNA polymerase II, hordein, and gluten. *Proteins.* 1990; 7(2): 125–155. <https://doi.org/10.1002/prot.340070204> PMID: 2139224
14. Belton PS. The molecular basis of dough rheology. *Bread making: Improving quality.* 2003; 13: 273–287. <https://doi.org/10.1533/9781855737129.2.273>
15. Chwastyk M, Jaskólski M, Cieplak M. The volume of cavities in proteins and virus capsids. *Proteins.* 2016. <https://doi.org/10.1002/prot.25076> PMID: 27231838
16. Chwastyk M, Jaskólski M, Cieplak M. Structure-based analysis of thermodynamic and mechanical properties of cavity-containing proteins—case study of plant pathogenesis-related proteins of class 10. *FEBS J.* 2014; 281(1): 416–429. <https://doi.org/10.1111/febs.12611> PMID: 24206126
17. Kröger M. Shortest multiple disconnected path for the analysis of entanglements in two- and three-dimensional polymeric systems. *Comput Phys Commun.* 2005; 168(3): 209–232. <https://doi.org/10.1016/j.cpc.2005.01.020>

18. Shanbhag S, Kröger M. Primitive Path Networks Generated by Annealing and Geometrical Methods: Insights into Differences. *Macromolecules*. 2007; 40(8): 2897–2903. <https://doi.org/10.1021/ma062457k>
19. Karayiannis NC, Kröger M. Combined Molecular Algorithms for the Generation, Equilibration and Topological Analysis of Entangled Polymers: Methodology and Performance. *Int J Mol Sci*. 2009; 10: 5054–5089. <https://doi.org/10.3390/ijms10115054> PMID: 20087477
20. Hoy RS, Foteinopoulou K, Kröger M. Topological analysis of polymeric melts: Chain-length effects and fast-converging estimators for entanglement length. *Phys Rev E*. 2009; 80: 031803. <https://doi.org/10.1103/PhysRevE.80.031803> PMID: 19905139
21. Shewry PR, Halford NG. Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot*. 2002; 53(370): 947–958. <https://doi.org/10.1093/jexbot/53.370.947> PMID: 11912237
22. Ghavani A, van der Giessen E, Onck PR. Coarse-grained potentials for local interactions in unfolded proteins. *J Chem Theor Comp*. 2013; 9: 432–440. <https://doi.org/10.1021/ct300684j>
23. Tozzini V, Trylska J, Chang C, McCammon JA. Flap Opening Dynamics in HIV-1 Protease Explored with a Coarse-Grained Model. *J Struct Biol*. 2007; 3(157): 606–615. <https://doi.org/10.1016/j.jsb.2006.08.005> PMID: 17029846
24. Szymczak P, Cieplak M. Stretching of proteins in a uniform flow. *J Chem Phys*. 2006; 125(16): 164903. <https://doi.org/10.1063/1.2358346> PMID: 17092135
25. Allen M, Tildesley D. *Computer simulation of liquids*. Oxford: Clarendon Press; 1992.
26. Sułkowska JI, Cieplak M. Selection of optimal variants of Go-like models of proteins through studies of stretching. *Biophys J*. 2008; 95: 317491. <https://doi.org/10.1529/biophysj.107.127233> PMID: 18567634
27. Poma AB, Chwastyk M, Cieplak M. Polysaccharide-protein complexes in a coarse-grained model. *J PhysChem B*. 2015; 119: 12028–12041. <https://doi.org/10.1021/acs.jpcc.5b06141>
28. Korkut A, Hendrickson WA. A force field for virtual atom molecular mechanics of proteins. *Proc Natl Acad Sci USA*. 2009; 106(37): 15667–15672. <https://doi.org/10.1073/pnas.0907674106> PMID: 19717427
29. Hung NB, Le DM, Hoang TX. Sequence dependent aggregation of peptides and fibril formation. *J Chem Phys*. 2017; 147(10): 105102. <https://doi.org/10.1063/1.5001517> PMID: 28915764
30. Mioduszewski Ł, Cieplak M. Protein droplets in systems of disordered homopeptides and the amyloid glass phase. *Phys Chem Chem Phys*. 2020; 22: 15592–15599. <https://doi.org/10.1039/D0CP01635G> PMID: 32613961
31. Robustelli P, Piana S, Shaw DE. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci USA*. 2018; 115(21): E4758–E4766. <https://doi.org/10.1073/pnas.1800690115> PMID: 29735687
32. Qin M, Wang W, Thirumalai D. Protein Folding Guides Disulfide Bond Formation. *Proc Natl Acad Sci USA*. 2015; 112(36): 11241–11246. <https://doi.org/10.1073/pnas.1503909112> PMID: 26297249
33. Tilley K, Benjamin R, Bagorogoza K, Okot-Kotber B, Prakash O, Kwen H. Tyrosine Cross-Links: Molecular Basis of Gluten Structure and Function. *J Agric Food Chem*. 2001; 49(5): 2627–2632. <https://doi.org/10.1021/jf010113h> PMID: 11368646
34. Cazalis R. Homology modeling and molecular dynamics simulations of the N-terminal domain of wheat high molecular weight glutenin subunit 10. *Protein Sci*. 2003; 12(1): 34–43. <https://doi.org/10.1110/ps.0229803> PMID: 12493826
35. Consortium TU. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 2018; 47(D1): D506–D515. <https://doi.org/10.1093/nar/gky1049>
36. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: Protein structure and function prediction. *Nat Methods*. 2015; 12: 7–8. <https://doi.org/10.1038/nmeth.3213> PMID: 25549265
37. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res*. 2015; 43: W174–W181. <https://doi.org/10.1093/nar/gkv342> PMID: 25883148
38. Pedrazzini E, Mainieri D, Marrano CA, Vitale A. Where do Protein Bodies of Cereal Seeds Come From? *Front Plant Sci*. 2016; 7: 1139. <https://doi.org/10.3389/fpls.2016.01139>
39. Kremer K, Grest GS. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J Chem Phys*. 1990; 92(8): 5057. <https://doi.org/10.1063/1.458541>
40. Densities of Materials. 2010 [cited 30 July 2020]. In: *Engineering Toolbox* [Internet]. Available from: http://www.engineeringtoolbox.com/density-materials-d_1652.html.
41. Abang Zaidel DN, Chin NL, Yusof YA. A Review on Rheological Properties and Measurements of Dough and Gluten. *J Appl Sci*. 2010; 10: 2478–2490. <https://doi.org/10.3923/jas.2010.2478.2490>

42. Sikora M, Sułkowska JI, Cieplak M. Mechanical Strength of 17 134 Model Proteins and Cysteine Slip-knots. *PLoS Comput Biol*. 2009; 5: 1–15. <https://doi.org/10.1371/journal.pcbi.1000547> PMID: 19876372
43. Anjum F, Khan M, Din A, Saeed M, Pasha I, Arshad M. Wheat Gluten: High Molecular Weight Glutenin Subunits, Structure, Genetics, and Relation to Dough Elasticity. *J Food Sci*. 2007; 72(3): 56–63. <https://doi.org/10.1111/j.1750-3841.2007.00292.x> PMID: 17995810
44. Wu Y, Messing J. Proteome balancing of the maize seed for higher nutritional value. *Front Plant Sci*. 2014; 5. <https://doi.org/10.3389/fpls.2014.00240> PMID: 24910639
45. Wu Y, Wang W, Messing J. Balancing of sulfur storage in maize seed. *BMC Plant Biol*. 2012; 12(1): 77. <https://doi.org/10.1186/1471-2229-12-77> PMID: 22646812
46. Tsai CY. Genetics of Storage Protein in Maize. In: Janick J, editor. *Plant Breeding Reviews: Volume 1*. Boston, MA: Springer US; 1983. p. 103–138.
47. Chen P, Shen Z, Ming L, Li Y, Dan W, Lou G, et al. Genetic Basis of Variation in Rice Seed Storage Protein (Albumin, Globulin, Prolamin, and Glutelin) Content Revealed by Genome-Wide Association Analysis. *Front Plant Sci*. 2018; 9: 612. <https://doi.org/10.3389/fpls.2018.00612> PMID: 29868069
48. Jiang C, Cheng Z, Zhang C, Yu T, Zhong Q, Shen JQ, et al. Proteomic analysis of seed storage proteins in wild rice species of the *Oryza* genus. *Proteome Sci*. 2014; 12(1). <https://doi.org/10.1186/s12953-014-0051-4> PMID: 25505850
49. Muench DG, Okita TW. The Storage Proteins of Rice and Oat. In: Larkins BA, Vasil IK, editors. *Cellular and Molecular Biology of Plant Seed Development*. Dordrecht: Springer Netherlands; 1997. p. 289–330.
50. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3rd ed. USA: Cambridge University Press; 2007.
51. Licup AJ, Münster S, Sharma A, Sheinman M, Jawerth LM, Fabry B, et al. Stress controls the mechanics of collagen networks. *PNAS*. 2015; 112(31): 9573–9578. <https://doi.org/10.1073/pnas.1504258112> PMID: 26195769
52. Wellner N, Mills E, Brownsey G, Wilson R, Brown N, Freeman J, et al. Changes in protein secondary structure during gluten deformation studied by dynamic Fourier transform infrared spectroscopy. *Bio-macromolecules*. 2005; 6(1): 255–261. <https://doi.org/10.1021/bm049584d> PMID: 15638528
53. Starzyk A, Cieplak M. Denaturation of proteins near polar surfaces. *J Chem Phys*. 2011; 135: 235103. <https://doi.org/10.1063/1.3665930> PMID: 22191903
54. Aklonis JJ, MacKnight WJ. *Introduction to Polymer Viscoelasticity* Hoboken. NJ: Wiley-Interscience; 2005.
55. Grest GS, Pütz M, Everaers R, Kremer K. Stress-strain relation of entangled polymer networks. *J Non-Cryst Solids*. 2000; 274(1-3): 139–146. [https://doi.org/10.1016/S0022-3093\(00\)00224-6](https://doi.org/10.1016/S0022-3093(00)00224-6)
56. Létang C, Piau M, Verdier C. Characterization of wheat flour–water doughs. Part I: Rheometry and microstructure. *J Food Eng*. 1999; 41: 121–132. [https://doi.org/10.1016/S0260-8774\(99\)00082-5](https://doi.org/10.1016/S0260-8774(99)00082-5)
57. Khatkar BS. Dynamic rheological properties and bread-making qualities of wheat gluten: effects of urea and dithiothreitol. *J Sci Food Agric*. 2005; 85(2): 337–341. <https://doi.org/10.1002/jsfa.1974>
58. Song Y, Zheng Q, Zhang Q. Rheological and mechanical properties of bioplastics based on gluten- and glutenin-rich fractions. *J Cereal Sci*. 2009; 50(3): 376–380. Special Section: Enzymes in Grain Processing. <https://doi.org/10.1016/j.jcs.2009.07.004>
59. Dobraszczyk B, Morgenstern M. Rheology and the breadmaking process. *J Cereal Sci*. 2003; 38(3): 229–245. [https://doi.org/10.1016/S0733-5210\(03\)00059-6](https://doi.org/10.1016/S0733-5210(03)00059-6)
60. Sánchez H, Osella C, Torre MA. Optimization of Gluten-Free Bread Prepared from Cornstarch, Rice Flour, and Cassava Starch. *J Food Sci*. 2006; 67: 416–419.
61. Cieplak M, Robbins MO. Nanoindentation of virus capsids in a molecular model. *J Chem Phys*. 2010; 132(1): 015101. <https://doi.org/10.1063/1.3276287> PMID: 20078182
62. Parchment O, Shewry P, Tatham A, Osguthorpe D. Molecular Modeling of Unusual Spiral Structure in Elastomeric Wheat Seed Protein. *Cereal Chem*. 2001; 78(6): 658–662. <https://doi.org/10.1094/CCHEM.2001.78.6.658>
63. Koehler P, Wieser H, Konitzer K. *Celiac disease and gluten*. Burlington: Elsevier Science; 2014.
64. Dziuba M, Dziuba J, Iwaniak A. Bioinformatics-aided characteristics of the structural motifs of selected potentially celiac-toxic proteins of cereals and leguminous plants. *Polish J Food Nutr Sci*. 2007; 57(4): 405–414.
65. McAdam S. Getting to grips with gluten. *Gut*. 2000; 47(6): 743–745. <https://doi.org/10.1136/gut.47.6.743> PMID: 11076867