The structure of a high- M_r subunit of durum-wheat (*Triticum durum*) gluten

J. Michael FIELD,* Arthur S. TATHAM† and Peter R. SHEWRY†‡

*M.M.G. Agriseed, Docking, King's Lynn, Norfolk PE31 8LA, and †Biochemistry Department, Institute of Arable Crops Research, Rothamsted Experimental Station, Harpenden, Herts. AL5 2JQ, U.K.

A high- M_r subunit was prepared from durum wheat (*Triticum durum*). Viscometric analysis showed that the molecule is rod-shaped, with molecular dimensions of about 50 nm $\times$ 1.75 nm (500 Å $\times$ 17.5 Å) in 0.05 M-acetic acid/0.01 M-glycine and 49 nm $\times$ 1.79 nm (490 Å $\times$ 17.9 Å) in aq. 50 % (v/v) propan-1-ol ($\pm$ 0.01 M-glycine) at 30 °C. C.d. spectroscopy in the same solvents indicated the presence of β -turns, but little α -helix [7 % in 50 % (v/v) propan-1-ol] and no β -sheet. However, when dissolved in trifluoroethanol the protein contains about 30 % α -helix, and viscometric analysis gives dimensions of about 62 nm $\times$ 1.53 nm (620 Å $\times$ 15.3 Å). It is proposed, on the basis of these studies and previously published structural prediction, that the repetitive central domain of the high- M_r subunit forms a loose spiral based on repetitive β -turns, whereas the shorter non-repetitive *N*- and *C*-terminal domains are α -helical in trifluoroethanol, but random coil in other solvents. The M_r of the high- M_r subunit determined from the intrinsic viscosity in 6.0 M-guanidinium chloride was 65000, compared with 84000 determined in 5.0 M-guanidinium thiocyanate. The latter value is consistent with the M_r values for related proteins whose complete amino acid sequences are known, and it was concluded that the protein is incompletely denatured in the former solvent. This was confirmed by c.d. spectroscopy in increasing concentrations (1–6 M) of guanidinium chloride.

INTRODUCTION

Wheat gluten is the insoluble mass that remains when dough is washed to remove starch and components (including proteins) that are soluble in water and dilute solutions of salts. Its physical properties, a combination of elasticity and viscous flow (extensibility), are probably unique among materials of biological origin. Although the precise molecular basis of these properties is unknown, it is generally agreed that they are determined mainly by the storage proteins that account for about 70–80 % of gluten on a dry-weight basis [1]. Gluten is largely responsible for the ability of wheat flour to be baked into leavened bread, the proteinaceous matrix being expanded by trapped fermentation gases and finally fixed by baking to give a porous crumb structure. In addition, an increasing amount of gluten is being produced industrially for other uses in the food industry. Because of this, fundamental studies of gluten protein structure are likely to have commercial relevance as well as academic value.

One group of wheat gluten proteins has been intensively studied over the past few years because genetic variation in its component polypeptides has been shown to be closely associated with variation in breadmaking quality [2]. These proteins have been called the 'high molecular weight' ('HMW'), hereafter high- M_r , subunits of glutenin, and genotypes of hexaploid bread wheat (*Triticum aestivum*) contain from three to five individual subunits encoded by genes on the three homoeologous (i.e. homologous but non-pairing) genomes (called the A, B and D genomes) [2]. A number of individual subunits

have been purified and partially characterized [3], and further amino acid sequences have been deduced from the nucleotide sequences of partial cDNAs [4,5] and complete genes [6–9]. Preliminary studies of the secondary structure of a mixed preparation of high- M_r subunits have also been made using a combination of c.d. spectroscopy and computer prediction [10,11]. These indicated an unusual and intriguing structure. The amino acid sequence of each subunit can be divided into three clear structural domains. The central domain, which varies in length from 435 to 687 residues in the proteins encoded by the four sequenced genes, consists of tandem and interspersed repeats of nonapeptide [consensus sequence Gly-Tyr-Tyr-Pro-Thr-Ser-Pro (or Leu)-Gln-Gln] and hexapeptide (consensus sequence Pro-Gly-Gln-Gly-Gln-Gln) motifs. The precise organization of this domain has been discussed in detail by Halford *et al.* [9]. It is flanked at the *N*- and *C*-termini by non-repetitive regions of about 100 and 40 residues respectively. These are richer in charged residues than the repeat domain and contain most, or all, of the four to seven cysteine residues that are present in each subunit. Computer prediction and c.d. studies also indicate that the repetitive domain has an unusual secondary structure, with regularly repeated β -turns. We have speculated that these could form a regular helical structure, similar to the β -spiral formed by the synthetic polypentapeptide of elastin [10,11], and that this structure contributes to the elastic mechanism of whole gluten. Although computer predictions indicate that the non-repetitive *N*- and *C*-terminal regions are predominantly α -helical, this is not supported by c.d. analysis in an ethanol/trifluoroethanol

Abbreviation used: PAGE, polyacrylamide-gel electrophoresis.

‡ To whom correspondence and reprint requests should be sent.

solvent [10]. The precise roles of the cysteine residues are also unknown, although the observation that high- M_r subunits are present only in high-molecular-mass (above 1 MDa) polymers that dissociate in the presence of reducing agents [12] indicates that at least some must form intermolecular disulphide bonds.

We present here the first detailed conformational study of an individual high- M_r subunit. Results from hydrodynamic and spectroscopic studies are presented which are compatible with our hypothesis that the repetitive domain forms a regular helical structure based on repetitive β -turns, whereas the *N*- and *C*-termini are α -helical in some solvent systems.

MATERIALS AND METHODS

Protein purification and characterization

A high- M_r subunit from the French durum-wheat (*Triticum durum*) cultivar Bidi 17 was isolated by the procedure of Danno *et al.* [13]. Endosperm flour was prepared on a Brabander Quadrumat Senior laboratory mill, defatted with water-saturated butan-1-ol and extracted with 0.01 M-sodium phosphate buffer, pH 7.0, containing 0.5% (w/v) SDS to remove albumins, globulins and gliadins. Glutenins were then extracted with 0.5% (w/v) SDS containing 1% (v/v) 2-mercapto-ethanol and fractionated, without alkylation, by differential precipitation. The pH of the glutenin extract was adjusted to 4.0, ethanol was added to 70% (v/v) and the pH was raised to 5.2; this caused a mixed precipitate of high- and low- M_r subunits to form. This was removed by centrifugation and the pH of the supernatant increased to 6.7, resulting in the precipitation of almost pure high- M_r subunit. This precipitate was collected by centrifugation and dissolved in 0.05 M-Tris/acetate buffer, pH 7.8, containing 8 M-urea and 1% (v/v) 2-mercapto-ethanol before removal of SDS with Dowex AG1-X4 resin followed by alkylation with 4-vinylpyridine [14].

SDS/PAGE was carried out by using a gel system with Tris/borate running and separating gel buffers [15]. The amino acid composition was determined with an LKB 4400 amino acid analyser.

Viscometric analysis

Viscosity measurements were made in an Ostwald capillary viscometer [constant $\approx 3 \times 10^{-9} \text{ m}^2 \cdot \text{s}^{-2}$ (0.003 cSt/s)] on solutions of the pyridethylated high- M_r subunit in 50% (v/v) propan-1-ol, 50% (v/v) propan-1-ol/10 mM-glycine, 50 mM-acetic acid/10 mM-glycine, trifluoroethanol, 6.0 M-guanidinium chloride/0.02 M-sodium acetate, pH 5.0, and 5.0 M-guanidinium thiocyanate/0.02 M-sodium acetate, pH 5.5. For purposes of comparison, viscosity measurements were also carried out on solutions of 'C' hordein from the mutant barley (*Hordeum vulgare*) line Risø 56 [16] in trifluoroethanol and guanidinium thiocyanate. Viscosity determinations in guanidinium chloride were carried out at 25 °C; all other determinations were made at 30 °C. Density measurements were made using a graduated bicapillary pycnometer. Protein concentrations were determined spectrophotometrically. Absorbances at 280 nm were corrected for light-scattering by a straight-line extrapolation through the absorbances at 360 nm and 320 nm to the shorter wavelength [17]. Absorption coefficients at 280 nm were determined by using protein solutions

whose concentrations were determined by Kjeldahl nitrogen analysis (using a factor of 5.7 for the conversion of nitrogen to protein). The partial specific volume of the high- M_r subunit was calculated from compositional data (Table 1) [18].

C.d. spectroscopy

C.d. measurements were made with a Jasco J40CS dicograph. The temperature used was 30 ± 1 °C, maintained by a heated cell holder and a thermocouple of a Comark electronic thermometer inserted directly into solution. Results are expressed as mean residue ellipticities with the units degrees $\cdot \text{cm}^2 \cdot \text{dmol}^{-1}$.

RESULTS AND DISCUSSION

Purification and characterization of a high- M_r subunit

Detailed physico-chemical studies of single isolated high- M_r subunits have not been reported previously. This is at least partly due to the difficulty of preparing adequate amounts of individual subunits from the mixture of three to five present in genotypes of hexaploid bread wheat. To overcome this problem we identified a line (Bidi 17) of the related pasta wheat, *T. durum*, that contained only one major high- M_r subunit (Fig. 1, track a). This species shares the A and B genomes with *T. aestivum*, and the single high- M_r subunit of Bidi 17 had a similar mobility on SDS/PAGE to the 1B-encoded subunit (called 'subunit 7' by Payne *et al.* [20]) of the *T. aestivum* cultivar Broom (Fig. 1, track c).

SDS/PAGE of the purified subunit preparation (Fig. 1, track b) showed little contamination with other proteins. The amino acid composition (Table 1) was similar to those of other high- M_r subunits [3] with

Fig. 1. SDS/PAGE of total gluten proteins from *T. durum* cv. Bidi 17 (a), the purified high- M_r subunit from Bidi 17 (b), and a high- M_r subunit preparation from *T. aestivum* cv. Broom (c)

The four major high- M_r subunit bands (bracket in track c) present in cv. Broom are (in order of increasing mobility): subunits 5 (chromosome 1D-encoded), 7 and 9 (both chromosome 1B-encoded) and 10 (chromosome 1D-encoded).

Table 1. Amino acid composition (expressed as mol %) of the high- M_r subunit

Results are the mean of duplicate hydrolyses and determinations.

Asp*	1.33
Thr	3.20
Ser	7.11
Glu*	37.24
Pro	10.62
Gly	16.66
Ala	3.49
Cys†	1.50
Val	1.98
Met	0.30
Ile	1.30
Leu	3.79
Tyr	6.37
Phe	0.96
His	0.75
Lys	0.96
Arg	2.44

* Include the amides asparagine and glutamine respectively.

† Determined as cysteic acid.

high glutamate + glutamine (37 mol%) and glycine (17 mol%). Determination of the free carboxy-group content of a purified subunit by carbodi-imide modification [3] and deduction of amino acid compositions from the nucleotide sequences of cloned genes (see refs. [6–9]) show that most of the glutamate recovered after hydrolysis is present as glutamine.

It is not possible to determine accurate M_r values for high- M_r subunits by SDS/PAGE using globular proteins as standards [15]. Because of this we determined an approx. M_r for the Bidi 17 subunit by comparison of its mobility on a Tris/borate SDS/PAGE system [15] with those of three subunits whose true M_r values have been calculated from deduced amino acid sequences. These

are subunits 2 (M_r 87000), 9 (M_r 73500) and 12 (M_r 68700) [7–9]. The calculated M_r of 82000 agrees well with that determined by viscometric analysis under denaturing conditions (see below). This M_r was used for calculations in the viscometric study.

Viscometric analysis

The reduced viscosities (η_{sp}/c , where η_{sp} is specific viscosity and c is concentration) of the *S*-pyridylethylated subunit in several solvents are shown as a function of concentration in Fig. 2(b), and the intrinsic viscosities $[\eta]$ determined by extrapolation to zero concentration are given in Table 2. The values of $[\eta]$ in all solvents are considerably above those expected for compact globular molecules (3.5–4.0 ml/g [19]), indicating that the protein exists either as highly solvated random-coiled chains or as relatively rigid asymmetric particles. In the case of the solutions in acetic acid, the extent of polyelectrolyte chain expansion due to a polyelectrolyte effect, which at low pH is likely to be pronounced even in the presence of 10 mM-glycine acting to shield charged groups, is unknown. However, the polyelectrolyte effect would be expected to be minimal in aq. propan-1-ol and, in fact, the addition of 10 mM-glycine had essentially no effect on the $[\eta]$ measured in aq. 50% (v/v) propan-1-ol. Although it is usual to include salts to reduce the polyelectrolyte effect, this proved impossible in the present study as it resulted in precipitation of the protein.

C.d. spectroscopy of the high- M_r subunit dissolved in dilute acetic acid, aq. propan-1-ol or trifluoroethanol (see below) indicated that it has an ordered conformation. We therefore calculated its size and shape on the assumption that it exists as an asymmetric hydrodynamic particle rather than in the form of a random coil. The Simha shape factor [21] was calculated (assuming an effective solvation of 0.2 g/g of protein [19]) and used to determine the axial ratios (a/b) of the prolate ellipsoids best representing the real molecules in the various solvents [19,21,22]. Values were then assigned to the

Fig. 2. Reduced viscosity (η_{sp}/c) of C hordein of barley (a) and the high- M_r subunit of wheat (b) in a range of solvents

Key to symbols: ■, trifluoroethanol (30 °C); ○, 5.0 M-guanidinium thiocyanate/0.02 M-sodium acetate, pH 5.0 (30 °C); ●, 5.9 M-guanidinium chloride (25 °C) in (a) and 6.0 M-guanidinium chloride/0.02 M-sodium acetate, pH 5.0 (25 °C) in (b); ▲, aq. 70% (v/v) ethanol/0.01 M-NaCl (30 °C); △, 0.1 M-acetic acid (30 °C); □, 0.05 M-acetic acid/0.01 M-glycine (30 °C); ▽, aq. 50% (v/v) propan-1-ol/0.01 M-glycine (30 °C); ▼, aq. 50% (v/v) propan-1-ol (30 °C).

semi-axes (a and b) and to the length (L) and diameter (d) of the equivalent rod-shaped molecules [16,19].

The values for a/b , L and d are given in Table 2. They show that the molecule in 50% propan-1-ol (± 10 mM-glycine) is rod-like, with dimensions of about $49 \text{ nm} \times 1.8 \text{ nm}$ ($490 \text{ \AA} \times 18 \text{ \AA}$). There is a small increase in the axial ratio in 50 mM-acetic acid/10 mM-glycine, and a much greater increase in trifluoroethanol, where L is over 60 nm ($600 \text{ \AA}$).

As previously described for C hordein [16], the calculated axial ratio of the high- M_r subunit in aq. propan-1-ol can be used to test the initial assumption of an asymmetric hydrodynamic particle by means of the shape function $K_s/[\eta]$ [23,24] (K_s being the coefficient of concentration dependence of the reciprocal sedimentation coefficient). Comparison of the experimentally determined value of 16.4 ml/g for K_s (analogous to K_s) with the calculated values for K_s of 72 and 18 ml/g for random coil and asymmetric particle respectively clearly rules out the random-coil model.

The conformation of the barley storage protein C hordein is considered to be similar to that of the repetitive domain of the high- M_r subunits [16]. Results for this protein are included in Fig. 2(a) and Table 2. Being of lower M_r (52000), the overall dimensions of C hordein are smaller, but it too is rod-like, with a diameter of $1.7\text{--}1.9 \text{ nm}$ ($17\text{--}19 \text{ \AA}$). It is also more extended in trifluoroethanol than in the other solvents, the magnitude of the effect being about equal to that observed with the high- M_r subunit. However, the two proteins differ in their relative dimensions in aq. alcohol and dilute acetic acid, the high- M_r subunit being more extended in dilute acetic acid, whereas C hordein is more compact.

Where the protein is present as a highly solvated random coil, the M_r can be calculated from the intrinsic viscosity [25]. The M_r calculated in this way from the intrinsic viscosity of the high- M_r subunit in 6.0 M-guanidinium chloride (dissolved in 20 mM-sodium acetate, pH 5.5) was only 65000, much lower than would

be expected from the M_r values deduced from the nucleotide sequences of genes encoding other subunits [6–9]. When the determination was repeated with a stronger denaturant, 5 M-guanidinium thiocyanate, the intrinsic viscosity increased from 48.9 to 57.7 ml/g , and the calculated M_r to 84000. This agrees well with the M_r calculated on the basis of its electrophoretic mobility (see above), and is probably close to its true M_r . The discrepancy between the results obtained with guanidinium chloride and guanidinium thiocyanate indicates that the protein was incompletely denatured by the former solvent. This contrasts with C hordein, which had a similar intrinsic viscosity in the two denaturants (Table 2).

C.d. spectroscopy and structure prediction

The conformation of the high- M_r subunit, in a range of solvents, was studied by c.d. spectroscopy. Whereas the spectrum in the far-u.v. region (below 250 nm) arises principally from the peptide bond and is indicative of the backbone structure, the spectrum in the near-u.v. ($250\text{--}320 \text{ nm}$) arises from aromatic and disulphide chromophores.

The spectra were initially determined in 0.05 M-acetic acid ($\pm 0.01 \text{ M-glycine}$), aq. 50% (v/v) propan-1-ol ($\pm 0.01 \text{ M-glycine}$), 6 M-guanidinium chloride and trifluoroethanol at 30°C . The presence of glycine had no effect on the spectra in either solvent (spectra not shown). Also, although it was not possible to measure the far-u.v. c.d. spectrum in 0.05 M-acetic acid below 220 nm because of the absorbance of the acetate ion, the far-u.v. spectrum above this wavelength and the near-u.v. spectrum were essentially the same as those determined in aq. 50% (v/v) propan-1-ol. The spectra in 0.05 M-acetic acid are therefore not shown.

The c.d. spectra in aq. 50% (v/v) propan-1-ol are shown in Fig. 3. The far-u.v. spectrum has a minimum around 205 nm and pronounced shoulder around $220\text{--}230 \text{ nm}$. It differs from the spectrum in 6 M-guanidinium

Table 2. The intrinsic viscosities, calculated axial ratios and molecular dimensions of the high- M_r subunit of wheat and C hordein of barley

	Temperature (°C)	[η] (ml/g)	Axial ratio (a/b)	Molecular dimensions	
				Length (Å)	Diameter (Å)
High- M_r subunit					
0.05 M-acetic acid/0.01 M-glycine	30	48.4	23.4	504	17.5
50 % (v/v) aq. propan-1-ol	30	44.7	22.4	492	17.9
50 % (v/v) aq. propan-1-ol/ 0.01 M-glycine	30	44.8	22.4	492	17.9
Trifluoroethanol	30	75.1	32.9	618	15.3
6.0 M-guanidinium chloride/ 0.02 M-sodium acetate (pH 5.0)	25	48.9	—	—	—
5.0 M-guanidinium thiocyanate/ 0.02 M-sodium acetate (pH 5.5)	30	57.7	—	—	—
C hordein					
0.1 M-acetic acid*	30	16.8	12.1	282	19.1
70 % (v/v) aq. ethanol/0.01 M-sodium chloride	30	26.8	16.2	346	17.4
Trifluoroethanol	30	44.2	23.4	429	15.0
5.9 M-guanidinium chloride*	25	40.8	—	—	—
5.0 M-guanidinium thiocyanate/ 0.02 M-sodium acetate (pH 5.5)	30	41.4	—	—	—

* Data taken from ref. 16

chloride, which approximates to that typical of a random-coil structure [26]. This indicates that the protein has a regular conformation in aq. 50% (v/v) propan-1-ol, and this is confirmed by the presence in the near-u.v. c.d. spectrum of absorbances due to phenylalanine (256, 262 and 268 nm) and tyrosine (275 and 284 nm) (Fig. 3b) [27]. The c.d. spectrum of the aromatic side chains can indicate subtle changes in their environment; some absorbance was observed in the near-u.v. spectrum of the protein in 6 M-guanidinium chloride (Fig. 3b), suggesting that denaturation was incomplete.

The far-u.v. c.d. spectra of the protein in 0.05 M-acetic acid and aq. 50% (v/v) propan-1-ol were similar to those reported previously for a mixture of high- M_r subunits dissolved in 7:3 (v/v) ethanol/trifluoroethanol [10], and to C hordein in aq. 70% (v/v) ethanol, trifluoroethanol and dilute acetic acid [16,28]. The latter has been interpreted as indicative of a regular structure rich in β -turns, but with no contribution from the α -helix or β -sheet conformations. This is supported by computer predictions based on the amino acid sequences of the repetitive domains of these proteins [10,11,28].

When the high- M_r subunit was dissolved in trifluoroethanol, the far-u.v. spectrum was typical of that of an α -helical-rich protein, with minima around 208 and 221 nm and a cross-over point above 200 nm (Fig. 3a) [29]. The near-u.v. spectrum showed absorbances associated with tyrosine (Fig. 3b), with no evidence of phenylalanine band structure [27].

The spectra observed in 0.05 M-acetic acid, aq. 50% (v/v) propan-1-ol and trifluoroethanol indicate that the

protein can exist in two distinct stable conformations. The transition between these was studied by determining the spectra in 1:3, 1:1 and 3:1 (v/v) mixtures of aq. 50% (v/v) propan-1-ol and trifluoroethanol (Fig. 3a). Decreasing the concentration of trifluoroethanol resulted in a diminished intensity of the α -helical characteristics, with the greatest diminution occurring between 100 and 75% trifluoroethanol. An isocircular-dichroic point was observed around 204 nm, indicating that the conformational change that was occurring involved only the two forms, with a negligible concentration of intermediates.

Although the far-u.v. c.d. spectrum of the high- M_r subunit dissolved in 6 M-guanidinium chloride is consistent with a random-coil structure (Fig. 3a), the viscometric data and the near-u.v. c.d. spectrum (Fig. 3b) indicate that the protein is not completely denatured. To study the stability in chaotropic salts the spectra were determined in 1–6 M-guanidinium chloride; changes in the far-u.v. spectra indicated a progressive denaturation, with no indication that the process had reached completion at 6 M-guanidinium chloride (Fig. 4). It was, unfortunately, not possible to determine the far-u.v. spectrum of the protein in guanidinium thiocyanate, owing to the absorbance of the thiocyanate ion. These results indicate that the β -turn-rich conformation is unusually stable, and may be relevant to its role in gluten structure.

Structure prediction on the available amino acid sequences indicates that the non-repetitive *N*- and *C*-terminal domains of the high- M_r subunits are predominantly α -helical [10,11]. These domains comprise some 140–150 residues and represent 19–25% of the protein sequences. In aq. 50% (v/v) propan-1-ol these regions would not appear to be predominantly in the α -helical conformation, whereas in trifluoroethanol an α -helical-rich conformation is adopted. Arakawa & Godette [30] reported that certain organic solvents increase the α -helical content of proteins when added in high concentrations, the decreased dielectric constant of the solvent increasing protein structure contact [31]. The transition to α -helix may be preceded by protein unfolding [32], and the helical structures formed may

Fig. 3. Far-u.v. (a) and near-u.v. (b) c.d. spectra of the high- M_r subunit

Solvents were: (i) 6 M-guanidinium chloride; (ii) aq. 50% (v/v) propan-1-ol; (iii), (iv) and (v), 1:3, 1:1 and 3:1 (v/v) mixtures of aq. 50% (v/v) propan-1-ol and trifluoroethanol respectively, and (vi) trifluoroethanol. The temperature was $30 \pm 1^\circ\text{C}$, and the protein concentration was 1.0 mg/ml.

Fig. 4. Variation in $[\theta]_{222}$ for the high- M_r subunit in increasing concentrations (1–6 M) of guanidinium chloride

The temperature was $30 \pm 1^\circ\text{C}$ and the protein concentration 1.0 mg/ml.

include regions that are not α -helical in the native protein. The alkylated cysteine residues in the high- M_r subunit would not restrict formation of helical segments in the *N*-terminal domain. It is unlikely that the repetitive central domain could be induced to form an α -helix, owing to the high content of proline in the repeat motifs.

The intensities of the c.d. spectra of α -helices are markedly chain-length-dependent, the intensity decreasing with decreasing chain length [33]. This may account for some of the observed differences between the spectra in aq. 50% (v/v) propan-1-ol and trifluoroethanol; the minima associated with the α -helical spectrum are preserved, but red-shifted. Also, although structure prediction indicates that the central repeated domains are rich in β -turns, this spectrum would be obscured in trifluoroethanol, since the intensity of the α -helical spectrum is greater than those of the various β -turns types. The decreased α -helix content in aq. 50% (v/v) propan-1-ol would allow the β -turn spectra to be observed. Deconvolution of the spectra by the method of Chen *et al.* [34] yields a value of approx. 30% α -helix in 100% trifluoroethanol and 7% in aq. 50% (v/v) propan-1-ol, suggesting that, in 100% trifluoroethanol, the *N*- and *C*-termini are highly structured.

Comparison with C hordein of barley is again of interest. C hordein has only small non-repetitive domains (probably only 18 residues), which do not appear to have the potential to form α -helices. Whereas the c.d. spectra of C hordein in aqueous alcohols are similar to that of the high- M_r subunit in aq. 50% (v/v) propan-1-ol, the spectrum in trifluoroethanol indicates a second type of conformation that is also rich in β -turns [28]. The transition between these two conformations probably involves a change in the types (or proportions) of β -turns present and may result from the effect of the dielectric constant on the relative intensities of hydrophobic interactions and hydrogen-bonding [26]. A similar conformational transition in the repetitive domain of the high- M_r subunit would not be detected by c.d. spectroscopy because of the increasing contribution of the α -helical conformation to the spectrum.

General discussion

We have proposed a model for the molecular structure of C hordein of barley on the basis of similar studies (hydrodynamic, spectroscopic and predictive) as those reported here [16,28]. It was suggested that the repetitive primary structure results in regularly repeated β -turns and that these form a loose spiral similar to the β -spiral formed by the synthetic polypentapeptide of elastin [35]. Calculations based on the molecular dimensions of the β -spiral give overall molecular dimensions that are consistent with those determined viscometrically. This model is readily applied to C hordein because of the absence of extensive non-repetitive sequences and the presence of tandem repeats of a single octapeptide motif (consensus sequence Pro-Gln-Gln-Pro-Phe-Pro-Gln-Gln) with some related pentapeptides (consensus sequence Pro-Gln-Gln-Pro-Tyr). Although the hexapeptide and nonapeptide motifs present in the repetitive domains of high- M_r subunits have no sequence homology with those present in C hordein, the results reported here indicate that they may form a similar spiral conformation.

Assuming that the Bidi 17 subunit has an M_r of 82000 and similar-length non-repetitive domains to the four

subunits whose complete amino acid sequences are known, its repetitive domain should consist of about 620 residues. If these form a loose spiral with similar dimensions to the β -spiral formed by the synthetic polypentapeptide of elastin, the overall molecular dimensions of the domain will be about $44 \text{ nm} \times 1.8 \text{ nm}$ ($440 \text{ \AA} \times 18 \text{ \AA}$). This is slightly less than those determined hydrodynamically in aq. 50% (v/v) propan-1-ol or 0.05 M-acetic acid, and suggests that the non-repetitive domains at the *N*- and *C*-termini (which are unstructured in these solvents) may also contribute to the measured dimensions.

Two factors may contribute to the changed molecular dimensions of the protein in trifluoroethanol. Firstly, the decreased dielectric constant may result in changes in the proportions of different β -turns types in the repetitive domain, as described previously for C hordein [28]. These may well result in changes in the axial ratio of the β -spiral region. Secondly, the *N*- and *C*-terminal domains are highly structured (α -helical) in this solvent and may make a greater contribution to the dimensions measured hydrodynamically.

The relevance of the conformations determined for the high- M_r subunit in solution to those adopted *in vivo* and in dough is not known. The high- M_r subunits are deposited in the developing seed, together with other gluten proteins, as an insoluble mass in membrane-bound organelles called protein bodies [36]. N.m.r. studies of C hordein of barley [37] indicate that the protein has a regular conformation in the dry solid state, and we assume that the same would apply to the high- M_r subunits present in the hydrated state in protein bodies and in dough. The dielectric constants of the media in protein bodies and dough are likely to be closer to that of 50% (v/v) aq. propan-1-ol than to that of trifluoroethanol, and we therefore assume that the conformations of the high- M_r subunits are closest to that determined in the former solvent.

Because the high- M_r subunit is a three-domain protein, it is not as easy or reliable to interpret simple viscometric data in terms of overall structure as it is for a simple one-domain protein such as C hordein. Nevertheless we believe that the model for high- M_r subunit structure outlined above is consistent with all the available information on these proteins. It has important implications for the structure and functionality of wheat gluten (as discussed in ref. [38]) and provides a sound basis for further experimental studies.

We thank Dr. A. F. Drake, Birkbeck College, London, for the provision of c.d. facilities, Mr. R. P. White for statistical analyses and Mrs. S. Smith for amino acid analyses.

REFERENCES

1. Mifflin, B. J., Field, J. M. & Shewry, P. R. (1983) in *Seed Proteins* (Daussant, J., Mossé, J. & Vaughan, J., eds.), pp. 255–335, Academic Press, London.
2. Payne, P. I., Holt, L. M., Jackson, E. A. & Law, C. N. (1984) *Philos. Trans. R. Soc. London Ser. B*, **304**, 359–371.
3. Shewry, P. R., Field, J. M., Faulks, A. J., Parmar, S., Mifflin, B. J., Dietler, M. D., Lew, E. J.-L. & Kasarda, D. D. (1984) *Biochim. Biophys. Acta* **788**, 23–24.
4. Forde, J., Forde, B. G., Fry, R., Kreis, M., Shewry, P. R. & Mifflin, B. J. (1983) *FEBS Lett.* **162**, 360–366.

5. Thompson, R. D., Bartels, D., Barberd, N. P. & Flavell, R. B. (1983) *Theor. Appl. Genet.* **67**, 87–96
6. Forde, J., Malpica, J.-M., Halford, N., Shewry, P. R., Anderson, O. D., Greene, F. C. & Mifflin, B. J. (1985) *Nucleic Acids Res.* **13**, 6817–6832
7. Thompson, R. D., Bartels, D. & Harberd, N. P. (1985) *Nucleic Acids Res.* **13**, 6833–6846
8. Sugiyama, T., Rafalski, A., Peterson, D. & Soll, D. (1985) *Nucleic Acids Res.* **13**, 8729–8737
9. Halford, N. G., Forde, J., Anderson, O. D., Greene, F. C. & Shewry, P. R. (1987) *Theor. Appl. Genet.*, in the press
10. Tatham, A. S., Shewry, P. R. & Mifflin, B. J. (1985) *FEBS Lett.* **177**, 205–208
11. Tatham, A. S., Mifflin, B. J. & Shewry, P. R. (1985) *Cereal Chem.* **62**, 405–412
12. Field, J. M., Shewry, P. R. & Mifflin, B. J. (1983) *J. Sci. Food Agric.* **34**, 370–377
13. Danno, G., Kanazawa, K. & Natake, M. (1978) *Agric. Biol. Chem.* **42**, 11–16
14. Friedman, M., Krull, L. H. & Cavins, J. F. (1970) *J. Biol. Chem.* **245**, 3868–3871
15. Bunce, N., White, R. P. & Shewry, P. R. (1985) *J. Cereal Sci.* **3**, 131–142
16. Field, J. M., Tatham, A. S., Baker, A. & Shewry, P. R. (1986) *FEBS Lett.* **200**, 76–80
17. Beaven, G. H. & Holiday, E. R. (1952) *Adv. Protein Chem.* **7**, 319–385
18. Zamyatnin, A. A. (1972) *Prog. Biophys. Mol. Biol.* **24**, 109–123
19. Tanford, C. (1965) *Physical Chemistry of Macromolecules*, Wiley, New York
20. Payne, P. I., Holt, L. M. & Law, C. N. (1981) *Theor. Appl. Genet.* **60**, 229–236
21. Simha, R. (1940) *J. Phys. Chem.* **44**, 25–34
22. Mehl, J. W., Oncley, J. L. & Simha, R. (1940) *Science* **12**, 132–133
23. Creeth, J. M. & Knight, C. G. (1965) *Biochim. Biophys. Acta* **102**, 549–558
24. Rowe, A. J. (1977) *Biopolymers* **16**, 2595–2611
25. Tanford, C., Kawahara, K. & Lapanje, S. (1967) *J. Am. Chem. Soc.* **89**, 729–736
26. Dearbourn, D. G. & Wetlaufer, D. B. (1970) *Biochem. Biophys. Res. Commun.* **39**, 314–320
27. Strickland, E. (1974) *CRC Crit. Rev. Biochem.* **2**, 113–175
28. Tatham, A. S., Drake, A. F. & Shewry, P. R. (1985) *Biochem. J.* **226**, 557–562
29. Manavalan, P. & Johnson, W. C., Jr. (1983) *Nature (London)* **305**, 831–832
30. Arakawa, T. & Goddette, D. (1985) *Arch. Biochem. Biophys.* **240**, 21–32
31. Wilkinson, K. D. & Mayer, A. N. (1986) *Arch. Biochem. Biophys.* **250**, 390–399
32. Herskowitz, T. T., Gadegbeku, B. & Jaillet, H. (1970) *J. Biol. Chem.* **245**, 2588–2598
33. Chen, Y. H., Yang, J. T. & Chan, K. H. (1974) *Biochemistry* **13**, 3350–3355
34. Chen, Y. H., Yang, J. T. & Martinez, H. M. (1972) *Biochemistry* **11**, 4120–4126
35. Venkatachalam, C. M. & Urry, D. W. (1981) *Macromolecules* **14**, 1225–1229
36. Field, J. M., Shewry, P. R., Burgess, S. R., Forde, J., Parmar, S. & Mifflin, B. J. (1983) *J. Cereal Sci.* **1**, 33–41
37. Tatham, A. S., Shewry, P. R. & Belton, P. S. (1985) *Biochem. J.* **232**, 617–620
38. Shewry, P. R., Field, J. M. & Tatham, A. S. (1987) in *Cereals in a European Context* (Morton, D., ed.) (Proc. EFFoST Meeting, Bournemouth, July 1986), pp. 421–437, Ellis Horwood Ltd., Chichester

Received 18 March 1987/28 May 1987; accepted 1 July 1987