

# Rothamsted Repository Download

## A - Papers appearing in refereed journals

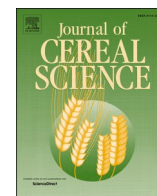
Shewry, P. R. and Belton, P. S. 2024. What do we really understand about wheat gluten structure and functionality? *Journal of Cereal Science*. 117, p. 103895. <https://doi.org/10.1016/j.jcs.2024.103895>

The publisher's version can be accessed at:

- <https://doi.org/10.1016/j.jcs.2024.103895>

The output can be accessed at: <https://repository.rothamsted.ac.uk/item/98zz0/what-do-we-really-understand-about-wheat-gluten-structure-and-functionality>.

© 20 March 2024, Please contact [library@rothamsted.ac.uk](mailto:library@rothamsted.ac.uk) for copyright queries.



# What do we really understand about wheat gluten structure and functionality?

Peter R. Shewry<sup>a,\*</sup>, Peter S. Belton<sup>b</sup>

<sup>a</sup> Rothamsted Research, Harpenden, Hertfordshire, AL5 JQ, UK

<sup>b</sup> School of Chemistry, University of East Anglia, Norwich Research Park, Norwich, NR4 7JT, UK

## ARTICLE INFO

### Keywords:

Wheat  
Gluten  
Gliadin  
Breadmaking  
Functionality  
Glutenin

## ABSTRACT

The structure and functional properties of wheat gluten have fascinated cereal chemists for over a century and a range of approaches have been taken to understand the structures and interactions of the gluten protein complex and how these are established. Nevertheless, our knowledge is still far from complete. We therefore review the current state of our knowledge and identify gaps and priorities for future research.

The evidence for the forces that determine the interactions of the individual proteins in the gluten complex is re-evaluated, which allows us to define the relative contributions of covalent disulphide bonds and non-covalent forces (hydrogen bonds, hydrophobic and electrostatic interactions) and to relate these interactions to the amino acid sequences, structures and properties of the individual protein subunits.

We also discuss the evidence for the pathway of gluten protein synthesis, deposition and assembly in the developing grain and how the assembly may be modified during the maturation of the grain.

## 1. Introduction

Wheat gluten is almost certainly the most widely studied plant protein: a search of the Web of Science database (1970-present) carried out in August 2023 identified over 10,000 articles with “gluten” in the title and over 14,000 if “gliadin” or “glutenin” were also used as search terms. Furthermore, research on gluten dates back almost 300 years, since Beccari described the first isolation of gluten in 1745.

Most studies of wheat gluten have focused on understanding its role in determining the unique biomechanical properties of dough made from wheat flour, a combination of viscosity, extensibility and elasticity, and how these properties in turn determine the ability to make bread, noodles, pasta and other processed foods. Although structurally-related proteins are present in rye and barley, doughs from these cereals do not have the same properties as wheat doughs and it is not possible to make similar foods.

Although understanding functional properties has been the driver for most studies of gluten proteins it is important to bear in mind that this is not their biological role. Gluten proteins are the major group of storage proteins which are deposited in the cells of the starchy endosperm during grain development. They are mobilised (hydrolysed) during germination and their biological role is to support the growth and

development of the seedling. This role clearly imposes some constraints on their structures, to ensure that they are efficiently deposited, stored and mobilised. However, their structures may be less constrained than those of structural and metabolic proteins such as enzymes and transporters.

An extra stimulus to the study of wheat gluten came from the development of omics technologies starting in the 1980s: cereal grains were attractive systems for early molecular cloning and proteomic analyses because of the high concentrations of gluten proteins in the developing and mature grain, leading to the availability of massive datasets of gluten protein sequences. These analyses facilitated attempts to understand the relationships between gluten protein sequences, glutenin polymer structure and dough functional properties, leading by the start of the 21st century to models for gluten structure and functionality which are now widely accepted as dogma.

But all knowledge is conditional and models that are developed to provide explanations for biological phenomena can, and should, be tested experimentally, as part of a reiterative process. In this review therefore we re-evaluate our knowledge of wheat gluten structure and properties, propose modifications to our current models and identify priorities for future research.

We will focus on answering three questions.

\* Corresponding author.

E-mail address: [peter.shewry@rothamsted.ac.uk](mailto:peter.shewry@rothamsted.ac.uk) (P.R. Shewry).

<https://doi.org/10.1016/j.jcs.2024.103895>

Received 30 November 2023; Received in revised form 6 February 2024; Accepted 18 March 2024

Available online 20 March 2024

0733-5210/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. What are the forces stabilising individual glutenin polymers, the interactions of individual glutenin polymers with other polymers and with gliadins, and the interactions of gluten proteins with other dough components
2. How do these forces and interactions account for dough physical properties
3. How are these interactions established in the developing grain and modified during grain maturation.

It is also important to note that the structures and interactions of gluten protein are dynamic and may be affected by the methods used to prepare the samples including the mixing of flour and water to form dough. These effects must be taken in account when designing experiments and interpreting data but are only briefly considered here.

However, before doing this we will briefly introduce the diversity of gluten proteins in mature grains.

## 2. Wheat gluten proteins

The combination of high throughput sequencing of DNA with highly sensitive mass spectrometry (MS)-based proteomics has generated massive databases of gluten protein sequences, many of which are overlapping or duplicates. For example, Bromilow et al. (2017a) assembled a database of only 630 gluten protein sequences from over 24,000 gluten protein-related sequences in the UniProt database. However, the number of proteins present in significant amounts in individual wheat grains is much smaller, generally between 50 and 100. For example, Bromilow et al. (2017b) identified 63 gluten proteins in a single cultivar, using MS. Most of these proteins are encoded by single genes, although some may be generated by post-translational modification, and their numbers, proportions and properties vary widely between wheat genotypes.

Gluten proteins are broadly classified into two fractions, the polymeric glutenins and monomeric gliadins, each of which comprises

several groups of proteins. Gluten proteins have been extensively reviewed (Shewry et al., 2009a; Scherf, 2023) and we therefore only briefly summarise their structures and properties here (Table 1).

The gliadin monomers are classically separated based on their electrophoretic mobility at low pH into three groups, the  $\alpha$ -type ( $\alpha$ - and  $\beta$ -) gliadins,  $\gamma$ -gliadins and  $\omega$ -gliadins, which have since been redefined based on their amino acid sequences as sulphur-rich (S-rich) ( $\alpha$ -type,  $\gamma$ -) and S-poor ( $\omega$ -) prolamins. The glutenin polymers are stabilised by inter-chain disulphide bonds (as discussed below) and reduction of these allows the individual subunits to be separated and characterised. This separates two major groups of subunits called high molecular weight subunits of glutenin (HMW-GS) (which constitute a separate prolamin group, the HMW prolamins) and the S-rich B-type low molecular weight subunits of glutenin (LMW-GS). The HMW-GS are further classified into two types, x-type and y-type, which differ in their molecular masses and cysteine contents. In addition, "mutant" forms of  $\alpha$ -type/ $\gamma$ -gliadins and of  $\omega$ -gliadins which are able to form inter-chain disulphide bonds are present: these are referred to as C-type and D-type LMW-GS, respectively (see below) (Table 1).

Despite their wide diversity all gluten proteins are related, being derived from the same ancestral protein. They also share a number of features which determine their three-dimensional structures, their interactions and the functional properties of gluten and dough.

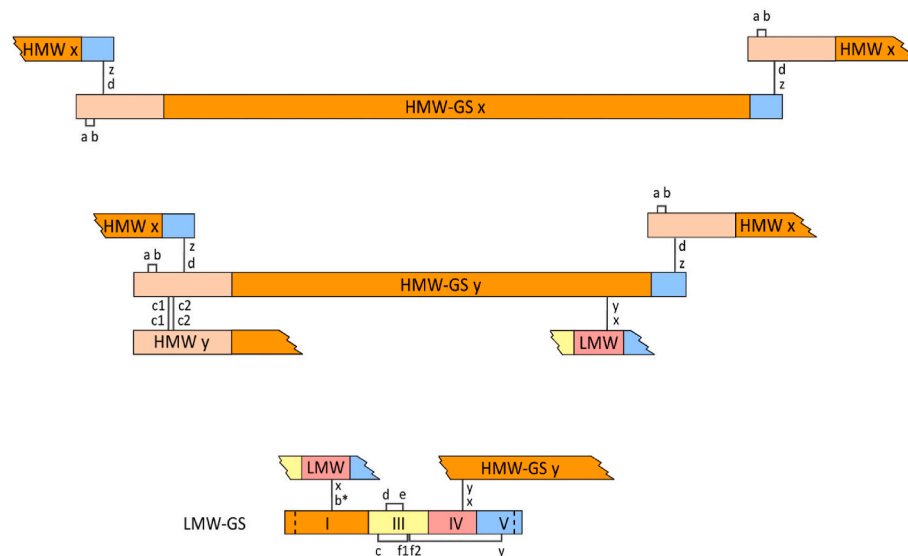
1. Their amino acid sequences can be divided into several "domains", one of which consists of tandem or interspersed repeats of one or more short (3–12 amino acid) peptide motifs.
2. The repetitive domains differ in length between and within different groups of gluten proteins, being longer in the HMW-GS and  $\omega$ -gliadins than in the S-rich  $\alpha$ -type and  $\gamma$ -gliadins and LMW-GS
3. The repetitive domains of all gluten proteins are rich in glutamine and proline and, in some proteins, other amino acids (notably phenylalanine in  $\omega$ -gliadins and glycine in HMW-GS).

**Table 1**  
Characteristics of wheat gluten proteins.

Gluten protein group	Protein nomenclature	% total gluten proteins*	Polymeric or monomeric?	Molecular mass	Partial amino acid composition	Domain structure	Consensus peptide repeat motifs
HMW	HMW subunits of glutenin (HMW-GS)	9	polymers	65,000–90,000	30–35% Q	N-terminal domain of 81–104 residues. Central repetitive domain of 400–700 residues.	PGQGQQ
					10–16% P	C-terminal domain of 42 residues.	GYYPSPorLQQ
					15–20% G 0.5–1.5% C 0.7–1.4% K		GQQ
Sulphur-rich	B-type LMW subunits of glutenin (LMW-GS)	24	polymers	30,000–45,000	30–40% Q 15–20 %P	N-terminal sequence of 5–12 residues.	B-type LMW-GS: PPFS/PQQ(QQ)
	$\alpha$ -type $\alpha$ -type gliadins ( $\alpha$ - and $\beta$ -gliadins)	35	monomers		2–3% C <1% K	Central repetitive domain of 80–150 residues.	$\alpha$ -type gliadins PForYPQQ(QQ)
	C-type LMW-GS	nd	polymers			C-terminal domain of 130–170 residues	PFPQ(Q)PQ(Q)
	$\gamma$ -type $\gamma$ -gliadins	21	monomers				$\gamma$ -gliadins: PFPQ(Q)
	C-type LMW-GS	nd	polymers				PQQ(PQQ)
Sulphur-poor	$\omega$ -type $\omega$ -gliadins	11	monomers	30,000–50,000	40–50% Q 20–30% P 8–9% F 0–0.5% K	Short N-terminal and C-terminal domains of 10–20 residues.	PQQPFPQQ QQQFP
	D-type LMW-GS	nd	polymers		0 cysteine (1 cysteine residue in D-type LMW subunits)	Central repetitive domain of 300–400 residues	

\*Data from Scherf (2023) who do not quote values for C-type and D-type LMW-GS; nd, not determined.

\*authors differ in their interpretation of repeat motifs, those included here are based mainly on Shewry et al. (2009). Residues in parentheses vary in their presence or absence. Standard single letter abbreviations are used: C, cysteine; F, phenylalanine; G, glycine; K, lysine; L, leucine; P, proline; Q, glutamine; S, serine; T, threonine; Y, tyrosine.



**Fig. 1.** Disulphide bonds involving HMW-GS and LMW-GS which have been mapped by proteomic analyses. The figure is reproduced with permission from Scherf (2023). a, b, b\*, c1, c2, d, e, f1, f2, x, y and z indicate individual cysteine residues.

- The amino acid compositions of the domains are therefore responsible for the unusual amino acid compositions of the whole proteins (Table 1).
- The degree of conservation of the repeat motifs varies between prolamin groups, being more highly conserved in the HMW-GS and  $\omega$ -gliadins than in the S-rich  $\alpha$ -type and  $\gamma$ -gliadins and LMW-GS.
- The compositions, sequences and degree of conservation of the repetitive domains also strongly influence the properties of the proteins: the amide group of glutamine is able to form hydrogen bonds (to water or to other proteins) while proline confers rigidity to the polypeptide chain.
- Gluten proteins have only low contents of acidic and basic amino acid residues, particularly lysine, and these are concentrated in the non-repetitive domains
- The numbers of cysteine residues vary: 0 in  $\omega$ -gliadins, 2 to 7 in HMW-GS (2–5 in x-type and typically 7 in y-type), 6 in  $\alpha$ -type gliadins and 8 in  $\gamma$ -gliadins and B-type LMW-GS. The cysteine residues in the  $\alpha$ -type gliadins and  $\gamma$ -gliadins form three and four intra-chain disulphide bonds, respectively, which are conserved in position between the two gliadin types. However, mutations may occur which result in the presence of additional unpaired cysteine residues in  $\alpha$ -type gliadins and in  $\gamma$ -gliadins, allowing the formation of inter-chain disulphide bonds to HMW-GS and LMW-GS. These “mutant” gliadins therefore form part of the glutenin fraction and are called C-type LMW-GS. Similarly, the D-type LMW-GS are “mutant” forms of  $\omega$ -gliadins with single cysteine residues which allow their attachment to glutenin polymers.
- The non-repetitive and repetitive domains of gluten proteins form different types of secondary and higher order structures. The non-repetitive domains are largely globular, with elements of  $\alpha$ -helix and  $\beta$ -sheet, and form tightly folded structures stabilised by intra-chain disulphide bonds. By contrast, the repetitive domains may form extended structures which comprise turns ( $\beta$ -reverse turns and possibly  $\gamma$ -turns) and poly-L-proline II-like structure which in the HMW-GS (and possibly also in the  $\omega$ -gliadins) may be sufficiently regular to form a helical super-secondary structure. These are discussed in detail by Shewry et al. (2003; 2009a).

### 3. Glutenin polymers are stabilised by covalent disulphide bonds

The structures of individual glutenin polymers have not been

determined, but it is clear that they vary widely in molecular mass, from oligomers of mass 100,000 to 150,000 to massive polymers of mass at least  $1 \times 10^6$  (reviewed by Shewry and Lafiandra, 2022). Furthermore, their subunit composition varies with size, from predominantly LMW-GS in oligomers to over 30% HMW-GS in larger polymers (Shewry and Lafiandra, 2022).

There is no doubt that glutenin polymers are stabilised by inter-chain disulphide bonds formed between cysteine residues as the structure is disrupted when these bonds are reduced leading to a loss of elasticity. Some of these bonds have been identified by proteomic analysis (Fig. 1), most of which involve cysteine residues in HMW-GS. These include head-to-tail bonds between two or more x-type subunits, head-to-tail bonds between x-type and y-type subunits, and bonds between the adjacent heads (N-terminal domains) of two y-type subunits. Bonds between LMW-GS and HMW-GS and between two LMW-GS proteins have also been identified. It has therefore been suggested that the HMW subunits form the “backbone” of gluten, providing a framework for interactions with LMW-GS and with gliadins, other glutenin polymers and other dough components by non-covalent forces.

Additional covalent cross-links may be formed between proteins under extreme processing conditions (dityrosine) or by enzyme treatments (notably between glutamine residues catalysed by trans-glutaminase) but there is no evidence that these types of cross-link exist in the mature grain or are formed during dough mixing.

It is therefore necessary to consider the types of non-covalent forces and their role in gluten structure and properties.

### 4. Noncovalent interactions stabilising protein structures and interactions

#### 4.1. Types of non-covalent forces

Non-covalent forces are less stable than covalent bonds and hence more difficult to characterise. Consequently, datasets may be open to different interpretations and results may differ due to effects of environment and processing, notably the choice of solvents and other conditions and the input of mechanical force. Hence, the types of non-covalent interaction, and their relative importance in stabilising gluten and determining its functional properties, are still debated.

Several types of non-covalent interactions may occur in gluten and dough.

1. Hydrogen bonds are a form of electrostatic interaction. They occur in water and between water and dissolved compounds, where they are constantly broken and reformed. It is notable that hydrogen bonds are thermally labile which is consistent with the greatly increased solubility of gluten when heated in alcohol-water mixtures.
2. Hydrophobic interactions refer to the propensity of non-polar molecules to aggregate in aqueous media to exclude water, and the removal of hydrophobic residues from exposure to water by burying them in the core of the protein is a major force which drives protein folding. The hydration energy of amino acids ranges from  $-037 \text{ kJ mol}^{-1}$  (tryptophan) to  $-28.65 \text{ kJ mol}^{-1}$  (arginine), with glutamine being moderately hydrophilic ( $-13.7 \text{ kJ mol}^{-1}$ ). Other abundant amino acids in the gluten protein repetitive domains are proline ( $0.96 \text{ kJ mol}^{-1}$ ), phenylalanine ( $-1.17 \text{ kJ mol}^{-1}$ ), tyrosine ( $-11.97 \text{ kJ mol}^{-1}$ ) and glycine ( $-0.96 \text{ kJ mol}^{-1}$ ) (Shewry et al. 2003).
3. Ionic interactions occur between amino acids of opposite charge, for example, the basic amino acid lysine and aspartic acid or glutamic acid. Charged residues occur rarely in the repetitive domains of gluten proteins (as discussed above) but ionic bonds could form between amino acid residues in the globular domains. The solubility of the proteins in acids, however, does indicate that they may have a significant role in inter-protein interactions (section 4.3.1)

#### 4.2. The “loop and train model”: importance of hydrogen bonds

The general features of gluten rheology have been explained by the “loop and train” model (Belton, 1995, 2005) which is now widely accepted. This proposes that dry gluten is disordered but that hydration of the HMW-GS leads to the formation of regular hydrogen bonded structures by orientation of the  $\beta$ -turns in adjacent  $\beta$ -spirals to form structures resembling “inter-chain”  $\beta$ -sheet. Further hydration results in the replacement of some of the inter-chain hydrogen bonds with hydrogen bonds with water, resulting in an equilibrium between aligned regions (trains) and loop regions. This equilibrium state is disrupted when force is applied, resulting in reformation of the equilibrium state when the force is relaxed.

This model focuses on the HMW-GS and their role in forming the “elastic backbone” of gluten. The HMW-GS have long repetitive domains comprising highly conserved repeated peptide motifs with regularly spaced glutamine residues. This allows the formation of regularly spaced hydrogen bonds which may form “glutamine zips” similar to those formed by protein deposits in the brain in neurodegenerative diseases of humans (Perutz et al., 1994). The importance of hydrogen bonds in stabilising the “loop and train” structures is supported by biophysical analyses of purified HMW-GS and synthetic peptides based on conserved repeat motifs (Tatham et al., 2001; Wellner et al., 2006; Haward et al., 2011).

However, the model is incomplete in two respects.

Firstly, it focuses on the role of the HMW-GS and does not define the roles of the other gluten proteins except to assume that they affect the degree of train formation (Shewry et al., 2003). However, other types of gluten proteins differ greatly from the HMW-GS, particularly in the extent of their repetitive domains and the sequences and degree of conservation of their repeat motifs.

Similarly, the “loop and train” model focuses on hydrogen bonds and does not consider the roles of other types of non-covalent forces in stabilising the structures and interactions of the polymers and monomers. For example, although the sulphur-rich  $\alpha$ -type and  $\gamma$ -gliadins are often considered to act as plasticisers for the glutenins, resulting in greater extensibility, these interactions have not been defined.

#### 4.3. Evidence for non-covalent forces from effects of solvents and salts

The loop and train hypothesis is largely based on evidence from analyses of gluten proteins and fractions in the hydrated solid state, which is the state in the mature grain and in dough.

However, is also often considered that the extractability of gluten proteins in specific solvents is an indication of the types of non-covalent bonds which stabilise their structures and interactions and conclusions drawn based on the extraction of fractions from dough or gluten. It is important to emphasise the distinction between extractability and solubility. Extractability measures the amount of protein that can be removed from a system containing protein in some kind of matrix, be it flour, gluten or some other protein-containing matrix. Solubility is strictly a thermodynamic term which is a measure of the amount of a substance in solution which is in equilibrium with the solid substance in a defined state. In practice true solubility is rarely measured for proteins and extractability is used as a proxy. Even when pure proteins are used, true equilibrium is very difficult to reach.

The extent to which the interaction of gluten with solvents elucidates the nature of protein/protein interaction needs careful analysis. Generally, three types of solvent have been examined: acids (mainly acetic acid), salt solutions and alcohols.

##### 4.3.1. Dilute acids

Acetic acid has been widely used as solvent for gluten proteins: typically a 50 mM (mM) solution has been employed but a wide variety of other concentrations have been used. It is commonly assumed that no distinct chemical effect is involved and that extraction into the solvent is due to a reduction of electrostatic attractive interactions between charged groups on the protein (Iwaki, et al. 2023). Since it is generally accepted that peptide bonds are neutral across the whole pH range (Martin, 2001) effects of this kind can be ignored.

The pH of 50 mM acetic acid is of the order of 3 which is well below the pI's of the gluten proteins, which are typically 6 to 10 (Gobaa, et al. 2008). Under these conditions a significant charge will be present on the proteins and there will be electrostatic repulsion leading to increased solubility. If, however, the effect of acetic acid is simply to give a pH at which charge/charge repulsion results in solubility, a 1 mM solution of a stronger acid, such as HCl, should result in the same solubility as in 50 mM acetic acid. Some support for a simple pH effect comes from the work of MacRitchie and co-workers (MacRitchie, 1987; MacRitchie et al. 1991) who extracted gluten in nine or ten stages with hydrochloric acid over a concentration range of 0.375–1.5 mM, corresponding to a pH range of about 6 to 3. At the lowest pH the cumulative extraction was between 80% and 90% of the total protein, without the disruption of disulphide bonds. The apparent solubilisation of very large polymers may indicate that a suspension was formed rather than a true solution.

The fraction extracted at high pH comprised mainly gliadins and LMW-GS. However, whereas the gliadin content decreased with decreasing pH the glutenin content remained approximately constant. The greatest extraction of HMW-GS and  $\gamma$ -gliadins varied slightly in pH range (4.5–3.5) between different cultivars, indicating variations in buffering capacity. The presence of buffering clearly indicates the attachment of protons to negatively charged groups on the protein and thus supports the suggestion that the solubility in acids is due to decreased attractive electrostatic interactions between charged groups.

Infrared studies of whole gluten and a purified HMW-GS showed a marked increase in the  $\beta$ -turn content in the presence of acetic acid, which is consistent with a high degree of hydration, but gave no evidence of a specific interaction between the acid and the protein (Pézolet, et al. 1992; Belton et al. 1995).

##### 4.3.2. Salts

The effects of salts on solubility and extractability are often referred to by the generic terms “salting in” and “salting out”. These effects are reflections of the specific effects of salts discussed below.

It is well known that addition of sodium chloride to dough increases dough strength. However, the interactions of salts and proteins are not straightforward. Many protein/salt interactions follow the Hofmeister series and Preston (1981) compared the effects of the sodium salts of  $\text{F}^-$ ,  $\text{Cl}^-$ ,  $\text{Br}^-$ ,  $\text{ClO}_4^-$ ,  $\text{I}^-$  and  $\text{SCN}^-$  on the extractability of gluten proteins. He



showed that extractability followed the expected behaviour for the Hofmeister series with extractability at constant salt concentration increasing with the anion size. Solubility also varied with anion concentration, with increasing solubility with increasing concentrations of  $\text{ClO}_4^-$ ,  $\text{I}^-$  and  $\text{SCN}^-$  but decreasing solubility with  $\text{F}^-$  and  $\text{Cl}^-$ . Addition of the salts to a solution of gluten in acetic acid resulted in turbidity of the solution but whereas this increased with ion concentration for  $\text{F}^-$  and  $\text{Cl}^-$ , it increased but then decreased with the larger anions.

Preston (1989) also compared the effects of the same group of ions on dough rheology. Whereas the dough strength was increased by low concentrations (0.05–0.1 mol/L) of all ions, higher concentrations of  $\text{Cl}^-$  and  $\text{Br}^-$  increased dough strength, whereas higher concentrations of  $\text{SCN}^-$  and  $\text{I}^-$  decreased dough strength. These effects of ions on dough strength and protein extractability were interpreted as a combination of two effects. Firstly, all concentrations of ions reduced electrostatic repulsion between charged groups on the proteins, and thus increased protein/protein interactions resulting in dough strengthening. However, at high concentrations the large ions affected water structure and changed inter-protein hydrophobic interactions.

This concept, that ions could be classified into “structure making” and “structure breaking”, was widely used in the earlier literature. However, more recent studies (Lo Nostro and Ninham, 2012; Zhang and Cremer 2006) suggest that the hypothesis is not tenable and that the effects of ions on water structure are very local, with the results suggesting specific interactions of ions with polar and charged groups on the proteins. In particular, large anions have been shown to bind to amide groups (Rembert et al., 2012). Tuhumury et al. (2016) suggested that this binding might extend to specific ion interactions with polar nitrogen atoms such as on the amide side chain of glutamine and evidence for this comes from the infrared spectra of gluten incubated with sodium iodide which showed a decrease in  $\beta$ -sheet content and increase in  $\beta$ -turn content (Wellner et al., 2003). Concomitantly, incubation with sodium chloride showed an increase in  $\beta$ -sheet content, which was also observed by Tuhumury et al. (2014).

Wellner et al. (2003) have applied these observations to the “loop and train” model (discussed above). It was proposed that the role of the anions was to reduce charge/charge repulsion between proteins and reduce the availability of water for protein hydration. Both these effects would increase the tendency for protein/protein interaction and therefore the formation of inter-chain  $\beta$ -sheet structures. This in turn would affect the rheology of the gluten, making it more resistant to extension.

However, evidence from the effects of Hofmeister series of ions has shown that a rather more detailed explanation of the role of small ions is required (Zhang et al. 2005). They considered the surface properties and thermodynamics of ion hydration and showed that small ions caused the dehydration of polar groups by specific ion interactions. Under these conditions the hydrogen bonding between polar groups would not be mediated by water and hence hydrogen bonds between polar groups, particularly amide groups, would be likely to form. A second effect would be to reduce the solubility of hydrophobic structures by increasing the surface tension at the hydrophobic structure/water interface. By contrast, the surface tension effect would be smaller for large ions and counteracted by the direct binding of large ions to amide groups which increases solubility in water.

The solubility of gluten in acids and salts provides clear evidence of the roles of charged groups and hydrogen bonding in the interactions of gluten proteins while the effects of salts may also indicate some role for hydrophobic interactions.

#### 4.3.3. Alcohols

Gliadins are defined as prolamins based on their solubility in alcohol-water mixtures and these are the most widely used solvents. It has also often been stated that the insolubility of gluten in water and the solubility in alcohols is evidence for the hydrophobic nature of gluten proteins. This is not the case, as demonstrated by calculations based on the hydrophilicity of the constituent amino acids and comparison of water

absorption with truly hydrophobic proteins (Shewry et al., 2003).

One of the earliest studies of the solubility of gluten proteins in alcohols, which is still the most detailed, was by Dill and Alsberg (1925) who measured the solubility of gliadins in ethanol, methanol and n-propanol (propan-1-ol). Their method was to dissolve a fixed amount of gliadin in different ratios of alcohols and water and determine the temperature at which the solution became turbid. The minimum temperature at which turbidity was observed in 60–65% ethanol-water was 5–10 °C, for 65–70% methanol-water was 65–70 °C and for 40–60% propan-1-ol-water was 3–5 °C. The dependence of the turbidity temperatures on gliadin concentration in ethanol water mixtures was quite small: a maximum difference of 11 °C between a 0.1% solution and a 25% solution of protein. The authors noted that if the turbid solution was heated the temperature at which it cleared was almost the same as that of turbidity onset which led them to hypothesise that they did not have a true solution but a suspension.

Significant temperature and solvent effects on the extraction of proteins from milled whole wheat were also observed by Byers et al. (1983). However, as noted above, extractability is not necessarily the same as solubility but may indicate solubility. The authors compared four solvent systems (70% ethanol/water, 50% propan-1-ol/water, 50% propan-1-ol/water with 1% acetic acid, 55% propan-2-ol/water) at three temperatures: (4, 20 and 60 °C) with extractions carried with and without 2-mercaptoethanol to reduce disulphide bonds. The results show the complex effects of solvent and temperature. For example, 70% ethanol at 60 °C extracted the same amount of total protein (measured as nitrogen) as 70% ethanol with 2-mercaptoethanol at 4 °C, but only the extract made at 60 °C contained proteins with molecular weights greater than 60,000. The amount of nitrogen extracted varied with solvent and temperature: 70% ethanol at 4 °C extracted 19.3 % of total nitrogen and at 60 °C extracted 35%. Similar amounts of nitrogen were extracted with 55% propan-2-ol but 50% propan-1-ol was more effective, extracting 41.9% at 60 °C. However, because the authors separated the extracted proteins by SDS-PAGE after reduction of disulphide bonds the proportions of monomers and polymers in the extracts are not known.

The addition of 2-mercaptoethanol to reduce disulphide bonds resulted in increased extraction, from 25.7% to 44.9% total nitrogen with 50% propan-1-ol at 4 °C, indicating that depolymerisation increased solubility but that not all polymeric species were dissolved, while the inclusion of 1% acetic acid and 2-mercaptoethanol resulted in the highest extraction.

Byers et al. (1983) also reported significant temperature-dependence of solubility in all solvents. This contrasts with the results of Dill and Alsberg (1925) and suggests that they were correct in that their samples were not true solutions.

However, it is important to recognise that Byers et al. (1983) determined how much nitrogen (protein) was extracted at various temperatures whereas Dill and Alsberg (1925) measured the protein that was not in solution. This is important as the work of Morel et al. (2020) suggests that the observation of turbidity is not due to precipitation but may result from a liquid/liquid phase transition separating fractions of different molecular weights. A careful study of the phase behaviour of ethanol solutions of proteins ranging in molecular weight from 25,000 to 300,000 showed that the behaviour of the gliadin fraction (molecular weight below 45,000) was typical of a system showing only long-range interactions, whilst the high molecular weight proteins showed behaviour typical of neutral linear polymers as described by the Flory Huggins theory (Boire, 2013). This is relevant to interpreting the data of Byers et al. (1983) who measured extractability, not solubility, as their extracts may have included large conglomerates as well as dissolved individual protein molecules. This conclusion is strongly supported by the work of Dahesh et al. (2014).

Even with these reservations some conclusions can be drawn about the solubility of gluten proteins in alcohol-water mixtures. Most notably, 50% aqueous propan-1-ol is a better solvent than 70% aqueous ethanol

**Table 2**

The most commonly occurring amino acids in gluten proteins together with their hydration energies and the total hydration energy of gliadins and glutenins. For comparison data for an average of 314 proteins is included. Data taken from Shewry et al., (2003).

Amino Acid	Free energy of hydration (Joules)	mol% in average of 314 proteins	mol% in gliadins	mol% in glutenins
Glutamine	−13.17	9.9	34.5	28.9
Proline	0.96	5.2	16.2	11.9
Glycine	−0.96	8.4	3.1	7.5
total hydration energy (kcal/mol protein)		−790.48	−771.2	−795.3

or 55% propan-2-ol, although the three mixtures have similar dielectric constants of about 40 at 25 °C. The differences in solubility cannot be attributed to differences in solution polarity and may lie in the relative propensity of the alcohols to form alcohol-rich clusters within the solution which may interact with hydrophobic regions on the proteins to increase solubility. This propensity is greater in propan-1-ol-water than in ethanol-water (Wakisaka et al., 2001).

The low dielectric constant of alcohol/water mixtures compared to water reduces the degree of ionisation of salts and acids and charged groups on proteins so there will be a general reduction of electrostatic attraction, which will tend to increase protein solubility. However, the fact that the addition of 1% acetic acid to propan-1-ol/water increases the amount of protein extracted suggests that the interaction of acetic acid with gluten proteins involves more than a simple ionic effect.

The question arises as to why the proteins are less soluble in propan-2-ol/water. The boiling point of the secondary alcohol is lower than that of the primary alcohol, indicating a less hydrogen bonded structure, and it is subject to phase separation on addition of salts to a mixture of it with water. This indicates that it has a less hydrophilic character than the primary alcohol and may indicate that a certain degree of hydrogen bonding capacity is required for solution of the proteins.

#### 4.4. Summary of forces stabilising protein:protein interactions

The results of solubility and extraction studies indicate firstly the subtlety of the interactions that govern the uptake of proteins in solvents and, secondly, to the existence of a range of interactions between gluten proteins.

There is clear evidence for a role of charge/charge interactions and hydrogen bonding between the HMW-GS. However, there is little direct evidence for inter-protein interactions between gliadins or between gliadins and glutenin polymers, with the exception of the binding of  $\omega$ -gliadin to glutenins (Morel et al., 2020). The molecular interactions of the LMW-GS have not been determined but they are likely to more

similar to those of the more closely-related  $\alpha$ -type and  $\gamma$ -gliadins. Molecular dynamics simulations of  $\alpha$ -gliadin (Yu et al., 2023) predicted regions of hydrophobic  $\alpha$ -helix, which increased in contact with solvent in the presence of ethanol. The simulations also suggested that the protein retains most of its structure but expands slightly when dissolved in alcohol-water (Yu et al., 2023) which consistent with the globular N-terminal and C-terminal domains being stabilised by intra-chain disulphide bonds.

The loop and train model emphasises the role of hydrogen bonding between HMW-GS. Similarly, hydrogen bonding is important in  $\omega$ -gliadins in the low water content state (Wellner et al., 1996) and may account for the binding of  $\omega$ -gliadins to HMW-GS discussed above. As water content increases the degree of  $\beta$ -sheet structure in both  $\omega$ -gliadins and HMW-GS decreases and is replaced by  $\beta$ -turn structure, indicating a marked decrease in interprotein interactions.

However, it is important to note that the HMW-GS and  $\omega$ -gliadins together represent only about 20% of the total gluten proteins and hydrogen bonds do not appear to be important in the interactions of the  $\alpha$ -type and  $\gamma$ -gliadins (and probably also the LMW-GS). In this respect, it may be relevant that the  $\omega$ -gliadins and HMW-GS both have long repetitive domains with highly conserved repeat motifs whereas the  $\alpha$ -type gliadins,  $\gamma$ -gliadins and LMW-GS have globular non-repetitive domains and shorter repetitive domains with more degenerate repeat motifs. Table 2 lists the three most commonly occurring amino acids in the gluten proteins, the corresponding hydration energies of the acids and the total calculated hydration energy of the protein molecule. It shows that, despite their unusual compositions compared to most proteins, the hydration energies of gluten proteins are comparable. In this case it seems anomalous that the proteins are water-insoluble. This may be due to the very large number of interprotein hydrogen bonds due to the highly abundant glutamine which makes it statistically very unlikely that all the bonds will be simultaneously broken, thus resulting in insolubility (Belton 1995).

The extent of hydrophobic interactions in gluten proteins is less clear. Their ready solubility in aqueous solutions of acids may imply that hydrophobic interactions are not important, or, at least are considerably less important than interactions between charged groups. However, if there are no hydrophobic interactions in gluten it is difficult to explain the insolubility in water and the solubility in alcohol-water mixtures, especially the increased solubility in propan-1-ol compared to ethanol. There still remains the problem of the solubility of the proteins in alcohol/water mixtures. The specific structure of the repeat units may hold some explanation. Proline, which is hydrophobic, occurs in the repeat motifs of all gluten proteins. The presence of alcohol may induce the hydrophobic proline to enter solution and reduce the interprotein interactions allowing dissolution.

It therefore appears that, in addition to disulphide bonds, the protein:protein interactions in gluten result from a balance between hydrophobic, electrostatic and hydrogen bonding. The relative

**Table 3**

Summary of the proposed forces that stabilise the structures of gliadins and glutenin subunits and their interactions to form gluten.

Subunits	Domains	Structure	Interactions
HMW-GS	Repetitive	Extended regular $\beta$ -spiral structure stabilised by regularly spaced hydrogen bonds	Regularly spaced hydrogen bonds between adjacent chains of HMW-GS (and possibly also $\omega$ -gliadins) form cross- $\beta$ -sheet (glutamine zips)
	N- and C-terminal	Globular structure stabilised by non-covalent forces	Inter-chain disulphide bonds stabilise high molecular mass cross-linked polymers
LMW-GS	Repetitive	Extended irregular structure with irregularly spaced hydrogen bonds	May interact with gliadins and other LMW-GS by hydrophobic and electrostatic interactions
	C-terminal	Compact globular structure stabilised by non-covalent forces and intra-chain disulphide bonds	Inter-chain disulphide bonds form oligomers with other LMW-GS and cross-link to polymers of HMW-GS
$\alpha$ -type gliadins and $\gamma$ -gliadins	Repetitive	Extended irregular structures with irregularly spaced hydrogen bonds	Bind to other gliadins and glutenin polymers by hydrophobic and electrostatic interactions
	C-terminal	Compact globular structure stabilised by intra-chain disulphide bonds and non-covalent forces	May interact with other gliadins and LMW-GS by hydrophobic and electrostatic interactions
$\omega$ -gliadins	Repetitive	Extended regular structure stabilised by regularly spaced hydrogen bonds	Regularly spaced hydrogen bonds between adjacent chains (and possibly also HMW-GS) form cross- $\beta$ -sheet

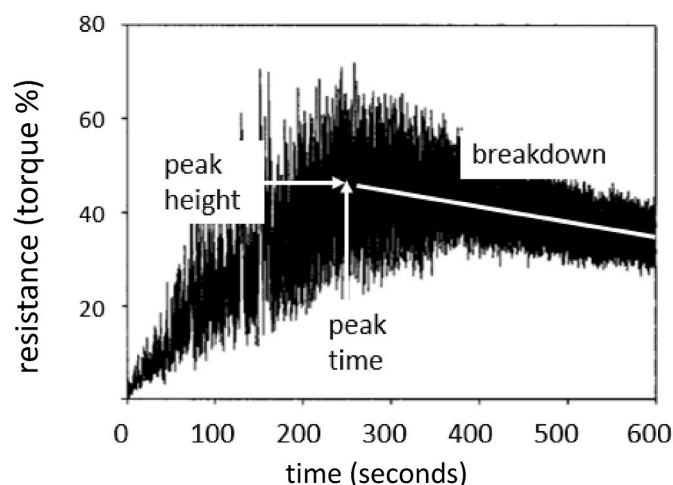


Fig. 2. Typical Mixograph curve for bread wheat.

contributions of these forces also vary between gluten protein types. Hydrogen bonds are important between the repetitive domains of  $\omega$ -gliadins and HMW-GS which account for high proportions of the whole proteins and have highly conserved sequence motifs. By contrast, hydrophobic and electrostatic forces are more important between the  $\alpha$ -type gliadins,  $\gamma$ -gliadins and LMW-GS which have more extensive non-repetitive domains and less well-conserved repeat motifs.

These interactions are summarised in Table 3.

## 5. Interactions of gluten proteins with other proteins and lipids

Isolated wheat gluten consists of about 80% protein on a dry weight basis. Most of the proteins present are “gluten proteins” but other proteins are present in small amounts, with small amounts of lipids and carbohydrates. These “non-gluten proteins” and other components may be entrapped in the gluten network but may also be bound to the gluten proteins. Consequently, they may influence the properties.

The non-gluten proteins include two groups of proteins with molecular masses below 30 kDa which are related to gliadins but lack repetitive domains. The low molecular weight gliadins/purinins have masses of about 17 kDa–19 kDa and are most closely related to the  $\gamma$ -gliadins in sequence. They may, therefore, be similar to the “ancestral” prolamins, before the development of the repetitive sequence domains. The farinins, also called avenin-like proteins, comprise two types, with the a-type having masses of about 17,000 and the b-type proteins masses of about 30,000 due to the presence of a duplicated sequence of about 120 residues. Whereas the purinins appear to behave similarly to gliadins in dough, over-expression of a transgene encoding a b-type farinin resulted in improved flour mixing properties and an increased proportion of large glutenin polymers, presumably due to the formation of inter-chain disulphide bonds (reviewed by Shewry, 2019).

The major group of low molecular mass proteins, accounting for 2–4% total grain protein, are  $\alpha$ -amylase and trypsin inhibitors (ATIs). Multiple forms are present, with molecular masses ranging from about

11,500 to 15,500, and they may be present as monomers, dimers or tetramers. They are of particular interest in relation to adverse reactions to wheat consumption (allergy, coeliac disease and non-coeliac wheat sensitivity) (reviewed by Geisslitz et al., 2021). ATIs have also been reported to contribute to the cooking quality of pasta, where they were initially reported to be glutenin components (called durum sulphur-rich glutenin, DSG) (Kobrehel and Alary, 1989).

It is important to distinguish between the effects of lipids on dough in the bread-making process, which are considerable, and specific interactions with gluten. Gluten contains non-polar lipids, mainly triacylglycerols, and polar lipids, mainly glycolipids and phospholipids.  $^{31}\text{P}$  NMR studies of phospholipids combined with freeze fracture electron microscopy (Marion, et al., 1987) showed that there was no specific lipid/protein binding and the phospholipids were in the form of vesicles. Similarly, a detailed study of extraction using solvents of differing polarity showed that, although the removal of lipids had significant effects on dough rheology, the effects on gluten were small (Georgopoulos et al., 2006).

## 6. Establishment of interactions during grain development

The precise pathway, location and timing of assembly of the gluten protein complex are still incompletely understood but it is clear that it is a complex multistage process. By contrast, the time course of synthesis of the gluten protein subunits (gliadins and glutenins) has been widely studied (see, for example, Shewry et al., 2009a). This shows that although the precise timing varies with the growth conditions, gluten proteins accumulate during the middle “grain filling” period of development with only small differences between the timing and rates of accumulation for individual subunits.

Gluten proteins are “secretory” proteins, being synthesised on the rough endoplasmic reticulum (ER) with N-terminal signal sequences which direct the polypeptides through the membrane into the lumen of the ER, the signal peptides being proteolytically cleaved to release the mature proteins. Secretory proteins are folded rapidly after synthesis in the lumen of the ER and this process may be aided by the enzyme protein disulphide isomerase (PDI), which catalyses the formation and breaking of disulphide bonds, and one or more molecular chaperones, such as the HSP70-related binding protein BiP (Tosi, 2021), which bind to the newly synthesised proteins to ensure that they remain able to fold correctly (as opposed to forming insoluble mis-folded forms). Most gliadins are transported from the lumen of the ER into the Golgi apparatus and then in Golgi-derived vesicles to the vacuole where they aggregate to form protein deposits. This is the classic secretory protein pathway with proteins being sorted within the Golgi apparatus and directed either to the vacuole (as for gluten proteins) or the plasma membrane where they are secreted outside the cell.

However, a second pathway for the transport of storage proteins from the lumen of the ER to the vacuole occurs in many seeds, in which the Golgi apparatus is by-passed and small protein bodies/vesicles derived directly from the ER are internalised into the vacuoles by a process analogous to autophagy. It is clear that the wheat glutenin polymers are predominantly transported via this pathway, although the demarcation is not clear cut with some gliadins being transported via the

Table 4

Key questions relating to gluten protein structure and functionality.

Cell biology	Where and how are glutenin polymers initially assembled? Is the size of the polymers determined by the subunit composition and the site of assembly ? Do individual polymers “grow” by the addition of subunits or oligomers or “fuse” with other polymers? If so, where and how does this occur? How do the polymers increase in size during grain maturation and how is this affected by environmental conditions?
Protein chemistry	Do glutenin polymers have defined structures with conserved disulphide bonds? What non-covalent forces stabilise the structures and interactions of glutenin polymers? How do the structures and interactions of glutenin polymers determine gluten elasticity? What are the forces that stabilise the binding of gliadins to the glutenin network and how do these determine gluten viscosity and extensibility.
Food processing	What are the effects of processing on the structures and interactions of gluten proteins and how do these affect the functional properties?



Golgi-independent pathway and some glutenin polymers via the Golgi pathway.

There is a massive volume of literature on the mechanisms of protein sorting and trafficking in plant cells which has been reviewed by Xiang et al. (2013) and by Tosi (2021) for wheat. Studies have focused on the types of vesicles and the signals which are recognised by the sorting mechanisms, including vacuolar sorting determinants (VSDs) which are protein sequences or structural motifs which are recognised in the Golgi pathway. However, such studies have failed to identify sequences which are responsible for the trafficking of gluten proteins via the Golgi-dependant and Golgi-independent pathways and it is possible that the segregation results from differences in the propensity of the gliadins and glutenin polymers to aggregate, with the glutenin polymers forming insoluble aggregates within the lumen of the ER while the gliadins remaining soluble until they are concentrated in the Golgi apparatus. This is clearly a challenging topic for future research! However, in addition to academic interest, the pathway of gluten protein trafficking and deposition is important for grain utilisation as it is likely to affect the assembly, structure and interactions of the glutenin polymers.

It can be assumed that the initial folding of the glutenin subunits and disulphide bond formation, including the formation of inter-chain bonds that stabilise the glutenin polymers, occur in the lumen of the ER. However, it is unlikely that this initial assembly results in the mature polymer structure. In fact, it is widely considered that both polymer growth and rearrangement occur during grain maturation. The former is most easily observed as increases in the amount and size of glutenin polymers during grain maturation (Shewry et al., 2009b; Koga et al., 2020). It is particularly important to understand where and how this polymerisation occurs because it is strongly influenced by the environment and consequently has a major impact on quality (Ausennac et al., 2020; Branlard et al., 2020, 2023; Koga et al., 2020).

Consequently, there are two major questions: what are mechanisms of polymerisation and how is the polymerisation affected by the spatial location (or separation) of the proteins in the starchy endosperm cells of the developing grain.

## 1. Mechanisms of polymer growth

We know that polymer size increases during the desiccation of the grain and that this increase also occurs when immature grains are harvested and artificially dried (Koga et al., 2017). Furthermore, it is becoming increasingly clear that non-covalent interactions between glutenin polymers results in highly stable structures which may only be broken down under extreme conditions (as discussed by Shewry and Lafiandra, 2022). However, the extent to which dehydration also drives the formation of non-covalent interactions between gliadins and glutenin polymers is not known and it is possible that different types of gliadins and glutenin polymers form separate phases, as demonstrated for films deposited by evaporation of mixed solutions of  $\alpha$ -type and  $\omega$ -gliadins (Tatham et al., 1999).

A further question is whether the growth of polymers also results from the formation of new disulphide bonds and/or the re-arrangement of disulphide bonds which are formed during the initial assembly. This is difficult to establish as direct mapping of disulphide bonds has only been carried out on a small number of samples and (as discussed above) it is difficult to ensure that all non-covalent interactions are disrupted during polymer separation. However, it is generally accepted that the formation of some new inter-chain disulphide bonds and the rearrangement of existing bonds do occur during grain maturation and that these processes are regulated by thiol oxidoreductase enzymes and low molecular mass redox pairs, notably glutathione/oxidised glutathione and ascorbic acid/dehydroascorbic acid (reviewed by Osipova et al., 2012). We do not know how these processes are controlled but we do know that the structures of the polymers in the mature grain are influenced by both the genotype (allelic variation in protein composition) and the environment during grain development (Branlard et al., 2023).

Finally, the ability of glutenin polymers to grow and interact will depend on their spatial locations within the cells of the developing starchy endosperm. In addition to large glutenin polymers enriched in HMW-GS, the mature wheat grain also contains smaller alcohol-soluble polymers consisting mainly, or totally, of LMW-GS, and it has been suggested that LMW-GS oligomers become attached to larger glutenin polymers (reviewed by Shewry and Lafiandra, 2022). It is also possible that small alcohol-soluble polymers are trafficked via the Golgi-dependant route, with gliadins.

The protein deposits from the Golgi-dependant and Golgi-independent pathways ultimately fuse to give a continuous protein matrix in the mature grain. However, we do not know whether this fusion results in mixing of the protein contents, which is perhaps unlikely with the falling water content and without mechanical input, or the proteins form microphases as discussed above (Tatham et al., 1999). Hence, it is possible that the contents of individual protein bodies are only brought together, allowing the growth and restructuring of polymers, during dough mixing.

## 7. Effects of sample preparation and processing on interactions

It is clear that the structure of gluten is not fixed but fluid, being affected by solvents used, other dough components and the input of mechanical energy. Hence, the “structure” determined for any sample will be a “snap shot” which reflects the methods used to prepare and analyse the sample as well as intrinsic differences between the flours used.

The effects of mechanical input are readily observed using a recording dough mixture such as a Mixograph (Fig. 2). Measurement of the resistance (torque) during mixing shows a steady increase, reaching a peak after several minutes (depending on the sample and conditions). This is considered to result from the optimisation of the structures and interactions of the gluten of the gluten proteins. Peak height and time are therefore used as measures of dough strength.

However, if the mixing is continued after the peak resistance is reached the resistance decreases. This is because the “optimum interactions” are being broken down, presumably by mechanical disruption of the non-covalent protein interactions and possibly also shearing of polymers.

In order to study the early development of protein interactions during processing, Wellner et al. (2005) prepared protein bodies from developing grain without mechanical input and followed the development of hydrogen bonding during several cycles of deformation and relaxation using FT-IR spectroscopy. This showed a build-up of persistent  $\beta$ -sheet structure stabilised by hydrogen bonding, with the structure being similar to that of gluten prepared from dough after 5 cycles. This is consistent with the “loop and train” model for structure development in dough but it should be noted that the protein body preparation had not undergone the dehydration which occurs during the latter stage of grain maturation and could also affect protein interactions (as discussed above).

A full discussion of the effects of sample preparation, including dough mixing, on gluten structure is outside the scope of this article but they must be taken into account when interpreting datasets.

## 8. Conclusions

Despite research carried out over more than a century, and the vast volume of information on gluten protein sequences and polymorphism from modern genomic and proteomic studies, we still do not understand many of fundamental aspects of gluten protein synthesis, assembly and remodelling during grain maturation and food processing. This understanding is important as it will underpin future the wider use of gluten to meet the increasing demand for plant proteins.

The central questions relate to the relationships between the structure and assembly of the gluten protein network and the functional

properties of wheat dough and in particular the forces that stabilise the network. The central role of covalent disulphide bonds in stabilising the folded structures of gliadin monomers and the network of glutenin polymers is generally accepted but we still have limited knowledge of the precise sizes and structures of the polymers. Even less is known about the non-covalent forces responsible for the aggregation of these polymers and the binding of gliadin monomers to them. The important of hydrogen bonding has been suggested by theoretical studies and supported by analyses of HMW-GS and  $\omega$ -gliadins. However, classical analyses of gluten using acids, salts and alcohol-water mixtures indicate that hydrophobic and electrostatic forces are more important between the  $\alpha$ -type gliadins and  $\gamma$ -gliadins. Little is known about the interactions of the LMW-GS, which account for about a quarter of the gluten proteins, except that they form oligomers and co-polymers with HMW-GS. However, they are likely to interact by similar forces to the  $\alpha$ -type and  $\gamma$ -gliadins which they resemble. Elucidating the precise contributions of the forces which stabilise the gluten network will enable rational modification of gluten protein structure and functionality.

Finally, it must be borne in mind that the gluten complex is synthesised and assembled in the developing grain and it is therefore important to understand the cell biology of these processes and how they are affected by environmental conditions during grain development and maturation. Key questions relating to gluten structure, assembly and functionality are summarised in Table 4.

#### CRediT authorship contribution statement

**Peter R. Shewry:** Writing – review & editing, Writing – original draft. **Peter S. Belton:** Writing – review & editing, Writing – original draft.

#### Declaration of competing interest

The authors declare no conflicts of interest.

#### Data availability

No data was used for the research described in the article.

#### Acknowledgements

Rothamsted Research receives strategic funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and this work is part of the Delivering Sustainable Wheat (BB/X011003/1) Institute Strategic Programme.

#### References

- Ausennac, T., Rhazi, L., Branlard, G., 2020. Molecular weight distribution of polymeric proteins in wheat grains: the rheologically active polymers. *Foods* 9, 1675. <https://www.mdpi.com/2304-8158/9/12/1905>.
- Belton, P.S., 1995. On the elasticity of wheat gluten. *J. Cereal. Sci.* 29, 103–107. <https://doi.org/10.1006/jcsc.1998.0227>.
- Belton, P.S., 2005. New approaches to study the molecular basis of the mechanical properties of gluten. *J. Cereal. Sci.* 41, 203–211. <https://doi.org/10.1016/j.jcs.2004.06.003>.
- Belton, P.S., Colquhoun, I.J., Grant, A., Wellner, N., Field, J.M., Shewry, P.R., Tatham, A. S., 1995. FTIR and NMR studies on the hydration of a high-M<sub>r</sub> subunit of glutenin. *Int. J. Biol. Macromol.* 17, 74–80. [https://doi.org/10.1016/0141-8130\(95\)93520-8](https://doi.org/10.1016/0141-8130(95)93520-8).
- Boire, A., et al., 2013. Phase behaviour of a wheat protein isolate. *Soft Matter* 9 (47), 11417–11426. <https://doi.org/10.1039/c3sm51489g>.
- Branlard, G., Faye, A., Rhazi, L., Tahir, V., Aussenac, T., 2020. Genetic and environmental factors associated to glutenin polymer characteristics of wheat. *Foods* 9, 683. <https://doi.org/10.3390/foods9050683>.
- Branlard, G., d'Orlando, A., Tahir, A., Schmutz, M., Rhazi, L., Faye, A., Aussenac, T., 2023. The conformation of glutenin polymers in wheat grains: some genetic and environmental factors associated with this important characteristic. *J. Exp. Bot.* 74, 2653–2666. <https://doi.org/10.1093/jxb/erad013>.
- Bromilow, S., Gethings, L.A., Buckley, M., Bromley, M., Shewry, P.R., Langridge, J.I., Mills, E.N.C., 2017a. A curated gluten protein sequence database to support development of proteomics methods for determination of gluten in gluten-free foods. *J. Proteomics* 163, 67–75. <https://doi.org/10.1016/j.jpro.2017.03.026>.
- Bromilow, S.N.L., Gethings, L.A., Langridge, J.I., Shewry, P.R., Buckley, M., Bromley, M. L., Mills, E.N.C., 2017b. Comprehensive proteomic profiling of wheat gluten using a combination of data-independent and data-dependent acquisition. *Front. Plant Sci.* 7, 2020. <https://doi.org/10.3389/fpls.2016.02020>.
- Byers, M., Mifflin, B.J., Smith, S.J., 1983. A quantitative comparison of the extraction of protein fractions from wheat grain by different solvents, and of the polypeptide and amino acid composition of the alcohol-soluble proteins. *J. Sci. Food Agric.* 34, 447–462. <https://doi.org/10.1002/jsfa.2740340506>.
- Dahesh, M., Banc, A., Duri, A., Morel, M.-H., Ramos, L., 2014. Polymeric assembly of gluten proteins in an aqueous ethanol solvent. *J. Phys. Chem. B* 118, 11065–11076. <https://doi.org/10.1021/jp5047134>.
- Dill, D.B., Alsberg, C.L., 1925. Preparation, solubility, and specific rotation of wheat gliadin. *J. Biol. Chem.* 65, 279–304. [https://doi.org/10.1016/s0021-9258\(18\)84841-4](https://doi.org/10.1016/s0021-9258(18)84841-4).
- Geisslitz, S., Shewry, P.R., Brouns, F., America, A.H.P., Caio, G.P.I., Daly, M., D'Amico, S., De Giorgio, R., Gilissen, L., Grausgruber, H., Huang, X., Jonkers, D., Keszthelyi, D., Larre, C., Masci, S., Mills, C., Moller, M.S., Sorrels, M.E., Svensson, B., Zevallos, V., Weegels, P.L., 2021. Wheat ATIs: characteristics and role in human disease. *Front. Nutr.* 8, 667370. <https://doi.org/10.3389/fnut.2021.667370>.
- Georgopoulos, T., Larsson, H., Eliasson, A.C., 2006. Influence of native lipids on the rheological properties of wheat flour dough and gluten. *J. Texture Stud.* 37, 49–62. <https://doi.org/10.1111/j.1745-4603.2006.00038.x>.
- Gobaa, S., Bancel, E., Branlard, G., Kleijer, G., Stamp, P., 2008. Proteomic analysis of wheat recombinant inbred lines: variations in prolamin and dough rheology. *J. Cereal. Sci.* 47, 610–619. <https://doi.org/10.1016/j.jcs.2007.07.001>.
- Haward, S.J., Shewry, P.R., Marsh, J., Miles, M.J., McMaster, T.J., 2011. Force spectroscopy of an elastic peptide: effect of D<sub>2</sub>O and temperature on persistence length. *Microscope Research and Techniques* 74, 170–176. <https://doi.org/10.1002/jemt.20888>.
- Iwaki, S., Fu, B.X., Hayakawa, K., 2023. Behavior of protein aggregates via electrostatic interactions or hydrogen bonds during dough formation. *J. Cereal. Sci.* 111, 103683. <https://doi.org/10.1016/j.jcs.2023.103683>.
- Kobrehel, K., Alary, R., 1989. The role of a low-molecular weight glutenin fraction in the cooking quality of durum-wheat pasta. *J. Sci. Food Agric.* 47, 487–500.
- Koga, S., Bocker, U., Wieser, H., Koehler, P., Uhlen, A.K., Moldestad, A., 2017. Polymerisation of gluten proteins in developing wheat grain as affected by desiccation. *J. Cereal. Sci.* 73, 122–129. <https://doi.org/10.1016/j.jcs.2016.12.003>.
- Koga, S., Aamot, H.U., Uhlen, A.K., Seehusen, T., Veiseth-Kent, E., Hofgaard, I.S., Moldestad, A., Bocker, U., 2020. Environmental factors associated with glutenin polymer assembly during grain maturation. *J. Cereal. Sci.* 91, 108065. <https://doi.org/10.1016/j.jcs.2019.102865>.
- Lo Nostro, P., Ninham, B.W., 2012. Hofmeister phenomena: an update on ion specificity in biology. *Chem. Rev.* 112, 2286–2322. <https://doi.org/10.1021/cr200271j>.
- MacRitchie, F., 1987. Evaluation of contributions from wheat protein fractions to dough mixing and breadmaking. *J. Cereal. Sci.* 6, 259–268. [https://doi.org/10.1016/S0733-5210\(87\)80063-2](https://doi.org/10.1016/S0733-5210(87)80063-2).
- MacRitchie, F., Kasarda, D.D., Kuzmicky, D., 1991. Characterization of wheat protein fractions differing in contributions to breadmaking quality. *Cereal Chem.* 68, 122–130.
- Marion, D., Le Roux, C., Akoka, S., Tellier, S., Gallant, D., 1987. Lipid-protein interactions in wheat gluten: a phosphorus nuclear magnetic resonance spectroscopy and freeze-fracture electron microscopy study. *J. Cereal. Sci.* 5, 101–115. [https://doi.org/10.1016/S0733-5210\(87\)80014-0](https://doi.org/10.1016/S0733-5210(87)80014-0).
- Martin, R.B., 2001. Peptide bond characteristics. *Met. Ions Biol. Syst.* 38, 1–23.
- Morel, M.H., Pincemaille, J., Chauveau, E., Loughichi, A., Violleau, F., Menut, P., Ramos, L., Banc, A., 2020. Insight into gluten structure in a mild chaotropic solvent by asymmetrical flow field-flow fractionation (AsFFFF) and evidence of non-covalent assemblies between glutenin and  $\omega$ -gliadin. *Food Hydrocolloids* 103. <https://doi.org/10.1016/j.foodhyd.2020.105676>.
- Osipova, S.V., Permyakova, M.D., Permyakov, A.V., 2012. Role of non-prolamin proteins and low molecular weight redox agents in protein folding and polymerisation in wheat grains and influence on baking quality parameters. *J. Agric. Food Chem.* 60, 12065–12073. <https://doi.org/10.1021/jf303513m>.
- Perutz, M.F., Johnson, T., Suzuki, M., Finch, J.Y., 1994. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proceedings of the National Academy of Sciences USA* 91, 5355–5358. <https://doi.org/10.1073/pnas.91.12.5355>.
- Pézolet, M., Bonenfant, S., Dousseau, F., Popineau, Y., 1992. Conformation of wheat gluten proteins: Comparison between functional and solution states as determined by infrared spectroscopy. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 299, 247–250. [https://doi.org/10.1016/0014-5793\(92\)80125-Z](https://doi.org/10.1016/0014-5793(92)80125-Z).
- Preston, K.R., 1981. Effects of neutral salts upon wheat gluten protein properties. I. relationship between the hydrophobic properties of gluten proteins and their extractability and turbidity in neutral salts. *Cereal Chem.* 58, 317–324.
- Preston, K.R., 1989. Effects of neutral salts of the lyotropic series on the physical dough properties of a Canadian Red Spring Wheat flour. *Cereal Chem.* 66, 144–148.
- Rembert, K.B., Paterova, J., Heyda, J., Hilty, C., Jungwirth, P., Cremer, P.S., 2012. Molecular mechanisms of ion-specific effects on proteins. *J. Am. Chem. Soc.* 134, 10039–10046. <https://doi.org/10.1021/ja301297g>.
- Scherf, K., 2023. Gluten proteins. In: Shewry, P.R., Koksel, H., Taylor, J.R.N. (Eds.), *ICC Handbook of 21<sup>st</sup> Century Cereal Science and Technology*. Elsevier Academic Press, London, UK, pp. pp71–78.
- Shewry, P.R., 2019. What is gluten? Why is it special? *Front. Nutr.* <https://doi.org/10.3389/fnut.2019.00101>.

- Shewry, P.R., Lafiandra, D., 2022. Wheat glutenin polymers 1. Structure, assembly and properties. *J. Cereal. Sci.* 106, 104386 <https://doi.org/10.1016/j.jcs.2022.103486>.
- Shewry, P.R., Halford, N.G., Tatham, A.S., Popineau, Y., Lafiandra, D., Belton, P.S., 2003. The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. *Adv. Food Nutr. Res.* 45, 221–302.
- Shewry, P.R., D'Ovidio, R., Lafiandra, D., Jenkins, J.A., Mills, E.N.C., Bekes, F., 2009a. Wheat grain proteins. In: Khan, K., Shewry, P.R. (Eds.), *Wheat: Chemistry and Technology*, fourth ed. AACC, St Paul Mn, pp. 223–298.
- Shewry, P.R., Underwood, C., Wan, Y., Lovegrove, A., Bhandari, D., Toole, G., Mills, E.N.C., Denyer, K., Mitchell, R.A.C., 2009b. Storage product synthesis and accumulation in developing grains of wheat. *J. Cereal. Sci.* 50, 106–112. <https://doi.org/10.1016/j.jcs.2009.03.009>.
- Tatham, A.S., Thomson, N.H., McMaster, T.J., Humphris, A.D.L., Miles, M.J., Shewry, P.R., 1999. SPM studies of cereal seed storage protein structures. *Scanning* 21, 293–298. <https://doi.org/10.1002/sca.4950210502>.
- Tatham, A.S., Hayes, L., Shewry, P.R., Urry, D.W., 2001. Wheat proteins exhibit a complex mechanism of protein elasticity. *Biochimica et Biophysica Acta* 1548, 187–193. [https://doi.org/10.1016/S0167-4838\(01\)00232-1](https://doi.org/10.1016/S0167-4838(01)00232-1).
- Tosi, P., 2021. Trafficking and deposition of prolamins in wheat. *J. Cereal. Sci.* 56, 81–90. <https://doi.org/10.1016/j.jcs.2012.02.004>.
- Tuhumury, H.C.D., Small, D.M., Day, L., 2014. The effect of sodium chloride on gluten network formation and rheology. *J. Cereal. Sci.* 60, 229–237. <https://doi.org/10.1016/j.jcs.2014.03.004>.
- Tuhumury, H.C.D., Small, D.M., Day, L., 2016. Effects of Hofmeister salt series on gluten network formation: Part II. Anion series. *Food Chem.* 212, 798–806. <https://doi.org/10.1016/j.foodchem.2016.06.015>.
- Wakisaka, A., Komatsu, S., Usui, Y., 2001. Solute-solvent and solvent-solvent interactions evaluated through clusters isolated from solutions: preferential solvation in water-alcohol mixtures. *J. Mol. Liq.* 90 (1–3), 175–184. [https://doi.org/10.1016/S0167-7322\(01\)00120-9](https://doi.org/10.1016/S0167-7322(01)00120-9).
- Wellner, N., Belton, P.S., Tatham, A.S., 1996. Fourier transform IR spectroscopic study of hydration induced structure changes in the solid state of  $\Omega$ -gliadins. *Biochem. J.* 319, 741–747. <https://doi.org/10.1042/bj3190741>.
- Wellner, N., Bianchini, D., Mills, E.N.C., Belton, P.S., 2003. Effect of selected Hofmeister anions on the secondary structure and dynamics of wheat prolamins in gluten. *Cereal Chem.* 80, 596–600. <https://doi.org/10.1094/CCEM.2003.80.5.596>.
- Wellner, N., Mills, E.N.C., Brownsey, G., Wilson, R.H., Brown, N., Freeman, J., Halford, N.G., Shewry, P.R., Belton, P.S., 2005. Changes in protein secondary structure during gluten deformation studied by dynamic Fourier-transform infrared spectroscopy. *Biomacromolecules* 2005, 255–261. <https://doi.org/10.1021/bm049584d>.
- Wellner, N., Marsh, J., Savage, A., Halford, N.G., Shewry, P.R., Mills, E.N.C., Belton, P.S., 2006. Structure-function relationships in an elastomeric protein of wheat: synthesis, expression and characterisation of repetitive peptides based on sequence variants present in the HMW subunits. *Biomacromolecules* 7, 1096–1103. [https://doi.org/10.1016/S0167-4838\(01\)00155-8](https://doi.org/10.1016/S0167-4838(01)00155-8). <https://doi.org/10.1021/bm050893t>.
- Xiang, L., Etcheberria, E., Van den Ende, W., 2013. Vacuolar protein sorting mechanisms in plants. *FEBS J.* 280, 979–993. <https://doi.org/10.1111/febs.12092>.
- Yu, J.T., Li, X.-Y., Huang, J.-H., Yu, M.-Y., Wu, Z.-Y., Cao, S.-L., 2023. Molecular dynamics simulation of  $\alpha$ -gliadin in ethanol/aqueous organic solvents. *Int. J. Food Sci. Technol.* 58, e1–e9. <https://doi.org/10.1111/ijfs.15754>.
- Zhang, Y., Cremer, P.S., 2006. Interactions between macromolecules and ions: the Hofmeister series. *Curr. Opin. Chem. Biol.* 10, 658–663. <https://doi.org/10.1016/j.cbpa.2006.09.020>.
- Zhang, Y., Furry, S., Bergbreiter, D.E., Cremer, P.S., 2005. Specific ion effects on the water solubility of macromolecules: PNIPAM and the Hofmeister series. *J. Am. Chem. Soc.* 127, 14505–14510. <https://doi.org/10.1021/ja0546424>.